Title
The Impact of Local Media on Household Investment Decisions

Author
Cho, Yeon Sik

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

The Impact of Local Media on Household Investment Decisions

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Management

by

Yeon Sik Cho

2017

Abstract of the Dissertation

The Impact of Local Media on Household Investment Decisions

by

Yeon Sik Cho
Doctor of Philosophy in Management
University of California, Los Angeles, 2017
Professor Avanidhar Subrahmanyam, Co-Chair
Professor Stuart A. Gabriel, Co-Chair

Information is costly, and agents have different information-processing abilities. Households have a limited capacity to actively seek out information, and it is therefore likely that they are reliant on local media for information. In this dissertation, I explore the effect of local media on households' investment decisions in the context of the two asset classes most widely participated in: housing and equity. In the first part of the dissertation, I document the potential causal impact of local media on households' housing investment decisions, which in aggregate leads to a predictive relationship between local media and subsequent housing returns. In the second part of the dissertation, I document the potential causal impact of local media on households' equity trading behavior.

The dissertation of Yeon Sik Cho is approved.

Simon Adrian Board

Barney P Hartman-Glaser

Stuart A. Gabriel, Committee Co-Chair

Avanidhar Subrahmanyam, Committee Co-Chair

University of California, Los Angeles

2017

To Mom, Dad, In Young,
family and friends.
Thank you for all your support

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

ACKNOWLEDGMENTS

I am greatly indebted to my advisors Simon Board, Stuart Gabriel (co-chair), Barney Hartman-Glaser, and Avanidhar Subrahmanyam (co-chair) for their guidance and support. I am thankful to Darren Aiello, Chady Gemayel, Nimesh Patel, Geoff Zheng for their helpful comments; Chandler Lutz for FOMC meeting dates as an instrumental variable; participants at the UCLA Anderson Student Seminar for their feedback.

VITA

2008-2012	B.A. Economics and Mathematics, Cornell University, Ithaca, New York, *Magna Cum Laude*
2012-2016	Anderson Fellowship, University of California at Los Angeles

CHAPTER 1

Overview into Media in Finance Literature

Information is costly, and agents have different information-processing abilities. As one of the main channels through which information propagates, the media play an important role in not only the financial markets, but also society in general. Households have a limited capacity to actively seek out information, and therefore media, especially local media, are likely to have a great impact on households' information set and ultimately their decision making. In this dissertation, I explore the effect of local media, the most accessible form of media, in the context of the two financial markets in which households most widely participate: the housing and stock markets. Before discussing the local media's impact on the financial markets, I briefly review the genealogy of media studies. I conclude this section by discussing the current state of media research in the finance literature and my contribution to that literature.

Media have been used as a tool for mass communication since perhaps the dawn of human society. Hence, it is not surprising that the literature on media studies has a long history. Ouellette (2013) suggests that the earliest roots of current media studies, an interdisciplinary field encompassing a wide range of fields from humanities to social science, can be traced to the 1920s. The first studies of media emerged as industrialization led to "new" forms of mass communication, such as movies, radio, and television. These early studies were mostly concerned with criticizing the negative effects of these new forms of media. For example, the Payne Fund Studies (late 1920s to early 1930s) analyzed the effects of movies on the behavior of children and adolescents.

Perhaps the earliest notable theory in media studies is functionalism, which is largely influenced by the works of Paul Lazarsfeld. Functionalism describes the first-order effects of the media, which include the media's tendency to raise awareness of specific issues and the media's tendency to change what is considered "normal". An example of an influential study during this period is Lazarsfeld et al. (1944), which is widely known as "The People's Choice." Lazarsfeld et al. (1944) finds that the media reinforce the already-held beliefs of viewers and rarely alter their beliefs.

The first boom in media studies occurred in the 1970s, when the number of those studies grew exponentially and explored numerous effects of the media on human behaviors (Roberts and Bachen, 1981). Several theories during this period are worth noting. The two-step flow model adopted by Katz and Lazarsfeld (1966) suggests that the media affect leaders, who then affect individuals. Cultivation theory as adopted by by Gerbner and Gross (1976) suggests that the effects of TV are initially small but accumulate over time. Uses and gratification theory as adopted by Katz et al. (1973) suggests that viewers also influence the media because different media sources compete to meet viewers' needs. McLuhan (1965) is notable for suggesting that different forms of media can have different effects even if their informational content is identical.

Given the media's close relationship with politics and public policy, which in turn is closely related to the economy, it is difficult to identify when the media in the economics literature started to be distinguished from the media in other media studies. One of the earliest media studies, Lazarsfeld et al. (1944), can be considered to fall within the economics realm. One distinguishing feature of media studies within the economics realm may be that these studies are focused on news media (as opposed to other forms of media such as movies and TV shows). I henceforth refer to these studies as "economics media studies."

There are too many economics media studies to fully discuss here; hence, I discuss two major topics addressed by the literature[1]. One of the most widely debated topics in economics media studies is agenda setting—that is, the media's ability to affect policy priorities. One of the earliest studies of this subject, Funkhouser (1973), uses simple chi-square analysis to argue that an agenda-

[1] For a more complete list of topics explored by the media studies literature, see Roberts and Bachen (1981)

setting effect may exist. Later studies use various potentially causal inference methods such as lagged correlation analysis (Tipton et al., 1975; Sohn, 1978) and field experiments (Cook et al., 1983; Protess et al., 1985). Despite these efforts, however, there is not yet a consensus on the extent of the agenda setting effect.

A topic explored by economics media studies more recently is the factors that affect media contents. Contrary to typical generalizations (i.e., in the agenda-setting literature), media content is rarely exogenous. Although this theory is not new, empirical studies gained momentum only by the 1990s. One of the earliest empirical studies to test this hypothesis is Glasser et al. (1989), which argues that the ownership of a media company affects the content of the media. Recent literature focuses on other factors that can influence media content (also known as media slant), such as advertising pressure (Reuter and Zitzewitz, 2006) and competition for audience attention (Baron, 2005; Gentzkow and Shapiro, 2006; Mullainathan and Shleifer, 2008).

The finance literature's adoption of media as a subject has been rather slow. Studies of the role of media on the financial market can be traced back to the 1990s. The relatively slow adoption of media as a subfield by the finance literature can perhaps be attributed to the strong assumptions made under the traditional asset pricing models. These models typically assume that agents (individuals and institutions) are fully rational and that there is no informational friction, which means the media can play no role in the financial market.

Liu et al. (1990), Klibanoff et al. (1998), and Huberman and Regev (2001) are among the earliest studies investigating the media's role in the financial markets. Liu et al. (1990) is not directly interested in the role of the media, but its use of the *Wall Street Journal* to measure the information provided by investment advisory agencies makes the study relevant to this literature. In contrast, Klibanoff et al. (1998) hypothesizes that the media play a facilitating role in information transmission. Huberman and Regev (2001) documents through a case study of a pharmaceutical firm, EntreMed, that the media can have a significant impact on a company's stock price. These early studies lack causal inference tools, however, and can provide only suggestive evidence that the media matter in the financial markets.

Numerous studies explore various aspects of the media's effects on financial markets. The evidence appears to suggest that the media play an indispensable role in financial markets. Perhaps the most straightforward evidence is the media's power to predict market prices. The most influential study on this subject is Tetlock (2007), which uses text mining on a *Wall Street Journal* column to build a predictive index of the Dow Jones Index. As a pioneering paper that introduced text mining to the finance literature, Tetlock (2007) spawned numerous follow-up papers including Tetlock et al. (2008), which extends to different news outlets; Engelberg (2008), which considers soft information from earnings announcements; and Tetlock (2011), which finds that even stale information can affect prices.

Similarly, Da et al. (2011) is notable for documenting the predictive power of a new medium, Google Searches, on stock prices. Although not directly predictive, Fang and Peress (2009) provides further evidence. Through binning, Fang and Peress (2009) finds that a portfolio of stocks with high media coverage tends to have lower returns than a portfolio of stocks with low media coverage. Furthermore, the media's predictive power need not be limited to stock prices. Soo (2013) finds that a sentiment index on the media similar to the one used by Tetlock (2007) has predictive power for local house prices. Chauvet et al. (2016) finds that Google Searches have predictive power for housing derivatives such as the ABX market.

Despite the widespread use of the phrase "impact of media" in the finance literature, evidence for a causal impact of media on market prices is limited. For example, Dyck and Zingales (2003) shows that stocks react more to earnings announcements emphasized by the media, which is evidence of the media impacting stock prices. However, without a plausibly exogenous source of variation, endogeneity issues preclude causal inferences. To my knowledge, Engelberg and Parsons (2011) is the only study to provide causal evidence of the media affecting market prices.

Engelberg and Parsons (2011) takes the view that individual investors learn about earnings events through local media or through major newspapers in the region. Engelberg and Parsons

(2011) argues that earnings event related articles that report on particular earnings events affect individual investors trading volumes. Using plausibly exogenous differences in the timing of earnings event related articles (time scale of 1-2 days), which may be due to extreme weather that delays newspaper delivery or print deadline differences, Engelberg and Parsons (2011) shows that trade volume from a particular newspaper's region peaks on the day of earnings event related article delivery. Chapter 3, which analyzes the effect of local media on local individual investors trades, takes a similar approach to Engelberg and Parsons (2011), but with a slightly different identification mechanism. Namely, I focus on all earnings event related articles, which may remind readers of particular earnings events (prior to the event) or inform readers on the content of the earnings event (after the event), and their effect on individual investors trades.

The channels through which the media can affect financial markets are numerous. Perhaps the most straightforward channel is salience. There are thousands of stocks on the US stock market, and, as with the agenda-setting hypothesis, the media decide what stocks to give attention to. For example, Fang et al. (2014) finds that mutual fund managers are more likely to buy media-covered stocks. Similarly, Bhattacharya et al. (2009) argues that media hype played a role in the internet IPO bubble of 1996-2000. It is also possible that the media affect financial markets by facilitating information transmission (Engelberg, 2008; Bushee et al., 2010; Engelberg and Parsons, 2011).

Several other media roles are worth mentioning. First is the media as related to corporate governance, which indirectly affects stock prices. Dyck and Zingales (2002) documents a correlation between media coverage of corporate governance violations and subsequent reversals of these violations. Similarly, Miller (2006) finds that media coverage deters accounting fraud. Second is the media as a management tool. Gurun (2010) finds that firms with board members with media expertise tend to "manipulate" media coverage to affect stock prices. Similarly, Gurun and Butler (2012) shows that local media tend to use fewer negative words toward local firms, potentially because of advertisement revenue issues.

Despite the long history of media studies, much work remains to be done. One topic requiring attention is the causal effect of media. Media can have profound effects on a wide range of outcomes, such as financial market prices, policy outcomes, and individuals' expectations. Hence, to maximize the positive effects of media, the extent of the causal impact of media must be analyzed, and this dissertation contributes to this field. Chapters 2 and 3 argue that the media affect individual investors' trading decisions.

Another area that requires attention is the effects of different forms of media. With few exceptions, studies of media and financial markets use major newspapers as their proxy for media. However, major newspapers represent only a small fraction of the media universe, and early media studies attest to the different effects that different forms of media can have. In focusing on the difference between major and local newspapers, this dissertation provides preliminary evidence that non-major newspapers can have a material impact.

This dissertation is structured as follows. Chapter 2 presents evidence that local media play an information-providing role and thus affect short-term housing price dynamics. Chapter 3 investigates one facet of the relationship between local media and the markets: the incremental stock market influence of local media over nationwide media. Chapters 4 and 5 provide additional analysis for robustness.

CHAPTER 2

Housing Market

2.1 Introduction

The housing market is a significant part of the economy. Over two thirds of households[1] in the US own homes, and despite the benefits of diversification, houses remain an overwhelming portion of households' portfolios (Tracy et al., 1999; Nakajima et al., 2011). As such, understanding the dynamics of house price movements can have significant welfare implications. The goal of this chapter is to test whether media can substantially affect the movement of housing prices in the short term. I use housing prices from counties across the US to test the effects of housing-related news in the media on housing market returns in a panel setting.

Evidence from behavioral economics and other streams of literature indicate that people can be affected by the media. Examples include the impact of newspaper articles on stock prices (Huberman and Regev, 2001; Liu et al., 1990), the impact of mass media on politics (Graber and Dunaway, 2014), and the impact of media on violence in society (Anderson and Bushman, 2002). In the finance literature, while behavioral biases have been documented among institutional investors (Bailey et al., 2011), it is more accepted that households exhibit behavioral biases and are influenced by the media (Barber and Odean, 2011)[2].

[1] In this chapter, I use "individuals" and "households" to mean the same thing.

[2] The media may affect households' behavior by providing them with information (a rational story) or through other means (a behavioral story). I assume that if a household's behavior changes after receiving a piece of news, it is because of the news and not because of how it is presented. However, both mechanisms lead to the same empirical

In contrast to other financial markets, the residential housing market is dominated by households. According to a 2013 Federal Reserve report, Atlanta had the highest institutional investor activity in 2012 among Metropolitan Statistical Areas (MSAs) at a mere 16%[3], which suggests that the remaining 84% of the activity can be attributed to households. Given households' responsiveness to the media and their dominance in the housing market, I hypothesize that the media play an important role in households' home buying decisions and, in turn, the housing market.

The model is as follows. Due to the high cost of information processing and limited attention, households can process only a fixed amount of news, which affects their expectations about the future housing market. For instance, if each piece of news is assumed to contain positive or negative news with equal probability and news articles are sampled without replacement, the probability of obtaining a disproportional (mostly positive or mostly negative) combined signal increases as the number of news articles increases. Increased disproportional combined signal creates more heterogeneous beliefs about the housing market and increases the number of households that want to buy or sell. However, because of shorting constraints[4], only households that currently own a home are able to react to a negative signal, whereas both homeowners (investing in a second home or moving into a bigger home) and non-owners can react to a positive signal. This asymmetry results in a positive relationship between housing news intensity and the subsequent housing return. Further details of the model are outlined in Section 2.2.

The role of shorting constraints in the presence of heterogeneous beliefs has received a great deal of attention in the finance literature (Harrison and Kreps, 1978; Diamond and Verrecchia, 1987). Although earlier studies focus on the equity market, recent studies highlight shorting constraints as an important feature in the housing market (Nathanson and Zwick, 2014; Burnside et al.,

---

results, and hence I do not distinguish between these stories.

[3]For more details, please see https://www.federalreserve.gov/econresdata/notes/feds-notes/2013/business-investor-activity-in-the-single-family-housing-market-20131205.html

[4]There are several ways to short the housing market, including shorting the stocks of real-estate companies such as REITS and selling CME Case-Shiller futures. Yet these strategies at best give negative exposure to MSA-wide housing risks for major MSAs, and hence cannot be used to exploit the county-level local effects documented here.

2016; Favara and Song, 2014). Heterogeneous beliefs can originate from various sources, such as the different weights used on past realization extrapolation (Granziera and Kozicki, 2015) and different experiences in the housing market (Kuchler et al., 2015). Similarly, Bailey et al. (2016) finds that social media play an important role in the housing market. My findings are consistent with these studies, and I argue that the media are a source of news shocks that create heterogeneous beliefs among households.

Due to the importance of the housing market as a research question, the relevant literature is too diverse to fully represent here. Hence, only the studies deemed most relevant to this study are highlighted. One of the more relevant streams of literature about the housing market is time-series dynamics. House prices are known to exhibit short-term momentum and long-term reversal. Recent attempts to explain these phenomena include concave demand functions[5], which motivates sellers to move prices slowly (Guren, 2014) and leads to the existence of momentum traders (Piazzesi and Schneider, 2009).

More germane to this study may be the literature on cross-sectional dynamics. These studies consider why housing prices in certain areas move differently from prices in other areas. One possible approach is risk-based stories, or cross-sectional asset pricing. Cross-sectional asset pricing, which gained popularity through a series of papers by Fama and French (Fama and French, 1992, 1993), argues that differences in exposure to risk factors can explain cross-sectional differences. Cross-sectional housing asset pricing studies, which are relatively more recent (Cannon et al., 2006; Case et al., 2011), explain a significant proportion of heterogeneity in housing returns across US MSAs. However, as suggested by Ling et al. (2015), there are limits to the extent to which risk-based stories using publicly available data can explain cross-sectional house price dynamics. For example, Soo (2013) shows that fundamentals that explain nearly 70% of house price variations in 1987-2000 explain only 10% of the variation in the 2000s.

[5] Guren (2014) defines a concave demand function as a demand function that decreases more with a small price increase than it increases with the same sized price decrease.

Other approaches to cross-sectional dynamics focus on the specific characteristics of different regions. For example, Glaeser et al. (2005) argues that differences in the difficulty of obtaining building permits explain why certain cities experience high home price appreciation. Gyourko et al. (2013) argues that a limited housing supply combined with differences in income distribution explains house price differences across regions. Although related to this literature, this study is unique in that it focuses on short-term dynamics. I offer evidence of a short-term (less than a year) correlation between the media and housing returns, which is consistent with an informational story. Whether this informational story has a long-term effect is an interesting question, but it is beyond the scope of this study.

Another body of relevant literature concerns the role of media, which has received a great deal of attention in academia. Even within the finance literature, many studies document the predictive power of media. Perhaps the most influential such work is Tetlock (2007), which uses text mining to analyze a *Wall Street Journal* column, “Abreast of the Market.” Tetlock argues that a sentiment index created from the column can predict the Dow Jones Index. In line with Tetlock (2007), the predominant interpretation of media’s role is the reflection of market sentiment, and numerous studies document the impact of sentiment on the stock market (Edmans et al., 2007; Baker et al., 2012; Brown and Cliff, 2005).

To the best of my knowledge, the only study that links the media to the housing market is Soo (2013). Due to the similarities between Soo (2013) and this study, it is worth mentioning this chapter’s contribution over Soo. Soo analyzes housing articles from a major newspaper in each of 20 MSAs using text mining to build a sentiment index. This approach stems from the assumption that newspaper articles contain information about the housing market that is beyond fundamentals. Hence, Soo’s analysis centers on ruling out latent fundamentals as the driving factor of the predictive relationship between sentiment and housing return. Soo (2013) finds that the sentiment index is useful for forecasting housing prices at a horizon of one to three years.

This study takes a fundamentally different view of the relationship between the housing media and housing market. This study argues that the media have an influential role on households' home-buying decisions, which is reflected in the short-term housing price dynamics. This is different from Soo's view that the media's content reflects unobserved current market conditions, which are correlated with long-term house price dynamics. To document the short-term relationship between housing price dynamics and media news intensity, this study uses county-level housing prices and newspaper data from nearly 300 counties to run a time-series panel regression. Although Soo controls for changes in fundamentals through the inclusion of nearly a dozen controls, this study does so through MSA interacted with time fixed effects, which is possible due to the county-level sample. Furthermore, I provide evidence that the media have a causal role in households' home-buying decisions, which may affect housing returns through the mechanisms described in Section 2.2. As evidenced by numerous studies of media stories described in Chapter 1, the media effect can have numerous channels. I argue that this study documents a different (short-term informational) channel from Soo's (long-term).

My hypothesis is that the media affect households' home-buying decisions, which affects short-term housing price dynamics. Hence, a lagged increase in the number of housing news articles is positively correlated with housing returns. I first show this correlation through a county-level panel time-series regression using data from 269 counties. I use newspaper articles as a proxy for the media, as is standard in the literature. I find that the increase[6] in the media's housing content leads to a 8-10 bp (annualized) increase in local housing returns after 4-6 months, and this jump in housing returns mostly reverts during the next 6 months. Then, using Federal Open Market Committee (FOMC) meetings as an instrumental variable (IV), I show that local media have a potentially causal effect on households' home-buying decisions. As an extension, I classify articles into positive and negative housing news and explore the differential effects of positive and negative media in the appendix.

[6]The coverage level of housing news varies greatly across counties in my data sample. As I observe housing news only intermittently in many counties, my independent variable is the direction of the housing news level change rather than the magnitude change.

Due to high transaction costs and the lack of a shorting mechanism, it is unlikely that the predictive relationship between media and housing prices documented in this study will lead to an arbitrage opportunity. Furthermore, as the impact of media analyzed in this study is short-lived (it decays in 1 year), the lasting impact on long-term asset prices may be minimal. Nevertheless, I argue that this study is meaningful for establishing the media's substantive role in the housing market, which can facilitate an understanding of the housing market that can lead to policy implications.

This chapter proceeds as follows. In Section 2.2, I briefly describe the mechanism through which I hypothesize how the media affects housing prices. In Section 2.3, I describe the data sources and analysis sample. In Section 2.4, I document the short-term predictive power of local media on local housing returns. In Section 2.5, I discuss the causal relationship between housing media and households' home-buying interest. Section 2.6 concludes the chapter.

2.2 Model

In this section, I present a simple model that illustrates the mechanism through which the media affects the housing market. In this model, households obtain different news about the housing market because of inefficiencies in processing information. Furthermore, the likelihood of a disproportionate combined signal increases as the amount of news increases because households can process only a small proportion of the news. Although both homeowners and non-owners receive news shocks, due to shorting constraints, only current homeowners can react to negative news shocks. Hence, an increase in housing articles increases housing demand (from both current owners wanting a second home and non-owners) more than the housing supply (only from current owners)[7]. The number of housing articles is positively correlated with subsequent housing

[7]In the real world, the housing supply may vary with external factors such as new construction. However, external housing supply factors tend to be more inelastic. For example, it takes on average 6 months for new homes to be built according to the *Wall Street Journal* (`http://www.wsj.com/articles/`

prices and returns. The essence of the mechanism—namely, shorting constraints and heterogeneous beliefs—is similar to mechanisms explored by other studies (Favara and Song, 2014). The source of heterogeneous beliefs is unique to this chapter.

Consider a two-period economy ($t = 0, 1, 2$) with only two investment instruments: zero-coupon bonds and housing. This assumption is motivated by the observation that while a large number of households are homeowners, only a small portion of households own stocks. In this economy, there is a mass 1 of agents whose goal is to maximize their wealth, and $0 < \mu < 1$ of these agents currently own one house. Assume there are many rental facilities that provide rental housing at a rate of c^r each period for agents without a house. There are no transaction costs associated with buying or selling a house, and each house is assumed to be identical and hence to have the same price. Each agent is initially endowed with wealth W, which is sufficient to buy a house. An agent without a house can decide whether to buy a house, and an agent with a house can decide to sell his, do nothing, or buy an additional house to rent out.

The goal of each agent is to maximize his wealth at $t = 2$. At $t = 0$, each agent obtains news about the housing market, $\eta \sim N(m, n^2)$, which affects his expectation of home prices at $t = 2$, where n represents the number of housing news articles in the media. Each agent knows the setup of the economy but does not know the realization of η for other agents.

The rationale for more housing news leading to more belief dispersion is as follows. Assuming that additional news is equally likely to be positive or negative, if agents have limited attention and they sample news from the media without replacement, it is more likely that agents will receive disproportionate combined signals when there is more news. For example, when there are three pieces of positive news and three pieces of negative news, digesting only two of them yields a $1/5 = 20\%$ chance of getting strong news (i.e. both are positive news) about the housing market. However, when there are six pieces of positive news and six pieces of negative news, this chance

---

`average-time-to-build-a-house-6-months-1420652311`). I focus on the short-term effect of local media on housing prices and ignore external factors.

increases to $5/22 \approx 22.7\%$. If only positive or negative news increases, the actual distribution can be skewed; however, for simplicity, I assume that a disproportionate increase in positive or negative news only affects the distribution of η through m and n.

Based on the news at $t = 0$, each agent makes a decision to buy or sell a house at $t = 1$. Each agent can deduce the market clearing price P_1 based on the distribution of η; hence, the optimization problem for non-owners is

$$\max \begin{cases} W - c^r - P_1 - c^o + \mathbb{E}_1[P_2] & \text{: Buy} \\ W - 2c^r & \text{: Not Buy} \end{cases}$$

If the agent decides to buy a home in period 1, he or she pays rent of c^r at $t = 0$, as he or she is without a home, and pays the price of the home P_1 and a maintenance cost of c^o at $t = 1$. At $t = 2$, he or she gets the price of the home at the time. If he or she decides not to buy a home, then his or her final wealth is his or her initial wealth less two periods' rent.

For agents endowed with homes, the optimization problem is

$$\max \begin{cases} W - c^o - P_1 - 2c^o + c^r + 2\mathbb{E}_1[P_2] & \text{: Buy} \\ W - 2c^o + \mathbb{E}_1[P_2] & \text{: Do Nothing} \\ W - c^o + P_1 - c^r & \text{: Sell} \end{cases}$$

Similarly as before, if an agent with a home decides to buy another, then at $t = 0$, he or she pays the maintenance cost; at $t = 1$, he or she pays the price of the home and the maintenance costs of two homes and gains rent from one of them; and at $t = 2$, he or she gets the price of the home at the time.

As discussed previously, each agent's future expectation of the home price is determined by a fundamental value and news shock. Assume that the fundamental value of a home v is determined by economic factors that do not change in the timeframe of the model. Hence,

$$\mathbb{E}_1[P_2] = v + \eta, \eta \sim N(m, n^2)$$

Lastly, assume that because of unforeseen activities such as job relocation, there is also a probability λ that an agent has to relocate. The shock will come at $t = 0$, and for agents with homes, this will mean that they have to sell at $t = 1$. This relocation shock is independent from the news shock. The relocation shock is introduced to model the relatively inelastic supply but is not crucial for the result. Although not modeled here, the inelastic supply can also be attributed to households' loss aversion, as discussed in (Genesove and Mayer, 2001).

Consider the perspective of an agent without a house. He or she will decide to buy a house if

$$\begin{aligned} W - c^r - P_1 - c^o + \mathbb{E}_1[P_2] > W - 2c^r \\ \implies \eta > P_1 - c^r + c^o - v \end{aligned}$$

The problem for an agent with a house is similar. He or she will decide to buy an additional house if $\eta > P_1 - c^r + c^o - v_0$, sell his or her house if $\eta < P_1 - c^r + c^o - v_0$, and be indifferent if equality holds.

Given the normal distribution of the news shock η and that the λ of individuals with a high news shock will not be able to participate because of relocation, the demand function as a function of the $t = 1$ home price P_1 is

$$D(P_1) = (1 - \lambda) \left(1 - \Phi\left(\frac{P_1 - c^r + c^o - v - m}{n}\right)\right)$$

The supply function can be similarly computed.

$$S(P_1) = \underbrace{\mu\lambda}_{\text{Supply from Relocation Shocked Owners}} + \underbrace{\mu(1-\lambda)\Phi(A)}_{\text{Supply from Non-Shocked Owners}}$$

where, $A = (P_1 - c^r + c^o - v - m)/n$

Equating the supply and demand yields the housing price P_1 as a function of amount of housing news in the media n.

$$P_1 = n\Phi^{-1}\left(\frac{(1-\lambda)(1+\mu)-\mu}{(1-\lambda)(1+\mu)}\right) + v + c^r - c^0 + m \tag{2.1}$$

Equation 2.1 gives a positive relationship between the amount of housing news and housing prices when $((1-\lambda)(1+\mu)-\mu) \,/\, ((1-\lambda)(1+\mu)) > 1/2$. Using reasonable assumptions about μ and λ, such as the percentage of homeowners $\mu = 0.65$ and relocation rate $\lambda = 0.12$, provides this result[8]. Based on equation 2.1, it can be deduced that housing returns are positively correlated with changes in the housing returns. Namely,

$$\Delta P_1 = \Phi^{-1}\left(\frac{(1-\lambda)(1+\mu)-\mu}{(1-\lambda)(1+\mu)}\right)\Delta n$$

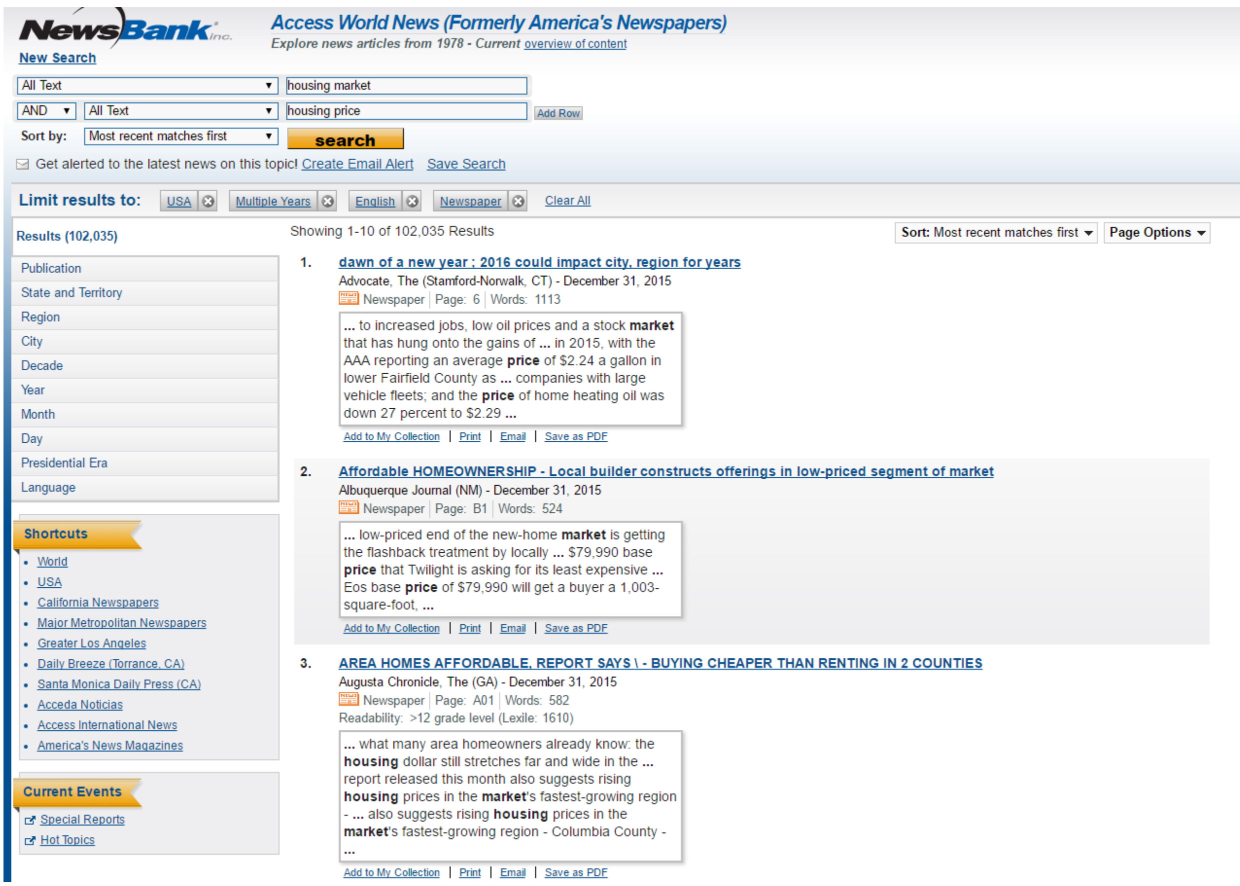
2.3 Data

For this study, I obtain data from multiple sources. Perhaps the most important data in this study are newspaper articles, which are sourced from NewsBank's Access World News (AWN) database, provides news articles for various media types from 1978 onward. Although there are two other major newspaper databases, Factiva and LexisNexis, AWN is the only database that provides com-

[8] According to the US Census, the US mover rate has remained stable at 12% since 2008. Similarly, according to the US Census, homeownership in the US was 62-68% in 2008-2015.

prehensive newspaper articles from local newspapers[9]. AWN has data on 863 US newspapers[10]. It contains all articles from newspapers with the exception of advertisements and articles by syndicated columnists and freelance writers. I focus on newspapers as a proxy for the media, which include radio, television, magazines, etc. Although the AWN does not include a comprehensive set of newspapers in circulation in the US, I assume that the articles included in the database are unbiased in their representation of the local media news flow.

Figure 2.1: NewsBank Screenshot



Notes: This figure shows a screenshot of the NewsBank database

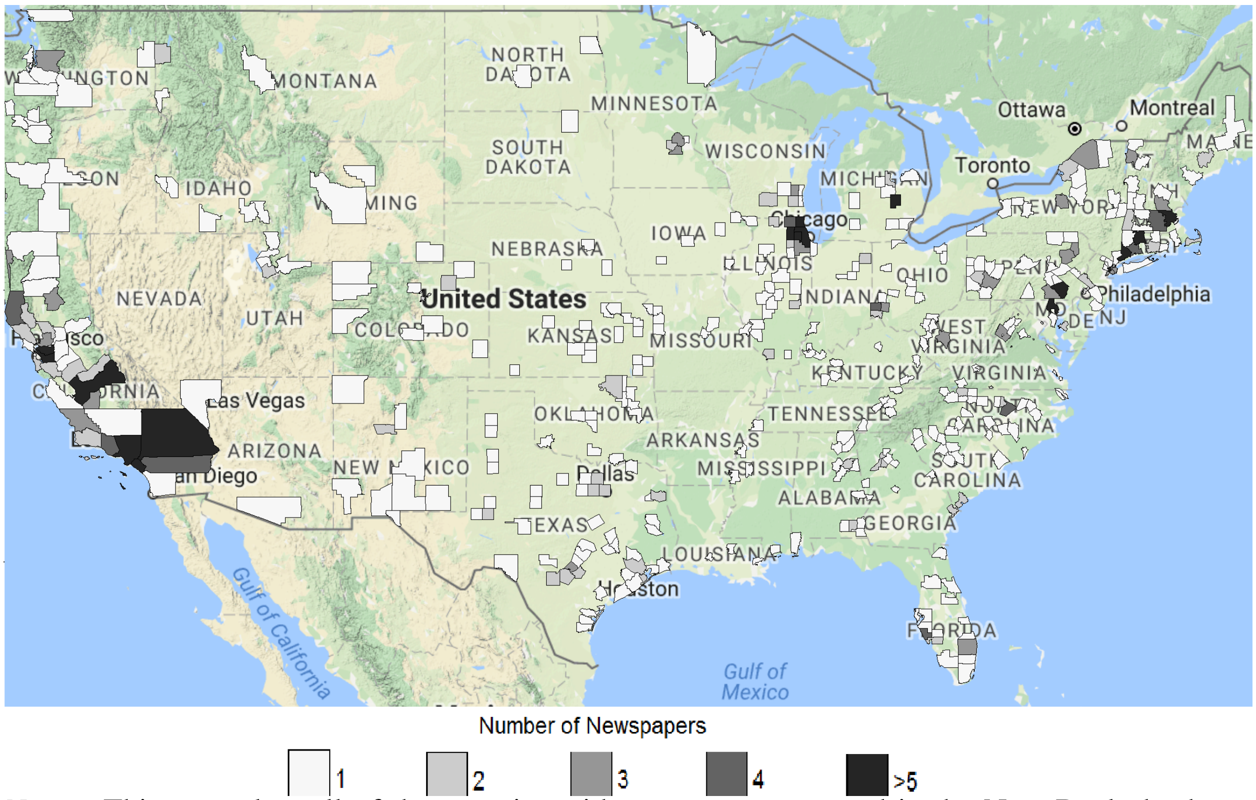
Housing articles are extracted from the database using the keywords "housing market" and

[9] I assume that local housing market-related news is best represented in local newspapers or in newspapers that are headquartered locally. Furthermore, it is likely that households are most receptive to local newspapers. As such, I use local newspapers to represent the media's effect on households.

[10] At least, UCLA's subscription with NewsBank at the time of access contained 863 US local newspapers.

“housing price”. The keywords “housing,” “market,” and “price” all had to be included in either the body or the headline for an article to be included in my sample. I limit the sample to articles from newspapers (i.e., I exclude articles from magazines, journals, and Web-only sources). I also drop articles from non-English-language newspapers. Figure 2.1 shows a screenshot from AWN.

Figure 2.2: Counties Covered by NewsBank



Notes: This map plots all of the counties with newspapers covered in the NewsBank database. The white shade indicates that NewsBank includes only one newspaper from the county, whereas the black shade indicates that NewsBank includes more than five. Not shown in the plot are four counties in Alaska and one county in Hawaii.

I obtain 210,964 articles published between 1998 and 2015 from 744 newspapers. These newspapers come from 447 counties. Figure 2.2 plots the counties represented in the newspapers from AWN during my sample period of 1998-2015. As shown in Figure 2.2, the database covers most major counties throughout the US. While some counties are part of large cities with more than five newspapers, such as Los Angeles, San Francisco, Chicago, Philadelphia, Boston, and New York (shaded black), the vast majority of the counties in the sample have only one newspaper.

AWN provides the town or city in which the newspaper is located. As mentioned previously, I assume that each newspaper provides news that is representative of the local media at the time. As my unit of analysis is the county, I assume each newspaper represents the media of the county. Each town or city is matched to the county via the county or counties listed on the town or city's Wikipedia page. For certain towns or cities located at the borders of multiple counties, I include the newspaper for all of the counties. Hence, some newspapers are counted in multiple counties. When there are different titled newspapers circulated by the same company (e.g., Daily vs. Sunday), I assume them to be the same newspaper.

The housing data are obtained from Zillow, as it is currently the only source of public data on county-level monthly house price indices. The data are updated with a 18-23 day lag every month. Zillow's HPI methodology differs from commonly used HPIs such as S&P Case-Shiller in several ways. Perhaps most importantly, S&P Case-Shiller is based on repeat sales, meaning that only houses that have been sold multiple times are included. The price change between the two or more transactions are aggregated using the weighted average to obtain the S&P Case-Shiller HPI. The Zillow HPI takes a fundamentally different approach in that it uses the Z-estimate, an estimated value of a home based on its proprietary machine learning algorithm.

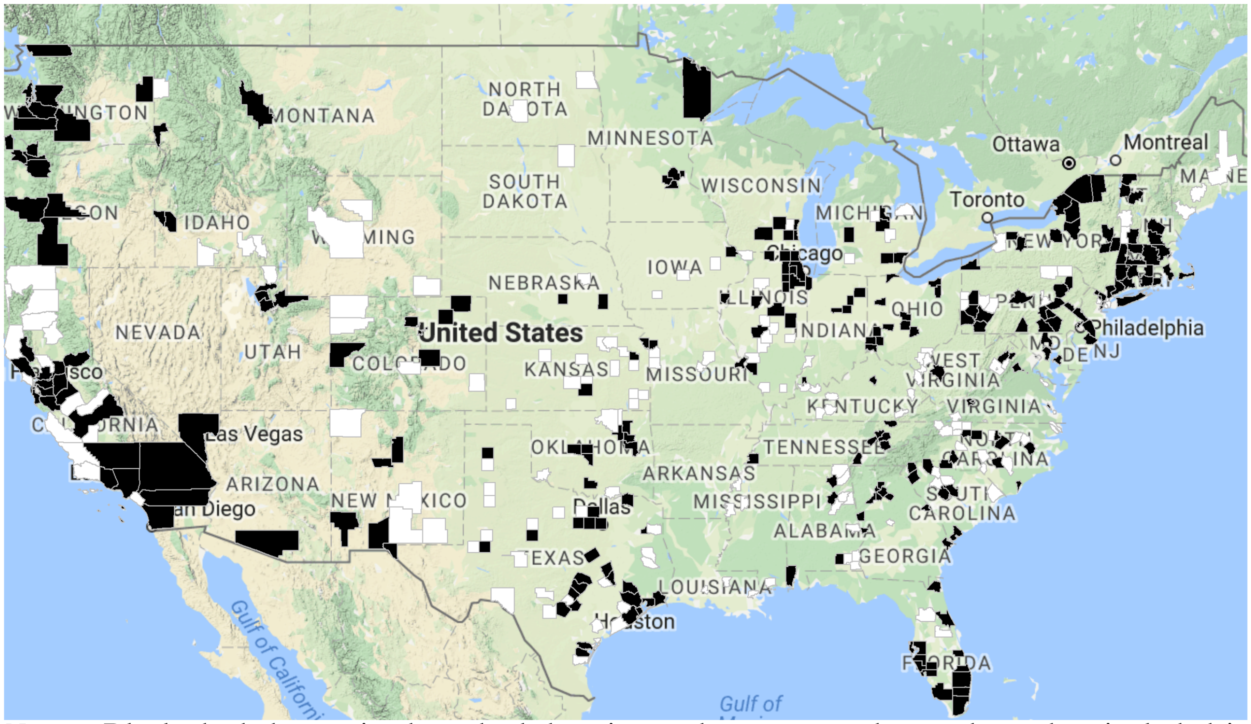
Zillow's Z-estimate uses multiple sources of data, including prior sales, county records, tax assessments, real estate listings, mortgage information, and GIS data. Furthermore, Zillow's website allows homeowners to view the entire history of Z-estimates and to report home improvements, which would otherwise take a long time to appear in official records. The Zillow HPI is the median of the Z-estimates of all of the homes in a region. Thus, the Zillow HPI represents 95% of the US housing stock by market value, while S&P Case-Shiller represents 71%. Other differences include that whereas S&P Case-Shiller uses foreclosed sales data, Zillow does not[11].

Zillow HPI is relatively new compared with S&P Case-Shiller, as the company was founded

[11] Zillow argues that this makes their HPI more accurate because foreclosure sales do not accurately represent health home transactions.

in 2006 and the current Z-estimate valuation model was released in June 2011. Hence, the Zillow HPI is not as established as S&P Case-Shiller in the finance literature. However, Zillow is the largest online real-estate database company and has the 30th highest traffic in the US[12] As such, Z-estimates receive extensive scrutiny and are unlikely to be significantly biased.

Figure 2.3: Counties Analyzed

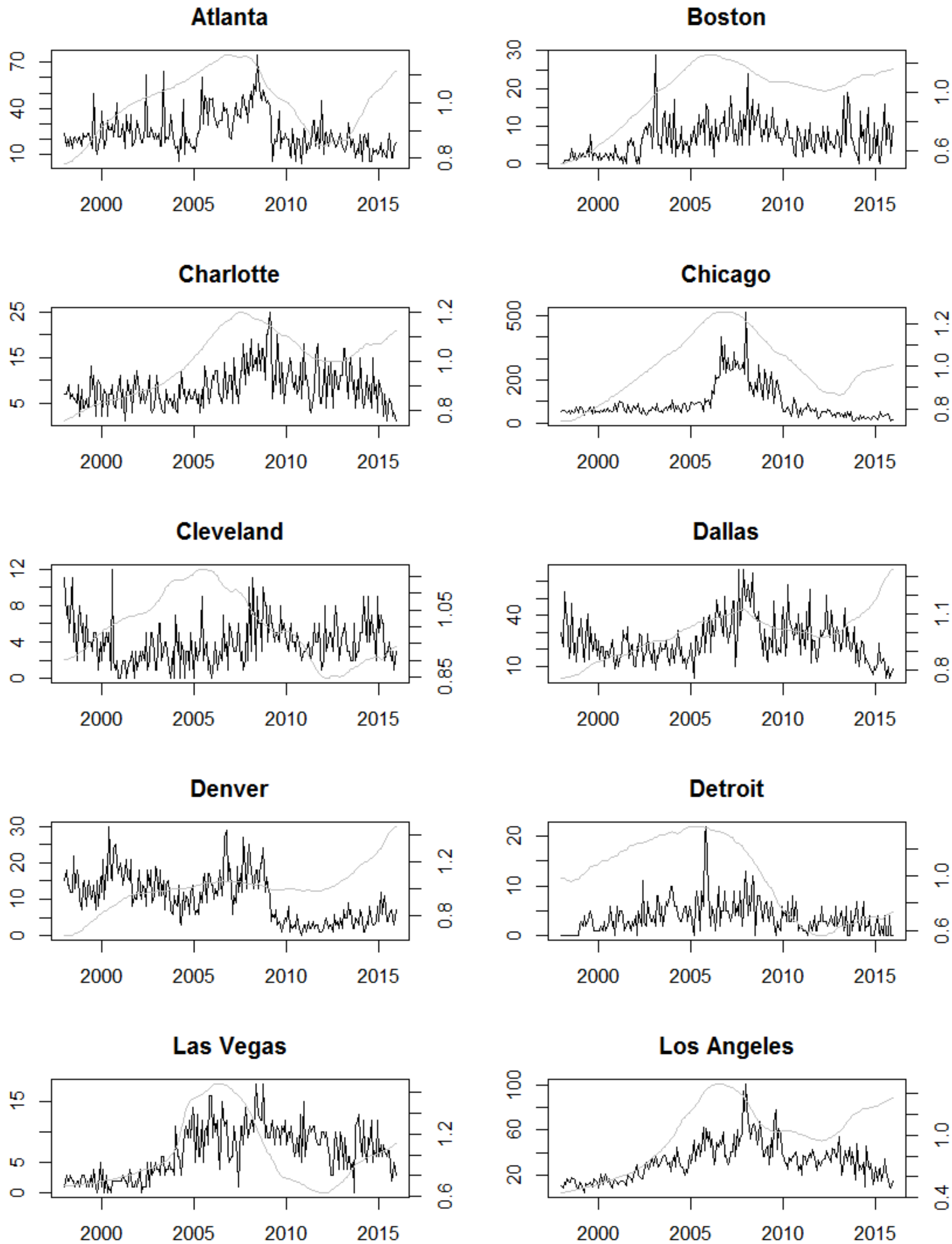


Notes: Black-shaded counties have both housing and newspaper data and are thus included in the analysis. The white-shaded counties are included in NewsBank but lack housing data. Four counties in Alaska and one in Hawaii are also included in the analysis.

Due to the constraints on the Zillow house price data, after matching the housing price data with the newspaper data, I obtain a sample of 269 counties with complete data between January 1998 and December 2015. During the analysis, I drop the January 1998 observations to compute changes in the media's housing content. Figure 2.3 plots the counties with data on both housing prices and media that are included in the analysis. The counties shaded in black are included, while the counties shaded in white are excluded because of a lack of housing price data.

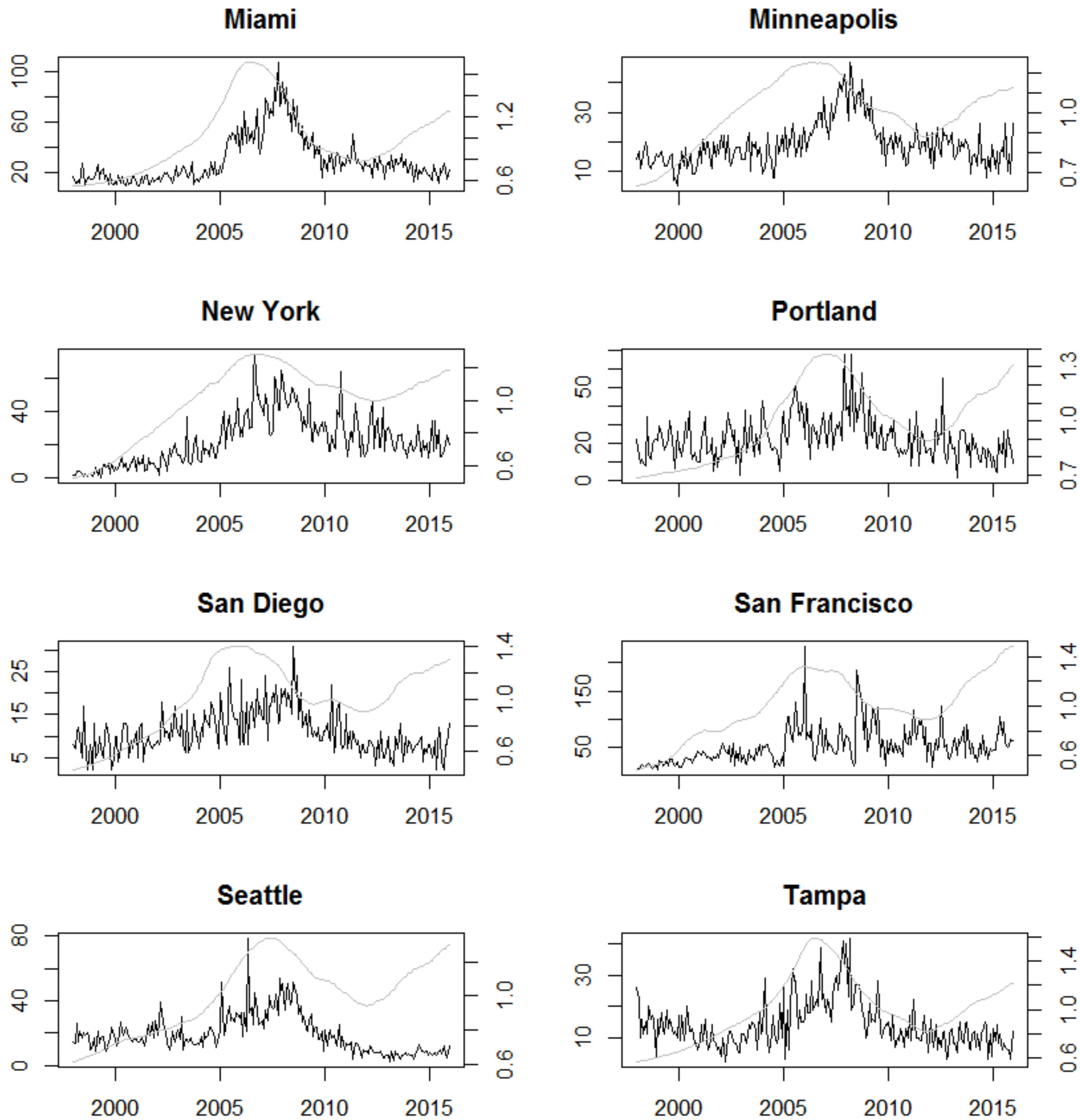
[12] As reported by Similar Web on August 2016.

Figure 2.4a: Housing Articles and Housing Return



Notes: These figures plot the total number of housing articles in each city (black) and the average HPI (grey). The average HPI is computed by averaging the demeaned county-level HPIs. The total number of articles uses the left scale and the average HPI uses the right scale

Figure 2.4b: Housing Articles and Housing Returns



Notes: These figures plot the total number of housing articles in each city (black) and the average HPI (grey). The average HPI is computed by averaging the demeaned county-level HPIs. The total number of articles uses the left scale and the average HPI uses the right scale

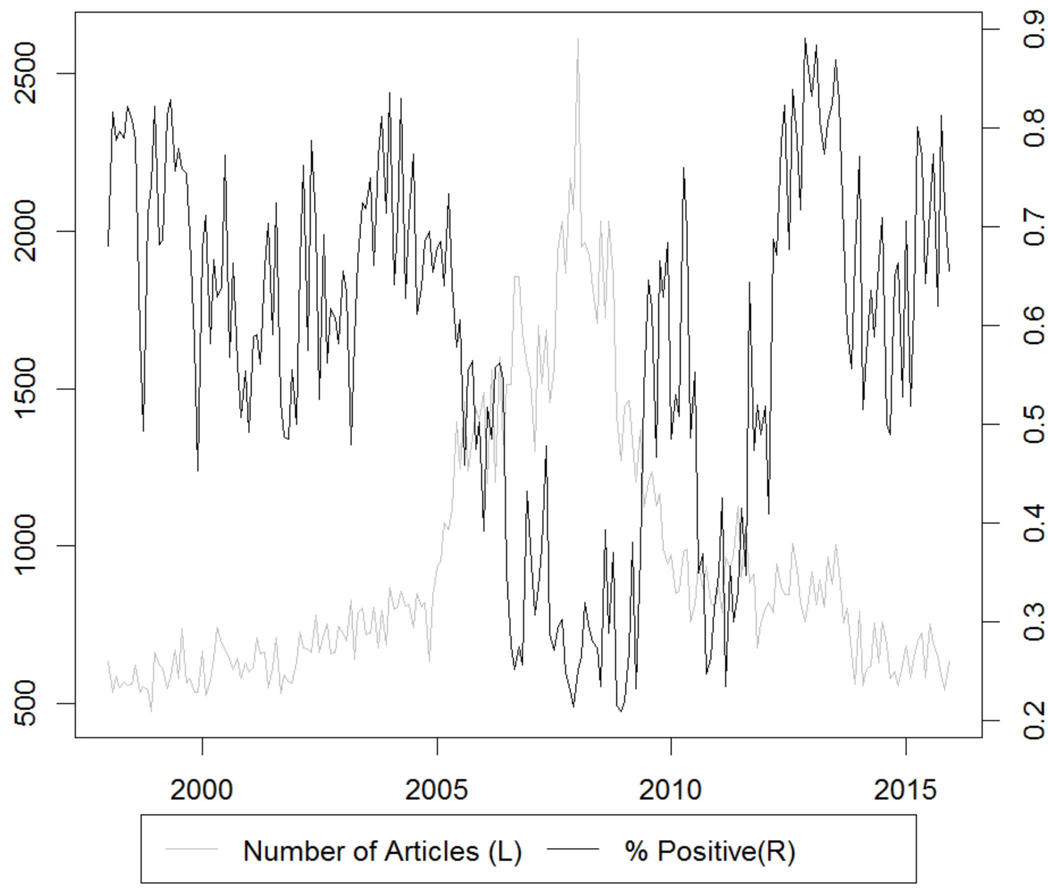
Figures 2.4a and b plot the total number of housing-related news articles over time for different cities and the average HPI of counties that are part of the city. As different counties of the same MSA may have HPIs on different scales, each county's HPI is divided by the whole sample average before averaging. The plots are for 18 major MSAs[13]. In each plot, the black line represents the total number of housing articles (scale on the left). The grey line represents the average HPI for counties in the MSA (scale on the right). There appears to be some contemporaneous movement between the two lines, especially during the crisis period. However, whether one leads the other is not immediately clear from the figures.

Figure 2.5 plots the total number of housing-related news articles across all counties and the proportion of news that is positive. Each housing article is sorted into positive, negative, and unclassified bins through the algorithm described in Section 4.B. The proportion of positive news is computed as the number of positive news articles divided by sum of positive and negative news. In line with assumptions made in Section 2.2, we can see that excluding the housing crisis, the proportion of positive news is relatively stable at around 60%.

Households' home buying interest is proxied by data from Google Trends, a search volume index (SVI) that gives the popularity of search terms in a normalized range between 0 and 100. The search volume of a particular keyword is divided by the absolute number of the search volume and the value 100 is assigned to the largest search volume during a given timeframe. Due to a privacy filter, when the search volume is insufficient, Google Trends reports 0. Google is the largest search engine in the US, accounting for over 66% of all US online searches, and hence has become credibly representative of the US online population. Recent studies use Google Trends to link households' mortgage default concern to mortgage delinquency indicators (Chauvet et al., 2016), home-buying intentions to housing prices (Beracha and Wintoki, 2013), and investor attention to stock prices (Da et al., 2011). The details of the Google Trends data used here are discussed in

[13] These are 20 counties that are part of the S&P Case-Shiller 20-City Composite index. Phoenix and DC are excluded because of a lack of newspaper or housing price data.

Figure 2.5: Average Normalized Articles and Proportion of Positive/Negative Articles



Notes: This figure plots the total number of housing articles in the data (grey) and the proportion of news that is positive (black). The number of articles has the scale on the left, and the proportion of positive articles has the scale on the right (%)

Section 2.5.

Aside from the three data sources, this study derives data from other sources. The complete list of data sources and the descriptions of those not discussed here are listed in Table 2.1.

Table 2.1: Data Sources

Data	Source	Notes
Housing News	NewsBank - “Access World News”	
Housing Price	Zillow	
Home Buying Interest	Google Trends	
National Articles	NewsBank - “Access World News”	Articles from “National” newspapers from “Access World News”
30 Yr Mortgage Rate	St Louis Federal Reserve	
New Constructions	U.S. Census - New Building Permits	I use total number of new units authorized
Mortgage Application	Mortgage Bankers Association - U.S. Purchase Index	Index measuring mortgage application volume. Based on 75% of mortgage applications. The index normalized to March 16th, 1990 level
FOMC Meeting Dates & Forecasts	Bloomberg	

Notes: Table 2.1 summarizes the data used in this study.

2.4 Local Media and Local House Prices

As discussed in Section 2.2, households get a heterogeneous news shock when there is a great deal of housing news due to costly information processing. With the presence of shorting constraints, only households with a positive news shock affect the housing market, creating a positive relationship between the number of housing news articles and housing returns. In this section, I test this hypothesis. I find that counties with higher numbers of housing news articles have housing returns 8-10 annualized basis points higher than other counties, and these appear 4-6 months afterward. This effect is persistent throughout the whole sample but is weakest during the housing crisis. Furthermore, the impulse response function from the panel VAR analysis suggests that this increase persists for 10 months after the increase in the housing content.

2.4.1 Does Housing News Content Predict Housing Returns?

I first use panel time-series regression to test whether housing-related content in the media predicts short-term local housing returns. I use the panel setup to identify the general relationship between the housing media and housing returns rather than the relationship at a certain regional setting. Furthermore, panel time-series regression is useful because numerous economic factors may affect housing returns. For instance, Soo (2013) includes numerous controls such as the real interest rate, mortgage rate, unemployment rate, housing starts, building permits, and population. However, including dozens of controls does not guarantee immunity from omitted variable bias. Using panel time-series regression can alleviate this issue by including county interacted with quarter fixed effects to control for local factors that vary at quarterly or less frequency.

The main goal of this analysis is to test whether a change in the proportion of housing content in the media as measured by the change in number of housing newspaper articles predicts the subsequent change in local housing returns in a horizon of less than a year.

The regression specification is as follows.

$$R_{it} = F.E. + \underbrace{\beta_1 \sum_{k=1}^{3} MI_{it-k}}_{\text{Quarter 1}} + \underbrace{\beta_2 \sum_{k=4}^{6} MI_{it-k}}_{\text{Quarter2}} + \underbrace{\beta_3 \sum_{k=7}^{9} MI_{it-k}}_{\text{Quarter 3}} + \underbrace{\beta_4 \sum_{k=10}^{12} MI_{it-k}}_{\text{Quarter 4}} + \sum_{k=1}^{36} \gamma_k R_{it-k} + X_{it} + \varepsilon_{it} \tag{2.2}$$

The main variable of interest is media increase MI_{it}, which measures the increase in housing news content in local newspapers between months t and $t-1$. For example, if housing-related newspaper articles increased between months t and $t-1$ for county i, MI_{it-k} would be 1, and 0 otherwise. I use the indicator for the increase in the number of housing articles rather than the amount of the increase because for small counties, housing articles are sparse and are essentially events[14].

[14] I also use the changes in the numbers of housing articles for larger counties and obtain qualitatively similar results.

Due to the significant processing time between a household's decision to buy or sell a house and the actual transaction, the unit of analysis is a quarter. Given that I am focused on the short-term (rather than long-term, as studied by Soo (2013)) predictive power of the media, I only include lags up to 1 year. According to the Home Buying Institute[15], on average, it takes at least 3-6 weeks for an individual to search for potential homes and another 1-3 weeks for mortgage underwriting and approval. Combined with the escrow and closing process, which takes 30-60 days, buying a home is at least a 2- to 3-month process. Due to this processing time, I expect most of the results to show up in β_2. Namely, I expect $\hat{\beta}_2$ to be positive and significant.

R_{it} is the one-month county-level housing return for county i at month t. This is computed as the log difference of the county's HPI. To account for the long-term auto-correlation exhibited by housing returns, 36-month lagged returns are included as controls. Although I present only results with 36 lagged returns, varying the number of lagged returns controls from 12 months to 48 months does not materially change the results. Lastly, to control for time-varying regional economic fundamentals, I include MSA interacted with year-quarter fixed effects to address MSA-specific economic factors such as income, unemployment, housing permits, and interest rates that affect housing prices but do not change significantly month to month. Furthermore, MSA interacted with year-quarter fixed effects also addresses seasonality. As a robustness check, I also try including a number of fundamentals—national media, 30-year mortgage rate, and new construction—X_{it} as controls.

Table 2.2 summarizes the estimation results for equation 2.2. As indicated by equation 2.2, Quarter 1 is i, $i \in \{0, 1, 2, 3\}$ if there were i months in which number of housing articles increased during the last 3 months. Quarter 2 is i if there were i months in which the number of housing articles increased 4-6 months ago. Quarters 3 and 4 are defined similarly. Namely, Quarter $1, \cdots, 4$ provides estimates for $\hat{\beta}_1, \cdots, \hat{\beta}_4$. Each regression includes lagged 1-month housing returns up to

[15]The article can be found at http://www.homebuyinginstitute.com/how-long-to-buy.php

Table 2.2: Housing Article Increase on Subsequent Housing Return

	1 Month Local Housing return		
	(1)	(2)	(3)
Quarter 1	0.0004	0.001	0.0005
	(0.002)	(0.002)	(0.002)
Quarter 2	0.006***	0.006***	0.006***
	(0.002)	(0.002)	(0.002)
Quarter 3	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)
Quarter 4	−0.003	−0.003	−0.003
	(0.002)	(0.002)	(0.002)
Lagged Housing Returns	✓	✓	✓
Fixed Effects	MSA×Yr/Qtr	MSA×Yr/Qtr, Yr/Month	MSA×Yr/Qtr
Other Controls	No	No	Yes
Observations	46,226	46,226	46,226
Adjusted R^2	0.907	0.909	0.907

Notes: Table 2.2 reports the estimates for $\beta_1, \cdots, \beta_4$ in equation 2.2. Quarter 1 is number of months during last 3 months where number of housing articles increased. Quarter 2,3,4 are defined similarly. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

36 months and MSA interacted with year-quarter fixed effects. Column (2) includes year-month fixed effects to see if more frequently varying national fundamentals change the results. Column (3) includes controls—national media, 30-year mortgage rate, and new construction—as a robustness check.

In all three specifications, I find $\hat{\beta}_2$ to be positive and significant. Over the whole sample, counties with increases in housing news experience 0.6 bp (7.8 bp annualized) higher housing returns 4-6 months later. Including year-month fixed effects or other controls does not change the results, which suggests that MSA interacted with year-quarter fixed effects control effectively for time-varying fundamentals. For future tests, only MSA interacted with year-quarter fixed effects and 36-month lagged housing return are included. A 0.6 bp increase in monthly housing represents about 3% of the average monthly housing returns.

2.4.2 Crisis vs. Post Crisis

The period analyzed, 1998-2015, is made up of different periods that represent different points in the housing cycle. The US housing market experienced a housing boom in 2003-2007 and a housing crisis in 2008-2011. The recent era of 2012-2015 is a post-crisis recovery period. In this section, I test whether the local media effect persists throughout different phases of the housing market cycle. I use the same method as before but divide the sample into different parts.

Table 2.3: Housing Article Increase on Subsequent Housing Return: Sub-Samples

	1 Month Local Housing Return			
	Entire Sample	Housing Boom (2003-2007)	Housing Crisis (2008-2011)	Post-Crisis (2012-2015)
	(1)	(2)	(3)	(4)
Quarter 1	0.0004	0.001	0.005	-0.002
	(0.002)	(0.004)	(0.005)	(0.005)
Quarter 2	0.006^{***}	0.006^{*}	0.003	0.008^{*}
	(0.002)	(0.004)	(0.005)	(0.005)
Quarter 3	0.002	-0.001	0.001	0.007
	(0.002)	(0.004)	(0.005)	(0.005)
Quarter 4	-0.003	-0.010^{***}	-0.005	0.013^{***}
	(0.002)	(0.004)	(0.005)	(0.005)
Lagged Housing Returns	✓	✓	✓	✓
Fixed Effects	MSA $\times$Yr/Qtr	MSA$\times$Yr/Qtr	MSA$\times$Yr/Qtr	MSA$\times$Yr/Qtr
Observations	46,226	13,920	11,136	11,136
Adjusted R^2	0.907	0.925	0.885	0.856

Notes: Table 2.3 reports the estimates for β_k in equation 2.2 in different time periods. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

Table 2.3 summarizes the regression results. Quarter $1, \cdots, 4$ is defined the same way as in Table 2.2. As with the whole-sample analysis (recapped in column (1)), the increase in local housing articles positively affects subsequent housing returns. As summarized in column (2), during the housing boom of 2003-2007, the Quarter 2 effect is 0.6 bp, which is identical to the whole sample, albeit with weaker statistical significance. There is a similar effect during the post-crisis period of 2012-2015. The Quarter 2 effect is slightly stronger at 0.8 bp.

During the housing crisis of 2008-2011, however, the local media effect disappears. One potential explanation is the proportions of positive and negative news. As can be seen in Figure 2.5, the proportion of positive news drops significantly during the housing crisis. As most of the housing news is negative during the crisis, the increase in housing articles does not substantially increase the proportion of households receiving a positive news shock. The differential effect of positive and negative news is considered as an extension and is discussed in Section 4.B.

Another interesting observation is that there is strong reversion during Quarter 4 for the housing boom period. One potential explanation for this is the change in the housing supply. The change in new construction following the increase in housing news is explored in the Section 4.A. Overall, the regression results show that the media effect is stable throughout the housing market cycle.

2.4.3 Impact of Housing Media Increase over Time

In the previous section, I showed that an increase in housing-related articles is correlated with increased housing returns 4-6 months afterward. In this section, to address issues with potential reverse causality and to analyze the lasting impact of housing news, I use panel vector auto-regression (VAR) analysis and the impulse response function.

The regression specification is as follows:

$$Y_{it} = \sum_{k=1}^{K} A^k Y_{it-k} + BX_{it} + e_{it} \tag{2.3}$$

Here, $Y_{it} = (R_{it}, MC_{it})'$, where R_{it} is the 1-month housing return for county i during month t. Media change, MC_{it}, is 1 if the housing content increased in county i during month t, -1 if the housing content decreased, and 0 if it did not change. A^k is the 2×2 matrix of coefficients. e_{it} is the 2×1 vector or error terms.

Lastly, $X_{it} = (\overline{R}_{it}, \overline{MC}_{it})'$ are proxies for local economic fundamentals. Similar to the panel time-series regression fixed effects, these are measured as the average housing return and average housing content increase of county months in a given MSA-quarter. For example, to obtain $\overline{R}_{it}$ in January 2009, I average all of the housing returns between January 2009 to March 2009 from all of the counties in the same MSA as county i.

Hence, equation (3) is identical to the system of equations

$$R_{it} = \sum_{k=1}^{K} A_{11}^{k} R_{it-k} + \sum_{k=1}^{K} A_{12}^{k} MC_{it-k} + B_{11}\overline{R}_{it} + B_{12}\overline{MC}_{it} + \varepsilon_{it} \tag{2.4a}$$

$$MC_{it} = \sum_{k=1}^{K} A_{21}^{k} R_{it-k} + \sum_{k=1}^{K} A_{22}^{k} MC_{it-k} + B_{21}\overline{R}_{it} + B_{22}\overline{MC}_{it} + \varepsilon_{it} \tag{2.4b}$$

Here, A_{ij}^{k} and B_{ij} is the ith row jth column element of matrices A^k and B.

I follow Abrigo et al. (2015) in the assumptions and panel VAR estimation. One necessary assumption is for the error terms to be non-correlated over time—that is, $\mathbb{E}[e_{it}' e_{is}] = 0, \forall t > s$. With the inclusion of sufficient controls for lagged returns, I assume this to be true. I report the regression result for $K = 12$ and, given other studies' use of 12 months to correct for autocorrelation (Soo, 2013), I argue that assuming zero serial correlation of error terms with 12-month lag controls is reasonable[16].

The estimation result for equation 2.3 is presented in Table 2.4. The first two columns report the coefficient estimates with R_{it} as the dependent variable, and the last two columns report the coefficient estimates with MI_{it} as the dependent variable. The coefficient estimates for B are not presented. As expected, due to the well-documented auto-correlation of housing returns, the first column contains many significant estimates. Toward the end of 12 months, the significance wanes,

[16] K between 12 and 36 does not materially change the result

Table 2.4: Panel VAR on Housing Return and Media Housing Content

	R_{it}		MC_{it}	
k	R_{it-k}	MC_{it-k}	R_{it-k}	MC_{it-k}
1	1.085***	-0.001	-0.004	-0.523***
	(0.017)	(0.001)	(0.014)	(0.004)
2	-0.838***	0.001	-0.003	-0.245***
	(0.028)	(0.002)	(0.02)	(0.005)
3	0.406***	0.003*	0.034	-0.071***
	(0.039)	(0.002)	(0.023)	(0.005)
4	0.047**	0.003*	-0.061**	-0.034***
	(0.041)	(0.002)	(0.024)	(0.005)
5	-0.308***	0.001	0.078***	-0.021***
	(0.038)	(0.002)	(0.024)	(0.005)
6	0.367***	0.003*	-0.073***	-0.01*
	(0.031)	(0.002)	(0.024)	(0.005)
7	-0.243***	0.002	0.037	-0.004
	(0.023)	(0.002)	(0.024)	(0.005)
8	0.094***	0.001	0.003	-0.013**
	(0.02)	(0.002)	(0.024)	(0.005)
9	0.031	0.001	-0.021	-0.001
	(0.02)	(0.002)	(0.024)	(0.005)
10	-0.048***	-0.001	0.018	-0.009*
	(0.019)	(0.002)	(0.023)	(0.005)
11	-0.003	0.001	-0.002	-0.012**
	(0.015)	(0.002)	(0.019)	(0.005)
12	-0.012	-0.001	0.007	0.01**
	(0.008)	(0.001)	(0.011)	(0.005)

Notes: Table 2.4 reports the estimates for equation 2.3. The first two columns represents estimates with R_{it} as the depedent variable. The last two columns represents estimates with MI_{it} as the dependent variable. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

giving more credibility to the assumption of zero auto-correlation among the error terms. More relevant are the coefficients in the second column. There are positive and significant relationships between the increase in housing-related articles and housing returns 3, 4, and 6 months afterward.

Housing returns also appear to have some reverse effect on housing articles. There also seems to be strong mean reversion in the MC_{it}, which is expected. Housing news for many counties comes out infrequently, and a decrease in housing news necessarily follows an increase.

Figure 2.6: Impulse Response Function

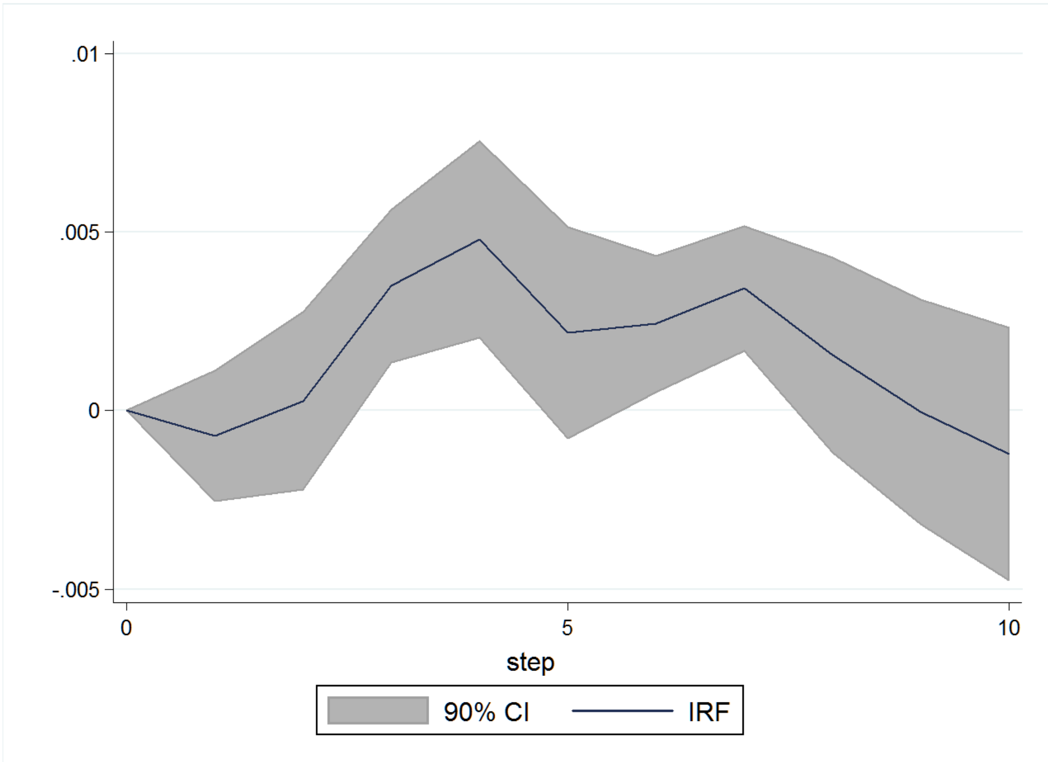


Figure 2.6 plots the impulse response function of increased media housing content on housing returns. The grey band represents the 90% confidence interval based on 10 simulations. Due to the contemporaneous effect of past housing returns on housing news, the media effect estimated is smaller, around 0.5 bp. However, an increase in the media's housing content still appears to lead to an increase in housing returns that peaks at month 4. The effect, however, mostly dissipates after month 9.

The dissipation of the media effect after month 9 can be partially explained by the change in supply. In the model described in Section 2.2, I assume that the housing supply is provided only by households selling homes because new construction takes 6 months on average. If new homes are constructed due to increased housing demand, the local media effect can dissipate. Assuming that the construction of new homes begins at the peak of the local housing effect (3-4 months after the news), new homes would start entering the market about 9 months after the news. I provide rough evidence of this in Section 4.A.

2.5 Do the Media Affect Household Home-Buying Decisions?

I have argued that there is a positive relationship between the lagged increase in housing articles and housing returns. This is consistent with the story that households' home buying decisions are affected by the media and that because of shorting constraints, only the upside is affected. Yet, claiming that media causes increased housing returns is difficult because the variation in the housing media is not exogenous.

In this section, I show evidence that suggests that housing media causally affect households' home-ownership decisions. I use FOMC meeting dates as an IV to suggest that an increase in local housing articles causally leads to an increase in home buying interest 2 weeks later. The causal relation thereby supports the story that the media affect the housing market through households' home-buying decisions. Incorporating the result by (Beracha and Wintoki, 2013) that housing interest predicts housing returns, the findings in this section provide evidence that the media affect short-term housing returns by affecting household decisions.

2.5.1 Identification Strategy

To establish a potential causal link between the number of housing articles and households' home-buying interest, I use FOMC meeting dates as an IV. One of the main tools used by the Federal Reserve System (the Fed) for monetary policy is open market operations. Through open market operations, by buying and selling securities in the open market, the Fed maintains the federal funds rate[17] near the target rate.

The FOMC is a committee within the Fed responsible for the open market operations[18]. The FOMC holds eight regularly scheduled meetings per year, and the exact date of each meeting is confirmed in the previous meeting, which is at least a month in advance. One important objective of these meetings is the setting of a target for the federal funds rate. The federal funds rate plays an important role in the yield curve, and hence in mortgage rates and the housing market. As such, FOMC meetings are likely to lead to increased housing-related news. As the meeting dates are announced well in advance, the timing is exogenous. Furthermore, the market expectation of the FOMC's next target rate can be observable well in advance through economist forecasts.

For this study, I focus on FOMC meeting dates with zero standard deviation in the target rate forecasts and where the mean forecast target rate equaled the actual decision. Namely, I focus on FOMC meetings without any informational surprise. As the financial market is forward looking, an FOMC meeting that confirms what the market has no uncertainty about (as indicated by zero standard deviation in the target rate forecast) should not change anything fundamental about the housing market. This is consistent with Kuttner (2001), which finds that the interest rate reaction to the anticipated target rate is small. Hence, my identification assumption is that FOMC meeting decisions in an environment without any uncertainty as measured by consensus among economists does not affect households' home-buying interest other than through the increased number of housing-related news articles. I consider 82 meetings between 1998 and 2015. The exact meeting dates and target rate decisions are listed in Table 2.5.

[17] "The interest rate at which deposit institutions lend reserve balances to other depository institutions overnight" (From the Fed website)

[18] While I focus on FOMC's setting of federal funds rate for this chapter, FOMC has different tools as well. These include long-term yield targeting such as operation twist

Table 2.5: No Surprise FOMC Meeting Dates

Date	Target Rate	Date	Target Rate	Date	Target Rate
12/22/1998	4.75	5/10/2006	5.00	11/2/2011	0.25
2/3/1999	4.75	9/20/2006	5.25	12/13/2011	0.25
3/30/1999	4.75	10/25/2006	5.25	1/25/2012	0.25
5/18/1999	4.75	12/12/2006	5.25	3/13/2012	0.25
6/30/1999	5.00	1/31/2007	5.25	4/25/2012	0.25
12/21/1999	5.50	3/21/2007	5.25	6/20/2012	0.25
3/21/2000	6.00	5/9/2007	5.25	8/1/2012	0.25
10/3/2000	6.50	6/28/2007	5.25	9/13/2012	0.25
11/15/2000	6.50	8/7/2007	5.25	12/12/2012	0.25
8/21/2001	3.50	6/25/2008	2.00	1/30/2013	0.25
3/19/2002	1.75	8/5/2008	2.00	3/20/2013	0.25
5/7/2002	1.75	8/12/2009	0.25	5/1/2013	0.25
6/26/2002	1.75	9/23/2009	0.25	6/19/2013	0.25
12/10/2002	1.25	11/4/2009	0.25	7/31/2013	0.25
8/12/2003	1.00	12/16/2009	0.25	9/18/2013	0.25
9/16/2003	1.00	1/27/2010	0.25	10/30/2013	0.25
10/28/2003	1.00	3/16/2010	0.25	12/18/2013	0.25
12/9/2003	1.00	4/28/2010	0.25	1/29/2014	0.25
1/28/2004	1.00	6/23/2010	0.25	3/19/2014	0.25
3/16/2004	1.00	8/10/2010	0.25	4/30/2014	0.25
5/4/2004	1.00	9/21/2010	0.25	6/18/2014	0.25
11/10/2004	2.00	11/3/2010	0.25	7/30/2014	0.25
2/2/2005	2.5	12/14/2010	0.25	10/29/2014	0.25
5/3/2005	3.00	1/26/2011	0.25	12/17/2014	0.25
8/9/2005	3.50	3/15/2011	0.25	1/28/2015	0.25
12/13/2005	4.25	4/27/2011	0.25	3/18/2015	0.25
3/28/2006	4.75	9/21/2011	0.25	4/29/2015	0.25
				10/28/2015	0.25

Notes: Table 2.5 lists all the no surprise FOMC meeting dates used in this study.

2.5.2 Google Trends Data

I use Google Trends data to proxy for households' home-buying interest. Buying a home is a complicated process that requires a great deal of research by individual households. Given the dominance of Google in online searching and the ease of online searching for the initial gathering of information, it is likely that households use Google to learn about the home-buying process. The search term I use to measure households' home-buying interest is "buying a home." A household completely new to the home-buying process is likely to start by asking "how do you buy a home?" After considering a number of search terms, which are listed in Table 2.6, the search term "buying a home" is chosen because it is the most frequent search phrase[19]. Using the keyword "buying a home," I obtain SVI for 11 MSAs[20].

Table 2.6: Different Search Terms Considered

Search Term	Relative Search Volume
Buying a Home	100.00
Buying a House	146.24
How to Buy a Home	54.86
How to Buy a House	96.58
Home Buying Process	6.07
House Buying Process	4.30
Buying First Home	15.62
Buying First House	8.94
Buying Second Home	3.66
Buying Second House	0.00
Mortgage Application	15.32
How to Apply for Mortgage	3.03
Home Ownership	25.68

Note: This table lists different keywords considered for analysis using Google Trends Data. Relative search volume is average weekly Google SVI between November 13th, 2011 to November 6th, 2016 (Accessed on November 12th 2016), normalized with "Buying a Home" SVI set to 100.

Although Google's search engine dominance strongly suggests that "buying a home" captures households' home-buying interest, I document the predictive power of the search term's popularity on the mortgage application level to further validate that the Google Trends data capture households' home-buying interest. To do so, I run a time-series regression of the lagged "buying a home" SVI on actual mortgage applications observed through the Mortgage Bankers Association US Purchase Index (MBA Purchase Index) as reported by Bloomberg Terminal. The MBA Purchase Index is based on weekly measurements of nationwide loan applications and is based on

[19]The phrase "buying a house" has a higher relative search volume than "buying a home," but the term "house" indicates a specific type of home, so I opt for "buying a home." The results presented in this section do not qualitatively change when "buying a house" is used instead.

[20]The MSAs are Atlanta, Boston, Chicago, Dallas, Detroit, Houston, Los Angeles, New York, Philadelphia, Portland, and San Francisco

about 75% of mortgage applications.

Table 2.7: "Buying a Home" SVI on Mortgage Applications

	Weekly MBA U.S. Purchase Index		
	2004-2007	2008-2011	2012-2015
	(1)	(2)	(3)
1 Week Lag	0.317	0.117	0.237*
	(0.206)	(0.255)	(0.135)
2 Week Lag	0.161	0.101	−0.018
	(0.199)	(0.233)	(0.108)
3 Week Lag	0.402**	0.118	−0.072
	(0.190)	(0.227)	(0.104)
4 Week Lag	0.175	0.189	−0.152
	(0.179)	(0.252)	(0.114)
5 Week Lag	0.239	0.375	0.060
	(0.194)	(0.267)	(0.143)
6 Week Lag	−0.069	0.223	0.012
	(0.196)	(0.243)	(0.110)
7 Week Lag	0.123	0.016	−0.123
	(0.178)	(0.232)	(0.102)
8 Week Lag	0.286	−0.040	0.119
	(0.179)	(0.234)	(0.112)
9 Week Lag	0.465**	0.353	0.037
	(0.201)	(0.266)	(0.135)
10 Week Lag	0.463**	−0.071	0.057
	(0.207)	(0.264)	(0.108)
11 Week Lag	0.065	0.097	0.156
	(0.179)	(0.234)	(0.095)
12 Week Lag	−0.048	−0.003	0.077
	(0.164)	(0.233)	(0.099)
Observations	197	196	197
Adjusted R^2	0.707	0.921	0.775

Notes: Table 2.7 documents predictive power of "Buying a Home" SVI on loan applications. To account for seasonality and time trend, year-month fixed effects are included. * represents $p<0.1$, ** represents $p<0.05$, and *** represents $p<0.01$.

Table 2.7 documents the regression results. Contrary to other analyses presented in this chapter, both the “buying a home” SVI and the MBA Purchase Index in this analysis are at the national level. To control for the time trend and seasonality, the regressions include year-month time fixed effects. Google Trends only provides weekly data for up to 5 years, and in contrast to other studies featuring a successful merging of daily or weekly series (i.e., Chauvet et al. (2016)) the replication yields an inconsistent merger. As such, I look at three periods—2004-2007, 2008-2011, and 2012-2015—separately.

Column (1) shows the strong predictive power of the “buying a home” SVI and mortgage applications. A 1-point increase in the “buying a home” SVI leads to 0.4- to 0.5-point increases in the MBA Purchase index 3, 9, and 10 weeks later. Although the coefficient estimates are mostly positive for the 2008-2011 and 2012-2015 samples, the strength of the correlation disappears. This is partially explained by how the Google Trends SVI is constructed. The Google Trends SVI measures the relative popularity of a particular keyword compared with other searches. Hence, in the later sample, in which the absolute search volume becomes more volatile, the “housing a home” SVI is likely to be more noisy. As such, I expect the strongest IV analysis result for the 2004-2007 sample.

2.5.3 Analysis

To establish the potential causal role of local housing media on households’ home-buying interest, I use FOMC meeting dates as an IV to conduct a panel two-stage least squares analysis (2SLS). Before running 2SLS, I need to document that housing media increases around FOMC meeting dates. To do so, I estimate the panel time-series regression of the following form.

$$HA_{it} = F.E. + \sum_{k=-4}^{10} \beta_k Meeting_{t-k} + \varepsilon_{it} \tag{2.5}$$

Here, the dependent variable is standardized housing articles (HA_{it}), which is the number of housing articles in MSA i during week t divided by the number of newspapers in MSA i. $Meeting_t$

is an indicator variable that takes a value of 1 if a no-surprise FOMC meeting took place in week i and 0 otherwise. Although I present the results for when a meeting takes place 4 weeks afterward ($k = -4$) compared with a meeting 10 weeks prior ($k = 10$), the analysis with k ranging between -12 and 12 does not materially change the result. MSA interacted with year-month fixed effects is included.

Figure 2.7: FOMC Meetings on Housing Articles

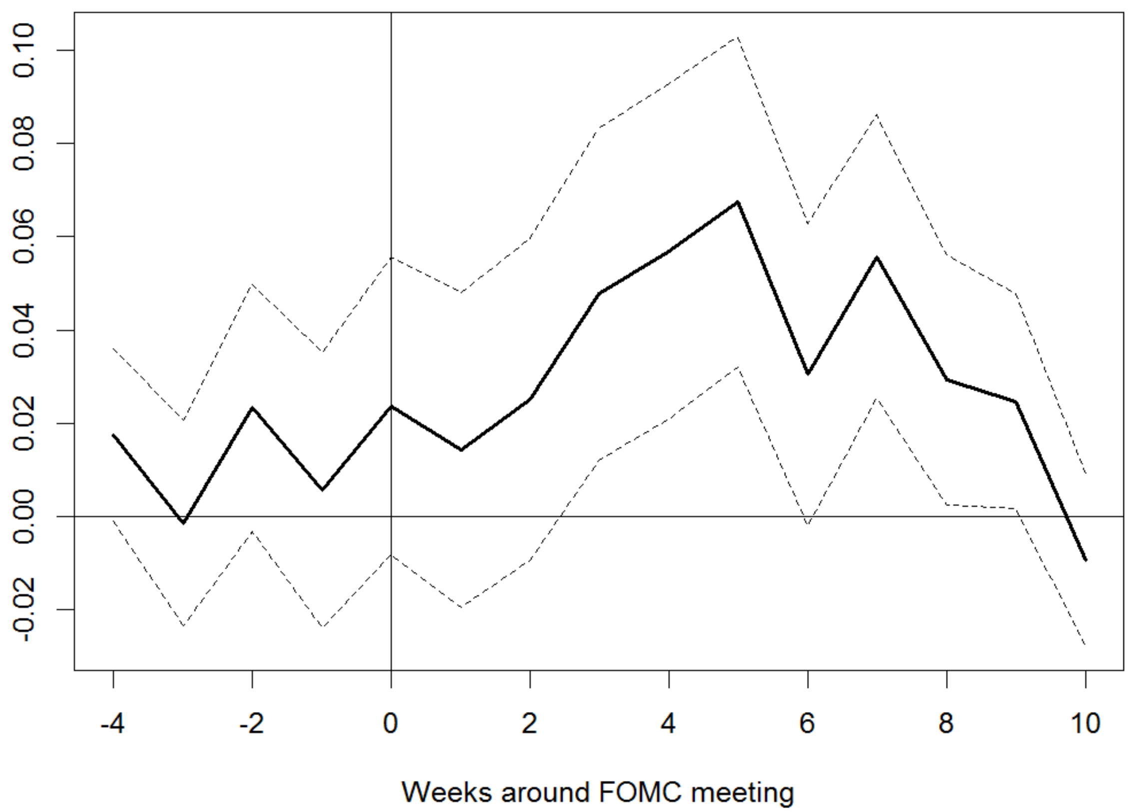


Figure 2.7 plots the coefficient estimates β_k. The thick black line represents the coefficient estimates and the dotted lines are the 10% confidence interval. Contrary to the popular view that information travels fast, the spike in housing articles seems to happen several weeks after the meeting. The number of housing articles starts to increase about 2 weeks after the meeting and peaks at week 5, with about 0.07 more articles per newspaper than in average weeks.

The 2SLS is specified as follows:

$$Media_{it} = F.E. + \sum_{k=1}^{6} \delta_k Meeting_{t-k} + \nu_{it} \tag{2.6a}$$

$$Interest_{it+2} = F.E. + \beta \widehat{Media}_{it} + \varepsilon_{it} \tag{2.6b}$$

Here, the IV is an indicator variable for whether there was a no-surprise FOMC meeting in week t ($Meeting_t$). The dependent variable in the first stage ($Media_{it}$) is the number of housing articles for MSA i in week t divided by the number of newspapers in the MSA. $\widehat{Media}_{it}$ is the predicted $Media_{it}$ by the first stage. The dependent variable for the second stage is the "buying a house" SVI. In both stages, MSA interacted with year-month fixed effects is included. I choose lags up to 6 months because the FOMC effect in Figure 2.7 starts to decline at this point. I consider home-buying interest 2 weeks after because the effect is strongest at this point.

Table 2.8: Housing Media on Housing Interest: IV Analysis

	Home-Buying Interest		
	2004-2007	2008-2011	2012-2015
	(1)	(2)	(3)
Fitted Weekly Housing Article / Newspaper	9.830*	12.280*	−0.879
	(5.711)	(6.749)	(5.030)
Observations	2,277	2,266	2,277
Adjusted R^2	0.143	−0.147	0.465

Notes: Table 2.8 documents IV analysis result described in Section 3.4 * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

Table 2.8 summarizes the result. Column (1) gives the estimate for $\hat{\beta}$ in the 2004-2007 sample. Columns (2) and (3) give the estimates for the 2008-2011 and 2012-2015 samples, respectively. As column (1) shows, there appears to be a potential causal link between housing media and subsequent home-buying interest: one more predicted housing article per newspaper predicts a 9.8-point increase in the "buying a home" SVI 2 weeks afterward.

In column (2), $\hat{\beta}$ is positive and significant, but the adjusted R^2 is negative, suggesting a spurious correlation. The poor explanatory power in this sample may be a combination of the "buying a home" SVI being noisy and negative housing articles dominating during the housing crisis. Furthermore, $\hat{\beta}$ in column (3) is not statistically different from 0. Again, this may be due to the noisiness of the "buying a home" SVI.

2.6 Conclusion

The residential housing market is dominated by households, and households have been documented to exhibit behavioral anomalies and high sensitivity to the media. Hence, it is likely that the media affect households' home-buying behavior, which in turn translates into short-term home price dynamics. I document several pieces of evidence, including a plausibly causal one, that support this hypothesis.

I first document a strong correlation between changes in lagged media coverage of the housing market and subsequent changes in local housing returns. Counties with positive housing news experience returns 8-10 annualized basis points higher than other counties. I suggest that this is due to the shorting constraint in the housing market. Households can respond only to positive housing news, and as positive news is more likely when the absolute number of housing news articles increases, housing news change is positively correlated with subsequent housing return change.

Furthermore, I provide evidence that the media affect households' home-buying interest through the use of no-surprise FOMC meetings as an IV and Google Trends' "buying a home" SVI as a proxy for households' interest in home buying. Although the results hold only for the early sample of 2004-2007, in this period under the identification assumptions, there appears to be some evidence of the media causally affecting households' home-buying interest.

Much work remains to be done in regard to this topic. Perhaps the most important work is to develop a setting that allows more robust testing of the causal story between the media and housing market. Further quantification of the lasting impact of the media on the economy and asset prices through their effects on households' home-buying behavior is also needed.

CHAPTER 3

Stock Market

3.1 Introduction

Do local media affect how individual traders trade? Historically, the main role of media has been to disseminate news. However, as traditional finance assumes that no informational friction exists in the market, the role of media in the stock market has only recently received attention in the literature. There are now many forms of media, ranging from newspapers to online blogs, and the literature related to media explores many aspects of media and their role in the financial market. However, although the majority of households receive most of their news from the local media[1] and that the content of this media differs significantly across regions, studies of the media and stock market focus mostly on individual nationwide press sources such as *The Wall Street Journal* (WSJ) and *DJ Wire*.

To fill the gap in the literature, this study investigates whether local media affect stock markets while controlling for the effects of other nationwide media. Using data from 1991 to 1996[2], I find a predictive relationship between local media and individual investors' trading patterns. In particular, in Metropolitan Statistical Areas (MSAs) in which the local media mention a particular stock, household trades of the stock increase immediately compared with other MSAs in which

[1] the Google Trends search volume index for local media such as the *Los Angeles Times* shows disproportionately high interest from the home state compared with other states.

[2] The main reason for focusing on 1991-1996 is the individual trading data constraint.

local media do not mention the stock. This increase in trading activity is strongest at first and lasts for several months after the local media's attention. Furthermore, using earnings announcement-related news as a plausibly exogenous news shock, I argue that this predictive relationship is due to the causal effect of local media. Lastly, separating the news shock into content and no-content news, I show that the local media effect operates mostly through a salience channel.

In this study, I use the term "local media" to describe the most readily accessible sources of media for individuals and households. During the sample period of this analysis, 1991-1996, the cost of obtaining information was much higher than it is today, especially because access to the Internet was limited[3]. As such, costs related to obtaining information from geographically distant sources were likely to be high, and the use of geographically local newspapers as a proxy for local media is likely to be valid during the 1991-1996 sample period. If the study were to be extended to a more recent period, local media could include Internet news sources such as Facebook and Twitter.

The literature assumes that the media play a causal role in financial markets and largely classifies this role in two ways. The first classification is information transmission. An example of this mechanism is Klibanoff et al. (1998), which documents the tendency of closed-end country fund prices to move more in line with their net asset value when related news appears in *The New York Times* (NYT). Klibanoff et al. argues that this is because the media aid the dissemination of news. Another example is Tetlock (2007), which builds a sentiment index based on a WSJ column. Tetlock argues that the sentiment index has predictive power over other fundamentals because the wording of the articles contains information. In short, the literature on the role of media as an information channel attributes the causal effect of media on financial markets to news updating market players' information sets.

The information transmission mechanism also plausibly explains the relationship between lo-

[3]Public access to the Internet started in the early 1990s, and one of the first search engines, Yahoo, was founded in 1994.

cal media and the market. In general, accessing updated financial news is costly, as it requires subscribing to relevant news sources and digesting the content. Given an average household's portfolio size of $100,000[4], it is conceivable that the majority of individual traders[5] choose not to actively seek the latest news but to passively update their information sets through local media channels. Under these assumptions, local media would serve as the dominant source of information for individual traders and would thus affect individual traders' trading behavior.

Media can also affect financial markets through salience. Portfolio theory provides guidance on how investors should trade. Traditional portfolio theory argues that all investors should hold market portfolios scaled appropriately for their risk appetite. However, in the current financial market, which consists of thousands of securities, solving the optimization problem to determine a market portfolio is impossible for individuals, let alone institutions. As such, the literature is filled with examples that document deviations from portfolio theory, which is evidence of the difficulty of optimization[6].

Salience is a heuristic used by individuals to aid their decision making. Salience refers to the tendency of individuals to focus on items or information that is more visible or prominent. In this setting, regardless of the informational content or context, a news article mentioning a particular stock may prompt an individual to start his or her portfolio search around the particular stock, which subsequently increases his or her likelihood of trading the stock. Hence, if salience is the dominant factor of the media effect, the media should affect households' trading behavior regardless of informational content.

A number of findings in the finance literature can be categorized into this salience channel. For instance, Barber and Odean (2008) documents individual investors' tendencies to buy attention-

[4]Based on Barber and Odean (2000), a study using the same individual trading data as this study, the wealthiest households that trade using a large discount brokerage have a mean starting portfolio size of $150,000.

[5]In this study, I use "individuals" and "households" interchangeably.

[6]Examples include households holding too few stocks and naive diversification strategies (Goetzmann and Kumar, 2008)

grabbing stocks, such as those with abnormally high trading volumes or extreme returns. Similarly, Hartzmark (2014) finds that individual investors tend to sell stocks that grab their attention—those that perform the best or the worst. Using Google Trends data as a proxy for investor attention, Da et al. (2011) documents a strong correlation between investor attention to a particular stock and the subsequent returns of the stock. Fang et al. (2014) documents evidence of salience among mutual fund managers, which suggests that the media's effect through the salience channel is not limited to individual investors. Frieder and Subrahmanyam (2005) documents individual investors' tendency to trade well-known brand stocks, which is consistent with salience. In line with these findings, local media coverage of a specific company can raise the salience of the company compared with other stocks or financial instruments, leading to increased trading activity for its stock.

This study is related to a number of strands of finance literature. One of the most closely related strands is the role of the media in the stock market. In addition to Klibanoff et al. (1998), Huberman and Regev (2001) provides evidence that the media play an important role in the stock market. Huberman and Regev conduct a case study that documents the changes in a biotechnology firm's stock prices after different media sources reported on its supposedly ground-breaking findings in cancer research. This case study shows clearly that different media sources can have various price impacts on a stock's prices and that even non-informational news can permanently affect stock prices; these findings support the salience channel. However, as methods to systematically quantify the informational content of media articles did not yet exist, these earlier studies do not provide conclusive evidence that the media causally affect financial markets.

The most recent wave of studies of the media's effects on the stock market can be attributed to Tetlock (2007). Tetlock uses text mining to assign sentiment value to the articles of a WSJ column and finds that sentiment (particularly high pessimism) has strong predictive power for the returns of the Dow Jones Index. Tetlock et al. (2008) builds on the result of Tetlock (2007) and finds that the sentiment index for a particular stock based on all of the relevant articles in the WSJ and *DJ Wire* has predictive power for firm-specific returns. More recent studies document different aspects of the media effect. For example, (Fang and Peress, 2009) documents higher return premiums from stocks with low media coverage and (Tetlock, 2011) finds that individual investors tend to overreact to stale news.

Another strand of literature related to this study is home bias. Investors are known to prefer investments that are geographically closer. Cooper and Kaplanis (1994) documents home bias in the form of investors investing large proportions of their portfolios with financial instruments from their home countries. More recent studies document home bias at a much more granular level. Coval and Moskowitz (1999) finds that US investment managers tend to invest more in companies in the same city and that these investments tend to perform better. Hong et al. (2008) argues that home bias leads to price dynamics in which the stock prices of companies that are the "only game in town" are higher. Other studies find similar effects in venture capital investments (Cumming and Dai, 2010), individual investors (Seasholes and Zhu, 2010), and different countries, such as Sweden (Bodnaruk, 2009). Although not directly related, Giannetti and Wang (2016) finds that households living near corporations with scandals are less likely to participate in the stock market.

Engelberg and Parsons (2011) presents an analysis that is closest to this study, using subtle differences in the timing of news delivery across different trading regions due to extreme weather and print deadlines as a source of plausibly exogenous variation in the potentially causal link between local media and local trading. The main difference between Engelberg and Parsons (2011) and this chapter is the identification mechanism. This chapter uses earnings announcement dates as an instrumental variable to obtain the exogenous variation in the local media. This approach addresses the potential limitation of Engelberg and Parsons (2011) that local news on an earnings announcement need not come after the announcement. Under salience, a local media story in anticipation of an earnings announcement can lead to same dynamics of individual trading. Differentiation between the salience and information channels is an additional contribution of this study. This study's evidence in support of the causal relationship between local media and individual trading reinforces the conclusion of Engelberg and Parsons (2011).

The chapter is structured as follows. Section 3.2 describes the data sources and criteria used in trimming the data. Section 3.3 documents the predictive power of local media on households' subsequent trading behavior. Section 3.4 presents causal evidence, and Section 3.5 presents evidence that the salience channel may be dominant. Section 3.6 concludes the chapter.

3.2 Data

This study sources data from multiple databases. Individual investor trading data come from a large discount brokerage (LDB) that is identical to the one used in a series of papers by Barber Barber and Odean (Barber and Odean, 2000, 2001, 2002, 2008). The data consist of all of the trades executed by the brokerage's account holders between January 1991 and November 1996. LDB data also contain non-household accounts such as those registered under corporations or investment clubs; however, most of the data appear to belong to households. About 98% of accounts are general brokerage or IRA accounts[7]. Hence, although the LDB data contain neither the entire individual trader population nor exclusively individual traders, I assume that the sample is representative of the individual trading population.

LDB data consist of multiple files: a file with all of the trades executed by a particular account, a file that links each account with a household number, and a file that provides details on each household, such as the zip code of the household when the account was opened. Household details are collected only for when the account was opened, and hence I cannot observe whether the household moved during the analysis period. Trading accounts are aggregated at the household level. The sample of households in LDB is restricted to those residing in a MSA. All households that placed at least one buy or sell order between January 1991 and November 1996 with a valid zip code[8] are aggregated at the MSA level. Only MSAs with 100 or more households in the LDB

[7]Various attempts to trim the data to isolate "ordinary" households do not materially change the results.

[8]Many accounts in the LDB data are linked with military zip codes, which cannot be geographically located.

data are analyzed, which leaves 62 MSAs.

LDB data tags each trade with an 8-digit CUSIP number. The first 6 digits of a CUSIP number represent the company that issued the financial security, and the last 2 digits represent the specific issue[9]. The first 6 digits of CUSIPs are matched to ticker codes based on a crosswalk file obtained from the Center for Research in Security Prices (CRSP), and household-level trades are aggregated using the ticker codes. Due to the CUSIP–ticker matching mechanism, this study does not differentiate between different types of securities (i.e., common stock, preferred stock, etc.). However, as the majority of trades involve common stocks, I henceforth refer to these trades as stock trades.

Table 3.1: LDB Data Summary Statistics

MSA Level Characteristic	Mean	SD	Min	Max
Mean Starting Equity	$107,731.50	$33,458.31	$61,412.59	$277,198.70
Median Starting Equity	$35,987.68	$5,224.42	$22,592.00	$50,045.00
Mean Number of Households	560.19	810.18	107.00	4103.00
Mean Household Head Age	42.71	2.54	36.42	49.75
Median Household Head Age	45.81	2.30	42.00	52.00
Mean Annual Trades	6.65	1.19	4.37	10.27
SD Annual Trades	14.67	5.55	7.22	30.14
Mean Number of Stocks in Portfolio	15.15	1.74	10.95	19.01

Note: This table summarizes the LDB data. All trading accounts are first aggregated at the household level and the households are aggregated at the MSA level. The summary statistics are computed using MSAs with at least 100 households.

Table 3.1 provides some summary statistics of households in the trimmed LDB data. Households residing in the 62 MSAs have $107,731.50 in starting equity on average[10]. There is a substantial heterogeneity in wealth between different MSAs. Households in the poorest MSAs have $61,413 in starting equity on average, while households in the richest MSAs have $277,199 in starting equity on average. As indicated by the much smaller median starting equity, the income distributions in the MSAs are skewed right.

Other household characteristics are relatively consistent across the MSAs. The average age of

[9]CUSIPS are typically 9 digits, but the 9th digit is an automatically generated checksum. Hence, the first 8 digits uniquely define a financial security.

[10]I report starting equity rather than income because the LDB data report household income in large brackets. Starting equity is the sum of the starting equities of all of the accounts belonging to a particular household

the heads of households is 42.71. The MSAs with the youngest households have heads of households aged about 36 years, while MSAs with the oldest households have heads of households aged about 50 years. Households make on average 6.65 trades per year, and, as can be observed by the relatively small standard deviation of 1.19, the average number of trades executed by households is largely constant across the sample MSAs. However, as may be noted by the mean standard deviation in the number of annual trades, there is a great deal of variation within each MSA in the number of household trades. Lastly, we note that within the 6-year timeframe of the LDB data, each household traded 15.15 different stocks on average.

Newspaper data comes from *NewsBank*'s "Access World News" (AWN) database. AWN provides news articles by various local media from 1978 onward. Although the database is not a comprehensive set of all of the printed media available to households, I assume that the sample of articles included in the database is unbiased in representing the local media news flow. I restrict my sample to English-based printed newspapers.

AWN uses a keyword-based search method. Articles containing specific keywords are returned when a user uses the AWN database. Although there is no guarantee that a specific set of keywords perfectly captures all news articles related to a stock, similarity in results using different well-motivated sets of keywords suggests that the stock article data used in this study are representative. I use three different sets of stock articles, which I refer to as "Ticker General News," "Keyword General News," and "Earnings News," respectively. Ticker general news articles are obtained using the firm's ticker symbol and "Stock" as keywords. For example, ticker general news articles on Apple Computers are obtained using the keywords "AAPL" and "Stock." All articles published between January 1, 1991 and November 30, 1996 are included in the sample. The ticker general news article set contains about 1,400 news articles related to 13 stocks.

As ticker symbols are often odd combination of letters, ticker general news articles are likely to be relevant to the intended stock. However, news articles with text containing a stock ticker symbol almost always have the stock name (i.e., "Bristol-Myers Squibb Co. (NYSE: BMY)"), but not vice versa. As a result, ticker general news may be too restrictive and may omit many relevant news articles. To alleviate this issue, I also present results using keyword general news articles, which are obtained using a stock-specific keyword and the term “Stock.” Table 3.3 lists the keywords for each of the stocks used in this study. The keywords are manually selected to be firm-specific, and the resulting article results are manually scanned to ensure that the keywords are not overly restrictive or general. The keyword general news article set contains about 298,000 news articles related to 34 stocks.

For earnings news articles, the keywords used are the stock-specific keyword, “Earnings,” and “Stock.” For example, earnings news articles on Bristol Myers Squibb are obtained via the keywords “Bristol,” “Earnings,” and “Stock.” Furthermore, in an attempt to capture only articles relevant to earnings announcements, I keep only articles that are printed between the day of an earnings announcement and 6 days afterward. There are about 10,000 articles in the earnings news article set related to 34 stocks.

Earnings data come from Thomson Reuters’ Institutional Brokers Estimate (I/B/E/S) Summary History - Surprise History file. I/B/E/S is a widely used database for studies related to analyst forecasts or recommendations (e.g. Ivković and Jegadeesh, 2004). I obtain all quarterly earnings announcement dates and the actual earnings figures for all stocks in the database between January 1993 to November 1996[11]. I also obtain the mean and standard deviations of analyst earnings forecasts available to the public before the earnings announcements.

The stocks included in this analysis are screened based on the following criteria. For the ticker general news article set, I first isolate the 30 stocks that are traded by the most households within the LDB data. To increase the article search accuracy[12], I further restrict the stocks to those with at least three characters in their ticker symbols. This selection criterion results in 13 stocks.

[11] I use the Surprise History file because it gives the most reliable analyst forecast standard deviation. However, the file has a relatively shorter period than other files, as it starts on January 1993.

[12] For example, Chrysler Corp (Ticker:C) was one of these 30 most widely traded stocks. However, the search terms “C” and “Stock” would result in articles that have the word “stock” and any word with the letter “C” in it.

The stocks represented by the keyword general news article set and the earnings news article set are chosen via slightly different selection criteria. As I use both sets of data in causal analysis involving earnings announcements, I first isolate stocks that are included in both the I/B/E/S and LDB data. Furthermore, to ensure sufficient data points across time and geography, I consider only stocks that are traded by at least 1,000 households. These selection criteria result in 34 stocks.

Table 3.2: Ticker General News Summary Statistics

Ticker	Num Articles	Ticker	Num Articles	Ticker	Num Articles
AAPL	42	HWP	66	NOVL	17
BMY	152	INTC	132	TMX	32
CCI	154	MOT	141	WMT	133
CPQ	80	MRK	139		
CSCO	56	MSFT	259		

Note: This table lists 13 stocks that included in ticker general news dataset. The 13 stocks are based on 30 most widely traded stock based on LDB data. Only stocks with at least 3 character-long tickers are left, resulting in the above 13 stocks.

Tables 3.2 and 3.3 list the stocks included in this analysis and the breakdown of the numbers of articles across different stocks. Table 3.2 lists stocks included in the ticker general news dataset. The company names represented by these ticker symbols are included in the Appendix [Add Appendix reference]. As expected, less stock news is searched through ticker symbols than through keywords. There are also large differences in the numbers of relevant stocks.

The newspaper sample consists of newspapers that printed at least one relevant stock article between January 1991 and November 1996 as reported by the AWN database, and only MSAs with at least one local newspaper represented in the newspaper sample are included in this analysis. As the keyword general news article set contains the most news articles, the most MSAs are analyzed using the keyword general news article set. After merging and trimming the data, the analysis sample contains 33, 34, and 37 MSAs for the ticker general news article set, earnings news article set, and keyword general news article set, respectively. The locations of these MSAs are plotted in Figure 3.1. As can be seen in the figure, the MSAs represent areas from all over the US.

Table 3.3: Keyword General News / Earnings News Summary Statistics

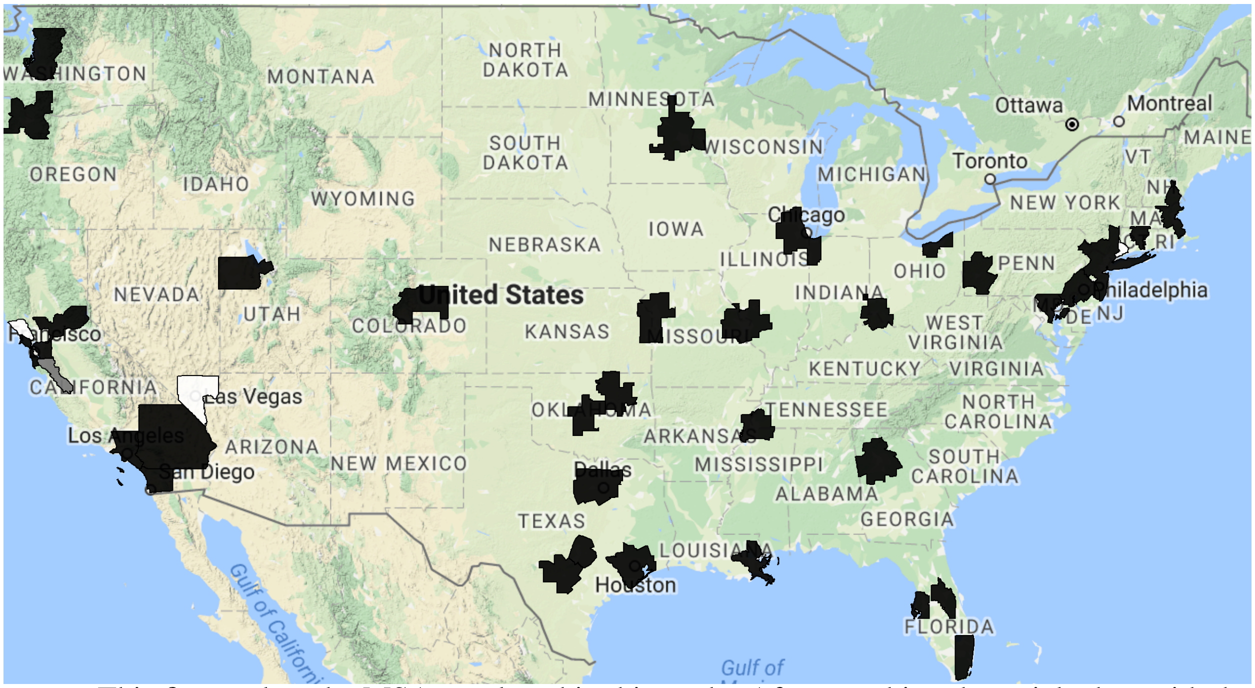
	Keyword General News		Earnings News	
Ticker	Keyword	Num Articles	Keyword	Num Articles
AAPL	Apple	23,928	Apple	634
ABT	Abbott	3,418	Abbott	98
AMGN	Amgen	1,704	Amgen	38
BA	Boeing	13,360	Boeing	471
BMY	Bristol	8,603	Bristol	157
C	Chrysler	18,727	Chrysler	838
CSCO	Cisco	2,713	Cisco	79
DEC	Digital Equipment	5,257	Digital Equipment	176
DELL	Dell	6,228	Dell	288
F	Ford	12,125	Ford	915
GE	General Electric	12,016	General Electric	174
GTE	GTE Corp	1,524	GTE	89
HAN	Hanson PLC	523	Hanson	9
IBM	IBM	33,364	IBM	1,082
INTC	Intel	14,969	Intel	616
JNJ	Johnson and Johnson	83	Johnson	551
KM	K Mart	2,240	K Mart	47
KO	Coca	16,377	Coca	296
LLY	Eli Lilly	2,415	Lilly	99
MCD	McDonald	16,860	McDonald	204
MCIC	MCI	6,880	MCIC	0
MO	Philip Morris	12,295	Philip Morris	320
MRK	Merck	8,035	Merck	380
MSFT	Microsoft	22,613	Microsoft	400
NOVL	Novell	2,712	Novell	107
OXY	Occidental	1,638	Occidental	48
PAC	Pacific Telesis	2,138	PAC	6
PEP	Pepsico	4,942	Pepsico	132
RN	Nabisco	7,739	Nabisco	187
S	Sears	16,517	Sears	646
SYN	Syntex	1,228	Syntex	34
WMT	Walmart	404	Walmart	20
WX	Westinghouse	6,097	Westinghouse	149
XON	Exxon	8,931	Exxon	374

Note: This table lists 34 stocks that included in keyword general news and earnings news dataset. The 34 stocks represent stocks that are traded by at least 1,000 households in LDB data and has I/B/E/S earnings announcement data.

3.3 Predictive Power of Local Media

In this section, to motivate analysis of the relationship between local media and households' trading, I document the predictive power of local media for households' trading patterns. The existence of such predictive power suggests that there may be a causal relationship between local media and

Figure 3.1: MSAs Analyzed



Notes: This figure plots the MSAs analyzed in this study. After matching the article data with the LDB data, 33 MSAs (black MSAs) remain for the ticker general news article set, 34 MSAs (black + grey MSAs) remain for the earnings news article set, and 37 MSAs (black + grey + white MSAs) remain for the keyword general news article set.

households' trading patterns. I mainly use weekly panel time-series regressions. I show that stock-related local news is correlated with an increase in the likelihood of local households trading the stock and that this increase lasts for several weeks. Causality is explored in Section 3.4.

3.3.1 Predictive Relationship between Local Stock News and Local Trading Activity

In this section, I present results using the ticker general news article set and the keyword general news article set obtained through the search criteria described in Section 3.2. For 1991-1996, there are about 1,400 articles from local newspapers related to the 13 companies in the ticker general news article set. Similarly, there are about 298,000 keyword general news articles related to the 34 companies in the keyword general news article set. Merging these sets with the LDB data yields 33 and 37 MSAs, respectively. To document the predictive power of local media beyond national indices such as national media, I use weekly panel regressions. Combining the LDB and AWN data yields a dataset that varies cross-sectionally between two dimensions—geography and stock—which allows me to control for both the time-varying local fundamentals and stock fundamentals through fixed effects. Based on the assumption that any local media effect would quickly interact with local trading patterns, the analysis period is set to weekly. In contrast to daily data, weekly data allow me to abstract away from local media content that may vary depending on the day of the week. However, I also explore long-term effects via monthly time-series analysis in the appendix.

The regression specifications are as follows.

$$Y_{ijt} = F.E. + \sum_{k=0}^{K} \beta_k N_{ijt-k} + \varepsilon_{ijt} \tag{3.1}$$

The dependent variable Y_{ijt} is the number of households in MSA i that traded stock j in week t. To account for differences in scale due to differing trading populations in the MSAs, I divide the number of trading holds by the sample mean. Hence, the dependent variable can be interpreted as the proportion of the active trading population in MSA i that traded stock j in week t. For brevity, I hereby refer to Y_{ijt} as local trades. Each sample mean is computed by averaging the non-zero observations during the whole sample (weekly aggregation yields many zero observations, especially for smaller MSAs). A week is defined as starting on Monday and ending on Sunday.

N_{ijt} is the number of local articles related to stock j in MSA i at week t. To account for differences in scale due to differing numbers of local newspapers in MSAs, I divide the number of local articles by the sample mean, which is computed in the same manner as Y_{ijt}. For brevity, I hereby refer to N_{ijt} as the number of local articles. I present regression results using up to a 6-week lag ($K = 6$), but including lags up to 12 months does not affect the significance of the coefficient estimates in weeks 1-6. Although the choice of $K = 6$ may seem arbitrary, as the main goal of this analysis is to document the predictive power of local media, the number of lags does not matter greatly.

I include several fixed effects. To account for time-varying stock specific fundamentals, I include stock interacted with week (i.e., year-week) fixed effects. To account for time-varying MSA-specific fundamentals, I also include MSA interacted with week fixed effects. Lastly, I include MSA interacted stock fixed effects to capture any remaining level effect among stock–MSA combinations.

Table 3.4: Local News and Subsequent Trading

	Weekly Local Trades at MSA i for Stock j	
	Ticker General News	Keyword General News
	(1)	(2)
Concurrent N_{ijt}	-0.004	0.023***
	(0.019)	(0.002)
1 Week Lag N_{ijt}	0.016	0.008***
	(0.019)	(0.002)
2 Week Lag N_{ijt}	0.048**	0.003**
	(0.019)	(0.002)
3 Week Lag N_{ijt}	-0.006	0.002
	(0.019)	(0.002)
4 Week Lag N_{ijt}	-0.022	0.003**
	(0.019)	(0.002)
5 Week Lag N_{ijt}	0.022	0.002
	(0.019)	(0.002)
6 Week Lag N_{ijt}	0.037*	0.003*
	(0.019)	(0.002)
Fixed Effects	MSA × Week, Stock × Week, MSA × Stock	MSA × Week, Stock × Week, MSA × Stock
Observations	129,987	381,174
Adjusted R^2	0.340	0.304

Notes: This table summarizes coefficient estimates for equation 3.1. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

Table 3.4 summarizes the regression results. Column (1) documents the results using the ticker general news article set and column (2) documents the results using the keyword general news article set. Although the timing of significant correlations between the local media and subsequent local trading is slightly different between the two columns, both regressions appear to indicate that an increase in stock-related local news is correlated with a subsequent increase in the local trading of the stock.

As shown in column (1), in an MSA with a 1-standardized-unit increase in local articles, there are 4.8% more local trades 2 weeks later compared with a MSA without news. There is also a statistically significant increase of 3.7% 6 weeks later. As column (2) shows, there appears to be a much stronger correlation. There is a steady increase in trading households that begins immediately after a local news increase and lasts for 6 weeks. In an MSA with a 1-standardized-unit increase in local articles, there are 2.8% more local trades immediately afterward compared with a MSA without news. Note that stocks in the ticker general news article set are more widely traded by households; hence, the higher coefficient estimates are as expected. However, the weaker statistical significance using the ticker general news set is consistent with attenuation bias. Furthermore, these results are robust to removing any single MSA or stock and hence do not seem to be driven by any single anomaly.

Numerous endogeneity issues prevent causal inference from this analysis. One possible issue can be attributed to reverse causality. Although specific individual trading data are not public, they are likely to be unavailable to local newspapers. However, because of the neighborhood effect documented by Ivković and Weisbenner (2007), local media may have a sense of which stocks are widely traded or held by local households. Local newspapers may be more inclined to write about stocks that local households widely hold, and the resulting positive correlation between local media and local trading patterns may be spurious. To address these concerns, I use plausibly exogenous variations in local articles related to earnings announcements to conduct causal analysis in Section 3.4.

3.4 Causal Effect of Local Media

In Section 3.3, I document a positive correlation between local articles and local trades. Although this correlation suggests that there may be a causal relationship, due to endogeneity issues such as reverse causality, the causal relationship cannot be tested without exogenous variations in local media. In this section, I test for the causal effect of local media on households' trading behavior using plausibly exogenous variations in local news articles as predicted by earnings announcements.

Earnings announcements are useful in this analysis mainly because the timing of these announcements is relatively exogenous. Securities and Exchange Commission (SEC) regulations require publicly traded firms to announce their quarterly earnings within 45 days of quarter ends[13]. As such, the relative timing of the announcements is known beforehand, although the exact day of an announcement may change. Furthermore, many of the stocks in this dataset pre-announce their earnings announcement dates. Hence, absent any systematic difference between MSAs in which local media decide to report earnings news and MSAs in which local media decide not to, the resulting difference is evidence of the causal effect of local media.

Although I cannot explicitly test whether there is any systematic difference between MSAs with earnings-news-related articles and MSAs without such news, the data do not show obvious patterns. Table 3.5 summarizes how many earnings announcements local media in each MSA reported. For example, MSAs on average had news articles reporting on 5.21 out of 16 earnings announcements for AAPL. Given that local media in most MSAs report on some but not all earnings announcements for each stock, there does not seem to be an explicit pattern, which suggests systematic differences among MSAs.

[13]Companies must report fourth quarter earnings announcements within 90 days.

Table 3.5: MSAs Covering Earnings Announcements

Ticker	Mean	SD	Total	Ticker	Mean	SD	Total
AAPL	5.21	4.42	16	KO	3.24	3.44	16
ABT	1.26	1.71	16	LLY	1.62	2.44	14
AMGN	0.74	1.4	14	MCD	2.74	3.73	16
BA	5.21	4.21	16	MO	3.85	3.55	15
BMY	2.38	2.73	16	MRK	4.21	4.77	16
C	7.5	4.17	16	MSFT	4.68	3.96	15
CSCO	1.26	1.91	16	NOVL	1.76	2.87	14
DEC	2.29	2.43	14	OXY	0.74	1.93	16
DELL	3.12	3.17	14	PAC	0.18	0.76	16
F	7.24	4.36	16	PEP	2.24	2.97	12
GE	2.35	2.75	16	RN	2.56	2.98	16
GTE	1.5	2.99	14	S	6.68	5.16	12
HAN	0.12	0.41	6	SYN	0.35	0.69	6
IBM	8.44	5.05	15	WMT	0.24	0.74	13
INTC	6.24	4.55	16	WX	2.03	2.42	15
JNJ	6.47	4.15	15	XON	4.21	4.53	16
KM	0.76	2.35	12				

Note: This table summarizes data on the number of earnings announcements that local newspapers from each MSA reported.

3.4.1 Earnings Announcement and Local Stock News

The main analysis in this section uses 2-stage least squares (2SLS) regression with earnings announcement dates as an instrumental variable (IV). Assuming that earnings announcement dates are exogenous, I assume that the exclusion restriction holds, but the relevance condition needs to be tested. In this section, I document the strong correlation between earnings announcement dates and stock-related news to justify the validity of the 2SLS framework in this setting.

To test for the correlation, I estimate the following regression.

$$N_{jt} = F.E. + I_{jt} + I^e_{jt} + \varepsilon_{jt} \tag{3.2}$$

Here, N_{jt} is the number of news articles nationwide related to stock j during week t. I aggregate the news articles at the weekly level to abstract away from potential differences between different weekdays. I_{jt} is an indicator variable indicating whether there was an earnings announcement for stock j during week t.

A number of studies document the information contained in the timing of earnings news. Studies such as (Begley and Fischer, 1998), (Kross, 1981), and (Kross and Schroeder, 1984) find that earnings announcements that come before the expected time often contain positive surprises, while earnings announcements that come after the expected time often contain negative surprises. In the spirit of these studies, I also include expected earnings announcements to test whether earnings-related articles are anticipatory or reactionary. I^e_{jt} is an indicator variable indicating whether an earnings announcement is expected for stock j during week t. The expected earnings announcement date is estimated by first averaging the delay between the quarter-end date and the announcement date for the past four quarterly earnings announcements. The average delay is added to the current quarter-end date to obtain the expected earnings announcement date.

Table 3.6: Earnings Announcement Week and Stock Related News

	Weekly Number of News Articles Related to Stock j			
	Keyword General News		Ticker General News	
	(1)	(2)	(3)	(4)
Ann. Week	12.974*** (0.797)	12.816*** (0.883)	0.275*** (0.070)	0.304*** (0.087)
Expected Ann. Week		0.367 (0.880)		−0.049 (0.088)
Fixed Effects	Stock × Yr/Mon, Week	Stock × Yr/Mon, Week	Stock × Yr/Mon, Week	Stock × Yr/Mon, Week
Observations	10,676	10,676	4,069	4,069
Adjusted R^2	0.703	0.703	0.256	0.256

Notes: This table summarizes coefficient estimates for equation 3.2. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

Table 3.6 summarizes the regression results. Columns (1) and (2) report the results using the keyword general news article set, and columns (3) and (4) report the results using the ticker general news article set. Each regression includes the stock interacted with year-month fixed effects to account for stock-specific time-varying factors and week fixed effects to account for more frequently varying nationwide factors.

As can be seen in columns (1) and (3), earnings announcement week is a statistically significant predictor of a spike in stock-related news. For the keyword general news article set, there are 13 more stock-related news articles nationally during an earnings announcement week on average compared with other weeks. Similarly, for the ticker general news article set, there are 0.28 more stock-related articles nationally on average. Given that the average number of weekly news articles is 21.5 for the keyword general news article set and 0.33 for the ticker general news article set, earnings announcement week predicts a significantly large spike in the number of stock-related news articles[14].

Comparing these results with columns (2) and (4), the coefficients for announcement week stay relatively constant when the indicator variable for expected announcement week is also included. However, expected announcement week does not have any additional predictive power for stock-related news, suggesting that the media react to actual earnings announcements rather than to forecasts of the announcements. I also include lagged indicator variables, but the only statistically significant predictor for news articles is the concurrent indicator variable for actual earnings announcements. This suggests that the media react to earnings announcements and that this reaction is short-lived.

3.4.2 Causal Effect

I document the causal effect between local media and local trading in two different ways. I first use earnings announcement dates as an IV in a 2SLS regression. Intuitively, the first stage of the 2SLS regression isolates variations in local news that can be attributed to earnings announcement dates. The second stage of the regression tests the correlation between local trading and the isolated variation in local news.

[14] I also try regressions with the dependent variable standardized by the sample mean, and the results are qualitatively similar.

The regression specification for the 2SLS regression is as follows.

$$Z_{ijt} = F.E. + \beta X_{jt} + \varepsilon_{ijt} \tag{3.3a}$$

$$Y_{ijt} = F.E. + \delta \widehat{Z_{ijt}} + \nu_{ijt} \tag{3.3b}$$

As in Equation 3.1, Y_{ijt} is the proportion of active trading households in MSA i that traded stock j in week t. $Z_{ijt} = (N_{ijt}, N_{ijt-1}, \ldots, N_{ijt-K})$ is a vector of the lagged number of standardized local stock news articles. As before, I report only the result for $K = 6$, but the results are qualitatively identical with K up to 12. $\widehat{Z_{ijt}}$ is the vector of predicted standardized local stock news articles as predicted by equation 3.3a. $X_{jt} = (I_{jt}, I_{jt-1}, \ldots, I_{jt-K})$ is a vector of the lagged indicator variables indicating whether there was an earnings announcement related to stock j in week t.

In both stages of the 2SLS regression, I include a number of fixed effects. First, MSA interacted with week fixed effects is included to control for time-varying MSA fundamentals, stock interacted with year-month fixed effects to control for time-varying stock fixed effects, and MSA interacted with stock fixed effects to control for any remaining level effect between stock–MSA pairs. Note that in this analysis, time-varying stock fixed effects are controlled via stock interacted with year-month fixed effects because there is no geographic variation in I_{jt}.

I also use the earnings news article set to check the robustness of the results from the 2SLS. As described in Section 3.2, the procedure to extract the earnings news set is intended to capture only earnings-announcement-related news. As such, variations in the earnings news article set are more likely to be exogenous than those in the general news sets, and the results of panel regression can enable causal inferences. The regression specification for this analysis is identical to Equation 3.1.

To exclude the potential informational channel from earnings announcement timing, I consider only earnings announcements that occur in the week they are expected. After trimming the data, I have observations from 37 MSAs, 34 stocks, and 323 earnings announcements for the 2SLS analysis and 34 MSAs, 33 stocks, and 311 earnings announcements for the earnings news article set analysis.

Table 3.7: Effect of Local Media on Local Trading

	Weekly Local Trades at MSA i for Stock j		
	Keyword General News IV	Earnings News	
	(1)	(2)	(3)
Concurrent N_{ijt}	0.458***	0.066***	0.118***
	(0.054)	(0.010)	(0.019)
1 Week Lag N_{ijt}	0.132**	0.019*	0.036***
	(0.061)	(0.010)	(0.013)
2 Week Lag N_{ijt}	0.117	0.01	0.003
	(0.090)	(0.010)	(0.011)
3 Week Lag N_{ijt}	0.102	0.020*	0.012
	(0.101)	(0.010)	(0.011)
4 Week Lag N_{ijt}	0.119	0.004	0.007
	(0.108)	(0.011)	(0.011)
5 Week Lag N_{ijt}	0.144	0.002	0.001
	(0.091)	(0.011)	(0.012)
6 Week Lag N_{ijt}	0.123**	−0.021**	−0.010
	(0.062)	(0.011)	(0.010)
Fixed Effects	MSA × Yr/Mon, Stock × Week, MSA × Stock	MSA × Week, Stock × Week, MSA × Stock	MSA × Yr/Mon, Stock × Week, MSA × Stock
Observations	381,174	339,966	339,966
Adjusted R^2	0.024	0.312	0.264

Notes: Column (1) summarizes coefficient estimates δ in equation 3.3. Columns (2) and (3) report coefficient estimates for equation 3.1 using earnings news article set. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

Table 3.7 summarizes the results. Column (1) reports the coefficient estimates for the 2SLS analysis, and columns (2) and (3) report the coefficient estimates for the analysis using the earnings news article set. Column (2) includes MSA interacted with week fixed effects and column (3) includes MSA interacted with year-month fixed effects. Across all of the specifications, local media appear to have a strong effect on trading behavior. As column (1) shows, a 1-standardized-unit increase in earnings announcement-predicting local news articles leads to an increase in local household trades of about 46%. This number quickly decreases to 13% 1 week after the announcement and remains mostly insignificant afterward. There is a slight increase of 12% 6 weeks after the increase in local media.

The coefficient estimates are much smaller in the analysis using earnings news article set compared to that using the 2SLS regression results. Part of the difference can be attributed to the difference in the fixed effects. The immediate effects of earnings-related local news with MSA interacted with year-month fixed effects appear to be twice as strong as those with MSA interacted with week fixed effects. However, even after accounting for the differences in fixed effects, there appears to be a great difference in the magnitude of the local media effect between the 2SLS regression result and the earnings news analysis result. This may be due to attenuation bias in earnings news analysis. Given that I used an ad hoc algorithm to parse out earnings-related news, there may be a great degree of noise[15], leading to attenuation bias.

I have previously suggested that the causal effect of local media on households' trading behavior can be explained mainly through salience and information. The underlying idea behind salience is that regardless of the content, a related article raises salience for a stock; hence, β_k in equation 3.1 would be positive under the salience hypothesis. Under the information hypothesis, β_k may be positive or negative[16]. As a result, the regression result in this section is consistent with both the salience and information hypotheses. In the next section, I test between the two hypotheses by considering informational earnings announcements and non-informational earnings announcements.

[15] For example, there may be earnings-announcement-related news without the word "earnings" used in the body of the text.

[16] This chapter does not take a stance on what the information channel may be.

3.5 Dissecting the Local Media Effect

In this section, I present evidence that the local media effect is largely driven by the salience channel. To do so, I isolate earnings announcements into informational and non-informational earnings announcements. Earnings announcements are ideal in this setting because market estimates of the earnings are readily available through the analyst earnings forecasts. Consider an earnings announcement that falls well within analyst expectations. If the analyst forecasts and relative timing of the earnings announcement are readily available ahead of the actual announcement, and if the actual earnings announcement is close to the analyst forecasts, there is no difference in the information set of households that receive local news covering the earnings announcement from that of households that do not receive the local news[17]. Hence, if there were a difference in trading behavior between MSAs with local news and MSAs without local news when earnings announcements were close to the analyst consensus, the salience channel would be probable[18]. Conversely, the lack of significance would suggest that an information channel is probable.

The earnings-related news is divided into informational and non-informational news according to the following rule. I label news related to earnings announcements that fell within 0.25[19] standard deviations of the mean analyst forecasts as non-informational (no shock) news. Similarly, I label news related to earnings announcements that fell outside of 1.5 standard deviations of the mean analyst forecasts as informational news. I also distinguish between positive shock news (earnings announcement higher than 1.5 standard deviations of the mean) and negative shock news (earnings announcement lower than 1.5 standard deviations of the mean). For this analysis, I use the keyword general news set to run 2SLS regressions. Of the 323 earnings announcements from

[17]One may argue that under my assumptions of informational frictions for households, households may have access to the analyst forecasts. In this scenario, the analysis setting can also be motivated by assuming that the analyst forecasts proxy for the market expectation.

[18]One may argue that confirming that the actual earnings are identical to the analyst consensus is information. I do not distinguish between this and salience

[19]I cannot choose a standard deviation smaller than 0.25 because of the statistical power issue. However, analysis with 0.5, 0.75, and 1 standard deviation does not affect the result materially.

34 stocks in my data set, there are 63 no-shock earnings announcements from 26 stocks. There are 76 positive shock earnings announcements from 27 stocks and 23 negative shock earnings announcements from 16 stocks.

3.5.1 Earnings Announcement and Local Stock News

In the previous section, I document the jump in local stock news around earnings announcements. Before running 2SLS analysis using informational and non-informational earnings announcements, I must show that local media react to both types of earnings announcements to test for both types of announcements' validity as instruments. The regression specification is identical to Equation 3.2.

Table 3.8: Earnings Announcement Week and Stock Related News - Announcement Type

	Weekly Number of News Articles Related to Stock j		
	No Shock (1)	Positive Shock (2)	Negative Shock (3)
Ann. Week	8.135^{***}	12.938^{***}	10.451^{***}
	(1.945)	(1.832)	(3.323)
Expected Ann. Week	-2.246	2.039	1.927
	(1.952)	(1.820)	(3.324)
Fixed Effects	Stock × Yr/Mon, Week	Stock × Yr/Mon, Week	Stock × Yr/Mon, Week
Observations	10,676	10,676	10,676
Adjusted R^2	0.694	0.696	0.694

Notes: This table summarizes coefficient estimates for equation 3.2. The parentheses include the standard errors. * represents $p<0.1$, ** represents $p<0.05$, and *** represents $p<0.01$.

Table 3.8 summarizes the regression results. Column (1) documents the results using non-informational news (earnings that fell within 0.25 standard deviations of the mean). Similarly, columns (2) and (3) document the results for positive shock and negative shock news, respectively.

Column (1) shows that there were eight more stock-related news articles nationally during the week of earnings announcements compared to other weeks. The jump is highest for positive shock earnings announcements, with 13 more articles during the week of earnings announcements.

There is a strong correlation between all three types of earnings announcements and stock-related news, which suggests that both the informational and non-informational earnings announcements are suitable as instruments.

3.5.2 Information vs. Salience

In this section, I use the variations in local stock news predicted by informational and non-informational earnings announcements to test whether the salience or informational channel drives the local media effect. I use 2SLS regression with a regression setup for the 2SLS identical to that in Equation 3.3.

Table 3.9 summarizes the regression results. Column (1) contains the results using no-shock earnings announcements (actual earnings within 0.25 standard deviations of analyst forecasts). Column (2) contains positive shock results and column (3) contains negative shock results.

As column (1) shows, a 1-standardized-unit increase in local stock news articles predicted by no-shock earnings announcements leads to an immediate 60% increase in the local trade of the particular stock compared to other MSAs without the increase in news. This effect appears to decay mostly after week 2. The strong response to increased stock news following no-shock earnings announcements is consistent with salience, as argued previously. The effects are similar for positive shock earnings announcements, but with smaller magnitudes. A 1-standardized-unit increase in local stock news articles predicted by positive shock earnings announcements leads to an immediate 35% increase in the local trade of the particular stock. This increase fades to about 17% after 2 weeks.

The pattern is slightly different for negative shock news. As column (3) shows, a 1-standardized-unit increase in local stock news articles predicted by no-shock earnings news leads to an immediate 82% increase in the local trade of the particular stock. However, this jump reverses almost right away, as the increase in stock news leads to a 44% decrease in local trades two weeks afterward. The differences in households' reactions to local newspaper article increase related to no-shock

Table 3.9: Effect of Local Media on Trading (Earnings Announcement Type)

	Weekly Local Trades at MSA i for Stock j		
	No Shock	Positive Shock	Negative Shock
	(1)	(2)	(3)
Concurrent N_{ijt}	0.606***	0.346***	0.819***
	(0.168)	(0.066)	(0.217)
1 Week Lag N_{ijt}	0.340	0.191***	−0.182
	(0.210)	(0.054)	(0.173)
2 Week Lag N_{ijt}	0.506**	0.166**	−0.438**
	(0.204)	(0.081)	(0.187)
3 Week Lag N_{ijt}	0.267	−0.042	0.017
	(0.262)	(0.071)	(0.200)
4 Week Lag N_{ijt}	0.101	0.014	0.463*
	(0.270)	(0.097)	(0.243)
5 Week Lag N_{ijt}	−0.061	0.046	0.464**
	(0.367)	(0.078)	(0.226)
6 Week Lag N_{ijt}	−0.123	0.087	0.126
	(0.415)	(0.054)	(0.161)
Fixed Effects	MSA × Week, Stock × Yr/Mon, MSA × Stock	MSA × Week, Stock × Yr/Mon, MSA × Stock	MSA × Week, Stock × Yr/Mon, MSA × Stock
Observations	381,174	381,174	381,174
Adjusted R^2	−0.427	0.119	−0.790

Notes: This table summarizes coefficient estimates for equation 3.3. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

earnings news and negative shock earnings news suggests that the informational channel may also play a role. However, the difference can also be explained by salience and stock price movements following earnings announcement shocks.

Lastly, I consider how local media affect the proportion of households that place a buy or a sell order. Table 3.10 summarizes the 2SLS regression results for buy trades, while Table 3.11 summarizes the results for sell trades. Consistent with the results in Table 3.9, there is a spike in the proportion of households that place buy trades immediately following no-shock-, positive shock-, and negative shock-related local news. However, the initial jump for the no-shock case is

Table 3.10: Effect of Local Media on Buy Trades

	Weekly Local Trades at MSA i for Stock j		
	No Shock (1)	Positive Shock (2)	Negative Shock (3)
Concurrent N_{ijt}	0.386*** (0.145)	0.340*** (0.057)	0.797*** (0.183)
1 Week Lag N_{ijt}	0.375** (0.184)	0.139*** (0.048)	−0.065 (0.146)
2 Week Lag N_{ijt}	0.263 (0.178)	0.124* (0.070)	−0.131 (0.173)
3 Week Lag N_{ijt}	0.298 (0.225)	−0.039 (0.063)	0.204 (0.187)
4 Week Lag N_{ijt}	0.199 (0.226)	0.027 (0.086)	0.487** (0.223)
5 Week Lag N_{ijt}	0.283 (0.314)	0.080 (0.068)	0.299 (0.198)
6 Week Lag N_{ijt}	0.213 (0.353)	0.003 (0.047)	0.032 (0.139)
Fixed Effects	MSA × Week, Stock × Yr/Mon, MSA × Stock	MSA × Week, Stock × Yr/Mon, MSA × Stock	MSA × Week, Stock × Yr/Mon, MSA × Stock
Observations	370,872	370,872	370,872
Adjusted R^2	−0.544	0.082	−0.916

Notes: This table summarizes coefficient estimates for equation 3.3. Buy trades are first isolated in LDB data and aggregated across MSAs. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

almost identical to the positive shock case. The initial jump in the proportion of households that place buy trades in the negative shock case is nearly 80%, which is higher than the jump for the no-shock and positive shock cases. A potential reason for the higher jump in the negative shock case is that stock prices tend to drop by a significant margin immediately following negative shock earnings announcements, and these lower prices might prompt households in salient MSAs to buy more.

The results are similar to the buy trade case, with two differences. First, although there are

Table 3.11: Effect of Local Media on Sell Trades

	Weekly Local Trades at MSA i for Stock j		
	No Shock	Positive Shock	Negative Shock
	(1)	(2)	(3)
Concurrent N_{ijt}	0.347**	0.204***	0.239
	(0.141)	(0.051)	(0.158)
1 Week Lag N_{ijt}	−0.021	0.089**	−0.184
	(0.176)	(0.043)	(0.131)
2 Week Lag N_{ijt}	0.310**	0.126*	−0.474***
	(0.155)	(0.065)	(0.144)
3 Week Lag N_{ijt}	−0.009	0.004	−0.213
	(0.203)	(0.058)	(0.155)
4 Week Lag N_{ijt}	−0.100	0.039	0.003
	(0.202)	(0.080)	(0.189)
5 Week Lag N_{ijt}	−0.472	0.007	0.255
	(0.288)	(0.063)	(0.171)
6 Week Lag N_{ijt}	−0.391	0.140***	0.106
	(0.301)	(0.043)	(0.122)
Fixed Effects	MSA × Week, Stock × Yr/Mon, MSA × Stock	MSA × Week, Stock × Yr/Mon, MSA × Stock	MSA × Week, Stock × Yr/Mon, MSA × Stock
Observations	329,664	329,664	329,664
Adjusted R^2	−0.603	0.090	−0.383

Notes: This table summarizes coefficient estimates for equation 3.3. Sell trades are first isolated in LDB data and aggregated across MSAs. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

initial increases in the proportions of households that place sell trades in the no-shock and positive shock cases, the coefficient estimate for negative shocks without a lag is not statistically significant. Second, although there are somewhat sustained increases in the proportions of households that place sell trades up until 2 weeks afterward for the no-shock and positive shock cases, there appears to be a decrease in the proportion of households that place a sell trade for the negative shock case. This may be due to depressed stock prices following a negative earnings shock and the disposition effect deterring households from selling losing stocks.

3.6 Conclusion

Information is costly, and agents have differing abilities to process it. Although institutions with billions of dollars in assets under management may have high information-processing abilities, making information friction less of an issue, the same may not be true of an average household investor. In this chapter, I use geographic proximity as a proxy for ease of access in defining local media during a period in which the Internet was not widely available.

I start my analysis by documenting the predictive power of local media on local trades. Specifically, the mention of a stock by local media increases the likelihood of trades of the particular stock by local households. Although the main analyses in this chapter are conducted using weekly horizons, the results in Section 5.B suggest that the local media effect may persist for several months after the news. In addition, using earnings announcement dates as an instrumental variable, I run 2SLS regressions and find evidence that the local media effect may be causal.

Lastly, having separated the earnings-related news into informational and non-informational news, I find evidence that the local media effect may be largely driven by the salience channel. The evidence in this chapter suggests that a stock's mention in local media matters more than the context for the effect on household trading behavior. Given that this study focuses on local media rather than national major media, the mechanism through which major media affect the stock market may be different.

Information propagation mechanisms are now very different from those that were prevalent during 1991-1996. With the rise of the Internet, news from faraway regions can be easily accessed. As such, geographically local media may no longer have the same causal effect on households' trading behavior. However, the underlying factors behind this effect, such as the cost of information processing, are likely to remain valid. As such, it may be useful for future studies to focus on modern forms of "local media" that households may use to obtain information.

CHAPTER 4

Housing Market Appendix

Appendix A. Local Housing Media and New Construction

In this section, I provide evidence that new construction may cause the local media effect to dissipate after about 9 months. I show this by documenting the increased number of building permit approvals 3-5 months after an increase in local housing news. As the construction of new units takes about 6 months on average, an increase in the housing supply 9-11 months later may explain why the local housing news effect dissipates after 9-11 months.

Building permits take 2 months to be approved on average. Hence, the correlation between the lagged increase in local housing news and building permit approval may be due to construction companies responding either to increased household home searches or to the local news shock. For example, construction companies receiving heterogenous signals from local media that are only able to respond to a positive housing shock provide an analogous result to the model described in Section 2.2.

The regression specification is similar to that used before, with several changes. The specification is as follows.

$$NP_{it} = F.E. + \sum_{k=1}^{12} \beta_k MI_{it-k} + \sum_{k=1}^{12} \gamma_k^1 R_{it-k} + \sum_{k=1}^{12} \gamma_k^2 NP_{it-k} + \varepsilon_{it} \tag{A1}$$

The dependent variable is the change in New Permits (NP) for county i at month t. The source of new permit data is shown in Table 2.1. The numbers of permit grants differ across different counties, so I divide by the whole sample average to standardize the data. As many of the data points are imputed, data for counties with few new permits are likely to be noisier than data for counties with more construction. As such, I use only data from counties with at least 100 monthly average permit grants in the sample. Matching the newspaper and building permit data yields 106 counties. As before, MI_{it} is the increase in housing-related media for county i at month t. I include up to 12-month lags of the local housing return R_{it} and new permits as controls.

Table 4.1: Lagged Local Housing Media on Building Permit Changes

	Change in New Building Permit			
	All	2004-2007	2008-2011	2012-2015
	(1)	(2)	(3)	(4)
Lagged 1 Month	−0.003	−0.012	0.002	−0.005
	(0.009)	(0.017)	(0.014)	(0.021)
Lagged 2 Month	−0.004	0.005	−0.006	−0.010
	(0.011)	(0.020)	(0.016)	(0.025)
Lagged 3 Month	0.024**	0.040*	0.011	0.019
	(0.011)	(0.021)	(0.017)	(0.026)
Lagged 4 Month	0.027**	0.024	0.032*	0.032
	(0.011)	(0.021)	(0.017)	(0.027)
Lagged 5 Month	0.018*	0.032	0.006	0.013
	(0.011)	(0.021)	(0.017)	(0.027)
Lagged 6 Month	0.013	0.039*	0.008	−0.007
	(0.011)	(0.021)	(0.017)	(0.027)
Lagged Returns	✓	✓	✓	✓
Lagged New Permits	✓	✓	✓	✓
Fixed Effects	County, Yr/Mon, MSA × Yr/Qtr	County, Yr/Mon, MSA × Yr/Qtr	County, Yr/Mon, MSA × Yr/Qtr	County, Yr/Mon, MSA × Yr/Qtr
Observations	18,974	18,974	18,974	
Adjusted R^2	0.418	0.417	0.429	

Notes: Table 4.1 predictive power of media increase on subsequent new permit approvals. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

Table 4.1 summarizes the regression results. Due to space constraints, I report only β_k for up to 6 months, but β_ks for 7- to 12-month lags are statistically indifferent from 0. Column (1) summarizes the result for the whole sample (2000-2015), and columns (2), (3), and (4) summarize the results for the subsamples. Looking first at the whole sample, media increase appears to have strong predictive power for subsequent changes in the number of new building permits. Increases in new building permits are 1.8-2.4% higher for counties that had increases in housing news articles 3-5 months prior.

Considering the subsamples, the predictive power of the media on new building permits appears to be strongest in the 2004-2007 sample. The change in the number of new building permits appears to be 4% higher for counties that had increased housing news 3 months prior. In the 2008-2011 sample, the change in the number of new building permits is 3.2% higher for counties with increased housing news 4 months prior. The significance disappears in the 2012-2015 sample. The predictive relationship is strongest in 2004-2007 and weaker in subsequent periods, which is consistent with the media effect reversion reaching its strongest state during the 2003-2007 period in Table 2.3.

Appendix B. Decomposing Housing Media

In this section, I explore an extension in which I decompose the housing media into positive and negative news via text analysis. Although this is an informative exercise, because of the noisiness of the decomposition due to my access to only the headlines, this exercise is included only in the appendix.

Appendix B.1. Classifying Articles

The biggest challenge in this analysis is classifying the articles into positive and negative categories. As mentioned previously, because of my institution's contract with NewsBank, I do not have access to the body of sample articles[1]. Instead, I rely on the headlines to determine the content of each article.

To objectively classify headlines as positive or negative, I follow the text-mining approach used by Tetlock (2007). I start with the Harvard IV-4 (HIV4) Dictionary of positive and negative words. As highlighted by numerous studies using this approach (Engelberg, 2008; Tetlock et al., 2008), using the HIV4 dictionary without any supplement can lead to inaccurate and noisy classifications. To increase the accuracy of my headline classification, I follow an approach suggested by Loughran and McDonald (2011)[2] and manually add the most used words that are likely to be used in positive and negative housing market contexts. Ultimately, a headline is classified as positive or negative depending on whether it contains more positive or negative words. Table 4.2 includes my addendum to the HIV-4 dictionary.

I limit the sample of articles to those with "home," "house," or "housing" in the headlines to

[1] Downloading article bodies for over 200,000 articles would require webscraping, which the contract prohibits

[2] Loughran and McDonald (2011) does not actually use this approach because it may lead to endogeneity problems. In the context of corporate finance, managers may decide to consider these results in future annual reports. However, this problem is not a concern with media and household behaviors.

Table 4.2: Words added to Harvard IV-4 Dictionary

Positive Words	Negative Words
up, climb, soaring,	down, worst, slowly,
rise, highest, highs,	fall, bottom, lose,
increase, increases, advance,	drop, lowest, dropped,
rising, positive, surges,	decline, downturn, slower,
higher, grow, recovering,	slow, tumble, lows,
jump, improve, increasing,	lower, bust, dips,
rebound, boosts, gains,	slump, concerns, loses,
strong, grows, upagain,	falls, slowdown, slips,
better, spike, hope,	bad, slows, dropping,
surge, hike, stabilize,	dip, declines, cloudy,
gain, increased, top-mover,	weak, distressed, ominous,
rises, upward, growing,	less, meltdown, gloom,
rose, recover, life,	falling, slowing, repossessions,
rally, improves	worries, worse, difficult,
	plunge, struggling, decrease,
	fell, sluggish

Note: This table lists words that are added to Harvard IV-4 dictionary to classify sentences into positive and negative sentences. These are words that appear most frequently in the headline dataset that are manually determined likely to have positive or negative meaning when used in housing market context.

filter out headlines such as "Online sales grow by 19% - Amazon.com leads retailers"[3], which has a positive sentiment but does not provide information about the housing market.

Table 4.3 lists a random sample of 25 selected headlines and their classifications. Of the 25 headlines, 16 include the terms "home," "house," or "housing" and are hence included in my sample. Although not perfect, the classification method appears to perform adequately. Aside from one error ("House sales decrease; unemployment rate rises") and several ambiguous cases (e.g., "Real estate sales slide in Myrtle Beach area - Home prices rise; condos continue to fall,"), classifications based on this method appear to be more or less accurate.

Appendix B.2. Positive and Negative Articles on Housing Returns

As with previous analyses, I use the panel time-series regression with controls for lagged returns and region-time fixed effects. Table 4.4 summarizes the regression results. As with Table 2.3, I divide the analysis into the whole sample and three different subsamples. The first panel reports the estimates for increases in positive housing articles. The second panel reports the estimates for increases in negative housing articles.

As seen in the column (1), an increase in positive news leads to an increase in housing returns. Although there seem to be similar effects for negative news, the coefficients are not statistically significant. In contrast to previous analysis, positive news seems to be strongest during the first quarter rather than the second quarter. Given the heterogeneity in home-buying process times, this result is not surprising. Considering the subsamples, there appears to be a strong effect from positive housing news during the housing crisis (column (3)) and post-crisis (column (4)). However, during these periods, negative housing news has no effect.

During the housing boom period (column 2), I find no effect for positive housing news. Instead, negative housing news has a strong effect during Quarter 3. The coefficient suggests that an increase in negative housing news during the housing boom is correlated with a 1.6 bp higher housing return 7-10 months afterward. Although negative news leading to higher returns may seem counterintuitive, it may indicate a reversal of the decline in housing returns accumulated during months 1-6.

While lacking statistical significance, these results are consistent with the implications of the model described in 2.2. Namely, positive news is likely to induce households to buy homes, positively affecting the local hpi, while negative news is less likely to affect the local hpi due to shorting constraints. At the very least, this result suggests that future work dissecting into the contents of

[3]"Online sales grow by 19% - Amazon.com leads retailers" talks about the sluggish housing market as a factor of consumer spending. Hence, if sentiment analysis on the article body were to be run, it would be classified as a negative housing article. For this analysis, the article is excluded.

local news articles may be beneficial.

Table 4.3: Sample Headlines

Headline	Positive Words	Negative Words
Clean Clicks vs. Dirty Data What would it take to build an environmentally sustainable internet? - Insurers and financial corporations are pushing for new laws and policies to fight climate change before it hits their bottom lines. It might not be altruism, but it just might work.	0	1
US homebuilders' confidence in sales surges in June	1	0
Existing home sales teeter in September - 1.9% drop blamed on higher prices and interest rates	1	1
Reel logic: Choosing a fishing reel boils down to skill level and application choices	0	1
US home prices up 9.3 pct., most in nearly 7 years	1	0
Oil falls as growth worries outweigh positive news	1	2
Rise in US home sales reflects steady improvement	1	0
Rise in home building suggests industry turnaround	1	0
Survey: Home prices increase in half of major U.S. cities	1	0
Mass. housing prices firm, sales fall	0	1
Real estate sales slide in Myrtle Beach area - Home prices rise; condos continue fall	1	1
BROWARD HOME SALES, PRICES DOWN IN THIRD QUARTER	0	1
Banks put up roadblocks to low mortgages	1	0
CHOPPY MARKET MOVES HIGHER AS DOW RISES 75	2	0
Real estate recovering, albeit slowly - In Darien	1	1
One step ahead - Real estate agent helps struggling homeowners in poor economyFACES OF IDAHO	0	1
Existing-home sales rise, fueling hopes of recovery Rebound in national market seen as critical step in exiting recession	2	0
New-home sales up 11% - Low prices, tax credits boost increase	2	0
Median home prices fall in 88 percent of cities	0	1
Stocks tumble as oil falls on economic worries	0	3
Analysts adjust oil-cost forecast - Global financial crisis is forcing them to lower their initially high price predictions	0	1
Coldwell Banker to drop home prices for 10 days in Valley - 173 houses will be discounted 5-12% in the latest promotion by a real estate company or builder.	0	1
House sales decrease; unemployment rate rises	1	1
Weak sales, write-downs send homebuilder Lennar to 1Q loss	0	1
Dixie sales tumble; profit slips	0	2

Note: This table lists 25 randomly picked headlines that has at least one positive or negative word. Postive and negative words attached to the Harvard IV-4 dictionary are listed in Table 4.2

Table 4.4: Increase in Housing Article (Positive / Negative) on Subsequent Housing Return

		1 Month Local Housing Return:			
		All	2003-2007	2008-2011	2012-2015
		(1)	(2)	(3)	(4)
Positive Article Increase	Quarter 1	0.006**	−0.0004	0.015**	0.006
		(0.003)	(0.005)	(0.006)	(0.006)
	Quarter 2	0.003	−0.007	0.004	0.013**
		(0.003)	(0.005)	(0.006)	(0.006)
	Quarter 3	−0.002	−0.005	−0.007	0.003
		(0.003)	(0.005)	(0.006)	(0.006)
	Quarter 4	0.001	0.001	−0.011*	0.013**
		(0.003)	(0.005)	(0.006)	(0.006)
Negative Article Increase	Quarter 1	−0.003	−0.008	−0.006	0.002
		(0.003)	(0.005)	(0.006)	(0.007)
	Quarter 2	−0.002	−0.001	−0.006	0.005
		(0.003)	(0.005)	(0.006)	(0.007)
	Quarter 3	0.002	0.016***	−0.002	−0.008
		(0.003)	(0.005)	(0.006)	(0.006)
	Quarter 4	−0.001	−0.001	−0.006	0.005
		(0.003)	(0.006)	(0.006)	(0.006)
	Lagged Housing Returns	✓	✓	✓	✓
	Fixed Effects	MSA ×Yr/Qtr	MSA ×Yr/Qtr	MSA ×Yr/Qtr	MSA ×Yr/Qtr
	Observations	46,226	13,920	11,136	11,136
	Adjusted R^2	0.907	0.925	0.885	0.856

Notes: Table 4.4 reports effect of increase in positive and negative articles on subsequent local housing return in different periods. * represents $p<0.1$, ** represents $p<0.05$, and *** represents $p<0.01$.

CHAPTER 5

Stock Market Appendix

Appendix A. Stocks Analyzed

The ticker codes and company names for the 39 stocks used are as follows. There are 34 stocks used in the keyword general news article set, 33 stocks used in the earnings news article set, and 13 stocks used in the ticker general news article set. Table 5.1 lists all of the stocks included in this study and the datasets that contain them.

Table 5.1: Stocks Analyzed and Their Tickers

Ticker	Company Name	Ticker General	Keyword General	Earnings
AAPL	APPLE COMPUTER INC	✓	✓	✓
ABT	ABBOTT LABORATORIES		✓	✓
AMGN	AMGEN INC		✓	✓
BA	BOEING CO		✓	✓
BMY	BRISTOL MYERS SQUIBB CO	✓	✓	✓
C	CHRYSLER CORP		✓	✓
CCI	CITICORP	✓		
CPQ	COMPAQ COMPUTER CORP	✓		
CSCO	CISCO SYSTEMS INC	✓	✓	✓
DEC	DIGITAL EQUIPMENT CORP		✓	✓
DELL	DELL COMPUTER CORP		✓	✓
F	FORD MOTOR CO DEL		✓	✓
GE	GENERAL ELECTRIC CO		✓	✓
GTE	G T E CORP		✓	✓
HAN	HANSON PLC		✓	✓
HWP	HEWLETT PACKARD CO	✓		
IBM	INTERNATIONAL BUSINESS MACHS CO		✓	✓
INTC	INTEL CORP	✓	✓	✓
JNJ	JOHNSON & JOHNSON		✓	✓
KM	K MART CORP		✓	✓
KO	COCA COLA CO		✓	✓
LLY	LILLY ELI & CO		✓	✓
MCD	MCDONALDS CORP		✓	✓
MCIC	M C I COMMUNICATIONS CORP		✓	
MO	PHILIP MORRIS COS INC		✓	✓
MOT	MOTOROLA INC	✓		
MRK	MERCK & CO INC	✓	✓	✓
MSFT	MICROSOFT CORP	✓	✓	✓
NOVL	NOVELL INC	✓	✓	✓
OXY	OCCIDENTAL PETROLEUM CORP		✓	✓
PAC	PACIFIC TELESIS GROUP		✓	✓
PEP	PEPSICO INC		✓	✓
RN	RJR NABISCO HOLDINGS CORP		✓	✓
S	SEARS ROEBUCK & CO		✓	✓
SYN	SYNTEX CORP		✓	✓
TMX	TELEFONOS DE MEXICO S A DE C V	✓		
WMT	WAL MART STORES INC	✓	✓	✓
WX	WESTINGHOUSE ELECTRIC CORP		✓	✓
XON	EXXON CORP		✓	✓

Appendix B. Long-Term Effects of Local Media

Weekly aggregation is used during the main analysis based on the assumption that local media quickly affect households' trading behavior. To test for any lasting local media effects, I replicate the analysis with monthly aggregation. Although the 2SLS/IV framework is harder to implement in monthly aggregations because the regression results in Table 3.6 suggest that the local media attention given to earnings announcements is short-lived. Thus, there would not be sufficient variation by earnings announcement month. However, I replicate the results using the earnings news article dataset.

The regression specification for predictive regression is identical to equation 3.1:

$$Y_{ijt} = F.E. + \sum_{k=0}^{K} \beta_k N_{ijt-k} + \varepsilon_{ijt} \tag{B1}$$

The main difference is that Y_{ijt} and N_{ijt} are aggregated monthly rather than weekly. Y_{ijt} is the total number of households that trade stock j in year-month t from MSA i divided by the sample mean. Similarly, N_{ijt} is the total number of local news articles from MSA i related to stock j in year-month t. MSA interacted with the year-month fixed effect, Stock interacted with the year-month fixed effect, and MSA interacted with stock fixed effects are included.

Table 5.2 summarizes the regression results. Column (1) summarizes the coefficient estimates for β_k using the ticker general news article set. Similarly, column (2) summarizes the coefficient estimates using the keyword general news article set. As expected, because of the strong local media effect in weekly aggregation, there is a strong concurrent effect of local media on subsequent local trades. As column (1) shows, a 1-standardized-unit increase in a stock-related local news article predicts a 7.5% increase in the local trade of the particular stock in the concurrent month. The corresponding increase according to column (2) is 1.1%.

In both datasets, the predictive power of local media appears to last at least until two months

Table 5.2: Local News and Subsequent Trading - Monthly

	Monthly Local Trades at MSA i for Stock j		
	Ticker General (1)	Keyword General (2)	Earnings (3)
Concurrent N_{ijt}	0.075^{***}	0.011^{***}	0.010
	(0.026)	(0.004)	(0.014)
1 Month Lag N_{ijt}	0.038	0.002	0.013
	(0.026)	(0.004)	(0.014)
2 Month Lag N_{ijt}	0.061^{**}	0.009^{**}	-0.011
	(0.027)	(0.004)	(0.014)
3 Month Lag N_{ijt}	0.056^{**}	0.005	0.028^{**}
	(0.028)	(0.004)	(0.014)
4 Month Lag N_{ijt}	0.080^{***}	-0.003	-0.026^{*}
	(0.028)	(0.004)	(0.014)
5 Month Lag N_{ijt}	-0.006	-0.002	-0.017
	(0.028)	(0.004)	(0.015)
6 Month Lag N_{ijt}	0.057^{**}	-0.003	0.006
	(0.029)	(0.004)	(0.015)
Fixed Effects	MSA $\times$ Yr/Mon, Stock $\times$ Yr/Mon, MSA $\times$ Stock	MSA $\times$ Yr/Mon, Stock $\times$ Yr/Mon, MSA $\times$ Stock	MSA $\times$ Yr/Mon, Stock $\times$ Yr/Mon, MSA $\times$ Stock
Observations	27,885	81,770	70,720
Adjusted R^2	0.429	0.380	0.436

Notes: This table summarizes coefficient estimates for equation B1. The parentheses include the standard errors. * represents p<0.1, ** represents p<0.05, and *** represents p<0.01.

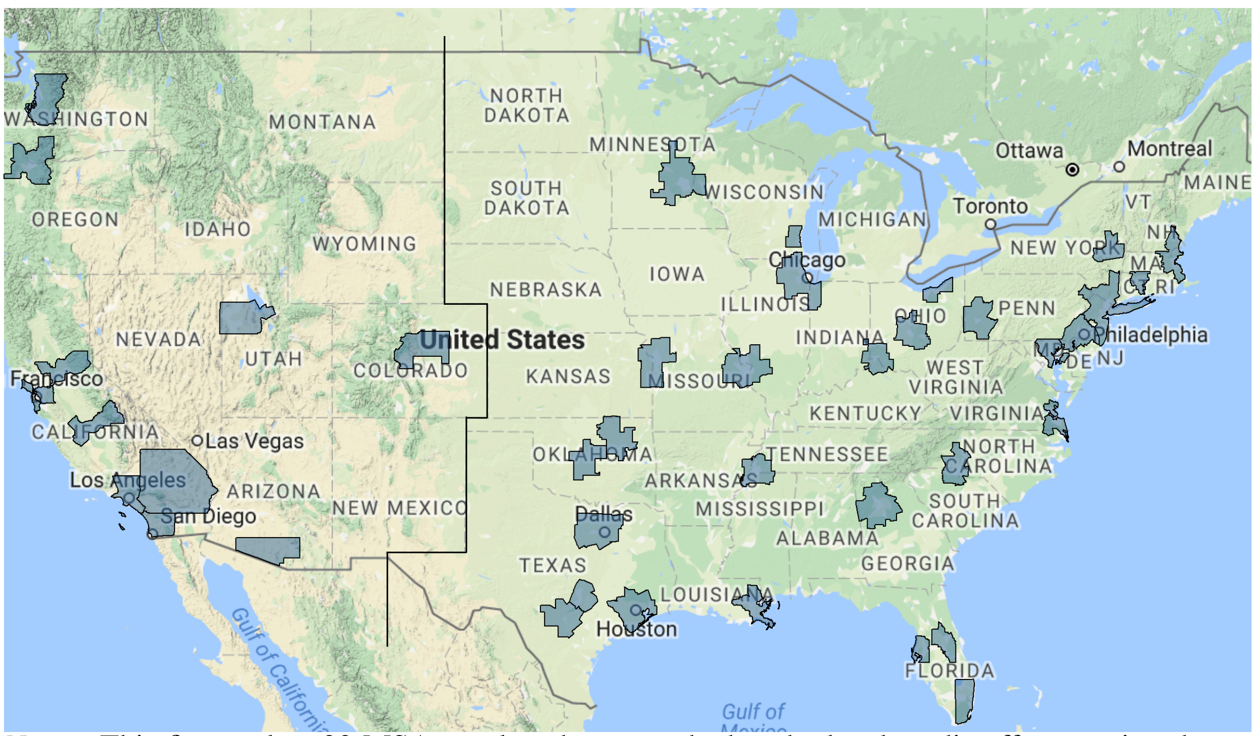
after the increase in local news articles. As column (1) shows, a 1-standardized-unit increase in a stock-related local news article predicts a 6.1% increase in the local trade of the particular stock 2 months afterward. The corresponding increase according to column (2) is 0.9%. The coefficient estimates using the ticker general news article set appear to be more persistent, with statistical significance lasting until 6 months afterward, and generally larger. This may be explained as follows: the stocks included in the ticker general news article set are much more widely and actively traded than stocks included in the keyword general news article set.

Column (3) summarizes the uses coefficient estimates using the earnings news article set. As described previously, the earnings news article set contains local news articles related to arguably exogenous earnings announcement dates. Hence, variation in the earnings news article set is plausibly exogenous. In contrast to the strong initial results in the weekly aggregation, there does not appear to be an increase in local trades related to the increase in local news articles on earnings announcements. This may be due to the overall increase in households' trading in months that include earnings announcements and the short-lived nature of the local media effect. For example, as Table 3.7 shows, the effect of earnings-announcement-related local media decays after 2 weeks; thus, averaging out the effect over 4 weeks yields smaller coefficient estimates. Consistent with the results in columns (1) and (2), the local media effect appears to be persistent, as evidenced by the statistical significance 3 months after earnings-announcement-related local news.

Appendix C. Difference from Home Bias

In this section, I provide evidence that the local media effect is not entirely driven by home bias. The potential relationship between home bias and the predictive power of local media is as follows. Studies of home bias document households' preference for local stocks, which comprise the overwhelming portion of investors' portfolios. Although not documented in the home bias literature, it is conceivable that local newspapers cover local firms more heavily. If this is the case, most of the variation in the data may be due to households responding to local news on local stocks, and the positive correlation may merely be a result of home bias.

Figure 5.1: East-West Division Line and MSAs Analyzed



Notes: This figure plots 33 MSAs analyzed to test whether the local media effect survives home bias. The black line in the figure divides the households into east and west.

To address this concern, I divide the household and company samples into two geographic regions: east and west. The thick black line in Figure 5.1 separates the states into east and west. To test whether the local media effect survives home bias, I test whether local articles by west (east) newspapers of firms headquartered in the east (west) affect the trading behavior of households in the west (east). For example, if households in Atlanta (east) receive local news about Apple (west) while households in Chicago (east) do not, I test whether households in Atlanta are subsequently more likely to trade Apple than households in Chicago.

The division line was chosen for ease of classification; moving the line to the left or right makes the division ambiguous. However, as a result of this division, the east has a much larger sample of 29 MSAs, while the west has 11 MSAs. The division of stocks is roughly equal, with four in the east and six in the west[1].

Table 5.3 summarizes the results. The specification is the same as equation B1. Column (1) documents the effect of eastern region stock news by western newspapers on western households. Similarly, column (2) documents the effect for eastern households. Each regression contains MSA interacted with year/month fixed effects, Stock interacted with year/month fixed effects, and MSA interacted with month fixed effects.

Consistent with the results in Appendix B, there appears to be a persistent local media effect even after accounting for home bias. In column (1), although not statistically different from 0, the coefficient estimate suggests that a 1-standardized-unit increase in local news articles about a stock in the east predicts a 20% increase in local trades by households in the west. Similarly, in column (2), the coefficient estimate suggests that a 1-standardized-unit increase in local news articles about a stock in the west predicts a 19% increase in local trades by households in the east. The stronger statistical significance in column (2) may a result of there being more MSAs in the east sample.

Although the exact timings are not identical, there appears to be evidence of a lasting local media effect in both columns (1) and (2). For example, local media coverage of non-local stocks leads to statistical significance 1 and 6 months afterward, as shown in column (1). Similarly, local media coverage of non-local stocks leads to statistical significance 2 and 4 months afterward, as

[1]The stocks headquartered in the east are BMY, CCI, MRK, and WMT. The stocks headquartered in the west are AAPL, CSCO, INTC, MSFT, and NOVL. TMX, which is headquartered in Mexico, is deemed non-local for both the east and west. CPQ, HWP, and MOT are not included in the analysis, as the division excludes all of the articles.

Table 5.3: Local Media and Subsequent Trading on Non-Local Stocks

	Monthly Local Trades at MSA i for Stock j	
	West Households	East Households
	(1)	(2)
Concurrent N_{ijt}	0.195	0.186**
	(0.176)	(0.078)
1 Month Lag N_{ijt}	0.369*	−0.134
	(0.201)	(0.086)
2 Month Lag N_{ijt}	0.245	0.169*
	(0.212)	(0.087)
3 Month Lag N_{ijt}	0.001	−0.061
	(0.184)	(0.094)
4 Month Lag N_{ijt}	0.026	0.266***
	(0.197)	(0.102)
5 Month Lag N_{ijt}	−0.107	0.040
	(0.196)	(0.098)
6 Month Lag N_{ijt}	0.332*	0.097
	(0.187)	(0.102)
Fixed Effects	MSA × Yr/Mon, Stock × Yr/Mon, MSA × Stock	MSA × Yr/Mon, Stock × Yr/Mon, MSA × Stock
Observations	1,826	4,770
Adjusted R^2	0.330	0.295

Notes: This table summarizes coefficient estimates for equation B1. Column (1) estimates the effect of local news by newspapers stationed in the west reporting about firms headquartered in the east on the local trading of households in the west. Column (2) is tests the opposite. The parentheses include the standard errors. * represents $p<0.1$, ** represents $p<0.05$, and *** represents $p<0.01$.

shown in column (2). Regardless, there is a statistically significant difference between households in MSAs without local news and households in MSAs with local news, which suggests that the local media effect is not entirely driven by home bias.

BIBLIOGRAPHY

Abrigo, Michael RM, Inessa Love, et al., 2015, Estimation of panel vector autoregression in stata: A package of programs, *manuscript, Febr 2015 available on http://paneldataconference2015. ceu. hu/Program/Michael-Abrigo. pdf* .

Anderson, Craig A, and Brad J Bushman, 2002, The effects of media violence on society, *Science* 295, 2377–2379.

Autor, David H, 2003, Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing, *Journal of Labor Economics, January* .

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, 2016, Social networks and housing markets, Technical report, National Bureau of Economic Research.

Bailey, Warren, Alok Kumar, and David Ng, 2011, Behavioral biases of mutual fund investors, *Journal of Financial Economics* 102, 1–27.

Baker, Malcolm, Jeffrey Wurgler, and Yu Yuan, 2012, Global, local, and contagious investor sentiment, *Journal of Financial Economics* 104, 272–287.

Barber, Brad M, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *The Journal of Finance* 55, 773–806.

Barber, Brad M, and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly journal of Economics* 261–292.

Barber, Brad M, and Terrance Odean, 2002, Online investors: do the slow die first?, *Review of Financial Studies* 15, 455–488.

Barber, Brad M, and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Barber, Brad M, and Terrance Odean, 2011, The behavior of individual investors, *Available at SSRN 1872211* .

Baron, David P, 2005, Competing for the public through the news media, *Journal of Economics & Management Strategy* 14, 339–376.

Begley, Joy, and Paul E Fischer, 1998, Is there information in an earnings announcement delay?, *Review of accounting studies* 3, 347–363.

Beracha, Eli, and M Babajide Wintoki, 2013, Forecasting residential real estate price changes from online search activity, *Journal of Real Estate Research* .

Bernard, Victor L, and Jacob K Thomas, 1989, Post-earnings-announcement drift: delayed price response or risk premium?, *Journal of Accounting research* 1–36.

Bhattacharya, Utpal, Neal Galpin, Rina Ray, and Xiaoyun Yu, 2009, The role of the media in the internet ipo bubble, *Journal of Financial and Quantitative Analysis* 44, 657–682.

Bodnaruk, Andriy, 2009, Proximity always matters: Local bias when the set of local companies changes*, *Review of finance* rfp014.

Brown, Gregory W, and Michael T Cliff, 2005, Investor sentiment and asset valuation, *The Journal of Business* 78, 405–440.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, 2016, Understanding booms and busts in housing markets, *Journal of Political Economy* 124, 1088–1147.

Bushee, Brian J, John E Core, Wayne Guay, and Sophia JW Hamm, 2010, The role of the business press as an information intermediary, *Journal of Accounting Research* 48, 1–19.

Cannon, Susanne, Norman G Miller, and Gurupdesh S Pandher, 2006, Risk and return in the us housing market: A cross-sectional asset-pricing approach, *Real Estate Economics* 34, 519–552.

Case, Karl, John Cotter, and Stuart Gabriel, 2011, Housing risk and return: Evidence from a housing asset-pricing model, *The Journal of Portfolio Management* 37, 89–109.

Chauvet, Marcelle, Stuart Gabriel, and Chandler Lutz, 2016, Mortgage default risk: New evidence from internet search queries, *Journal of Urban Economics* 96, 91–111.

Cho, Eric, 2016, Local media and the local housing market, *Working Paper* .

Cocco, Joao F, 2005, Portfolio choice in the presence of housing, *Review of Financial studies* 18, 535–567.

Cook, Fay Lomax, Tom R Tyler, Edward G Goetz, Margaret T Gordon, David Protess, Donna R Leff, and Harvey L Molotch, 1983, Media and agenda setting: Effects on the public, interest group leaders, policy makers, and policy, *Public Opinion Quarterly* 47, 16–35.

Cooper, Ian, and Evi Kaplanis, 1994, Home bias in equity portfolios, inflation hedging, and international capital market equilibrium, *Review of Financial Studies* 7, 45–60.

Coval, Joshua D, and Tobias J Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045–2073.

Cumming, Douglas, and Na Dai, 2010, Local bias in venture capital investments, *Journal of Empirical Finance* 17, 362–380.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *The Journal of Finance* 66, 1461–1499.

Diamond, Douglas W, and Robert E Verrecchia, 1987, Constraints on short-selling and asset price adjustment to private information, *Journal of Financial Economics* 18, 277–311.

Dyck, Alexander, and Luigi Zingales, 2002, The corporate governance role of the media, *The right to tell: The role of mass media in economic development* 107–37.

Dyck, Alexander, and Luigi Zingales, 2003, The media and asset prices, Technical report, Working Paper, Harvard Business School.

Edmans, Alex, Diego Garcia, and Øyvind Norli, 2007, Sports sentiment and stock returns, *The Journal of Finance* 62, 1967–1998.

Eisfeldt, Andrea, and Andrew Demers, 2015, Rental yields and hpa: The returns to single family rentals, Technical report, National Bureau of Economic Research.

Engelberg, Joseph, 2008, Costly information processing: Evidence from earnings announcements, in *AFA 2009 San Francisco Meetings Paper*.

Engelberg, Joseph E, and Christopher A Parsons, 2011, The causal impact of media in financial markets, *The Journal of Finance* 66, 67–97.

Fama, Eugene F, and Kenneth R French, 1992, The cross-section of expected stock returns, *the Journal of Finance* 47, 427–465.

Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of financial economics* 33, 3–56.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *The Journal of Finance* 64, 2023–2052.

Fang, Lily H, Joel Peress, and Lu Zheng, 2014, Does media coverage of stocks affect mutual funds' trading and performance?, *Review of Financial Studies* hhu056.

Favara, Giovanni, and Zheng Song, 2014, House price dynamics with dispersed information, *Journal of Economic Theory* 149, 350–382.

Frieder, Laura, and Avanidhar Subrahmanyam, 2005, Brand perceptions and the market for common stock, *Journal of financial and Quantitative Analysis* 40, 57–85.

Funkhouser, G Ray, 1973, Trends in media coverage of the issues of the 60s, *Journalism Quarterly* 50, 533–538.

Genesove, David, and Christopher Mayer, 2001, Loss aversion and seller behavior: Evidence from the housing market, Technical report, National bureau of economic research.

Gentzkow, Matthew, and Jesse M Shapiro, 2006, Media bias and reputation, *Journal of political Economy* 114, 280–316.

Gentzkow, Matthew, and Jesse M Shapiro, 2010, What drives media slant? evidence from us daily newspapers, *Econometrica* 78, 35–71.

Gentzkow, Matthew A, and Jesse M Shapiro, 2004, Media, education and anti-americanism in the muslim world, *The Journal of Economic Perspectives* 18, 117–133.

Gerber, Alan S, Dean Karlan, and Daniel Bergan, 2009, Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions, *American Economic Journal: Applied Economics* 1, 35–52.

Gerbner, George, and Larry Gross, 1976, Living with television: The violence profile, *Journal of communication* 26, 172–194.

Giannetti, Mariassunta, and Tracy Yue Wang, 2016, Corporate scandals and household stock market participation, *The Journal of Finance* .

Glaeser, Edward L, Joseph Gyourko, and Raven Saks, 2005, Why have housing prices gone up?, Technical report, National Bureau of Economic Research.

Glaeser, Edward L, and Charles G Nathanson, 2015, An extrapolative model of house price dynamics, Technical report, National Bureau of Economic Research.

Glasser, Theodore L, David S Allen, and S Elizabeth Blanks, 1989, The influence of chain ownership on news play: A case study, *Journalism Quarterly* 66, 607–614.

Goetzmann, William N, and Alok Kumar, 2008, Equity portfolio diversification, *Review of Finance* 12, 433–463.

Graber, Doris A, and Johanna Dunaway, 2014, *Mass media and American politics* (Cq Press).

Granziera, Eleonora, and Sharon Kozicki, 2015, House price dynamics: Fundamentals and expectations, *Journal of Economic Dynamics and Control* 60, 152–165.

Grinblatt, Mark, and Matti Keloharju, 2009, Sensation seeking, overconfidence, and trading activity, *The Journal of Finance* 64, 549–578.

Guren, Adam, 2014, The causes and consequences of house price momentum, *Cambridge, MA* .

Gurun, Umit G, 2010, Good news is endogenous, Technical report, Working paper, University of Texas at Dallas.

Gurun, Umit G, and Alexander W Butler, 2012, Don't believe the hype: Local media slant, local advertising, and firm value, *The Journal of Finance* 67, 561–598.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai, 2013, Superstar cities, *American Economic Journal: Economic Policy* 5, 167–199.

Harrison, J Michael, and David M Kreps, 1978, Speculative investor behavior in a stock market with heterogeneous expectations, *The Quarterly Journal of Economics* 323–336.

Hartzmark, Samuel M, 2014, The worst, the best, ignoring all the rest: The rank effect and trading behavior, *Review of Financial Studies* hhu079.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2008, The only game in town: Stock-price consequences of local bias, *Journal of Financial Economics* 90, 20–37.

Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer: A nonevent that made stock prices soar, *The Journal of Finance* 56, 387–396.

Ivković, Zoran, and Narasimhan Jegadeesh, 2004, The timing and value of forecast and recommendation revisions, *Journal of Financial Economics* 73, 433–463.

Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *The Journal of Finance* 60, 267–306.

Ivković, Zoran, and Scott Weisbenner, 2007, Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices, *Review of Financial Studies* 20, 1327–1357.

Katz, Elihu, Jay G Blumler, and Michael Gurevitch, 1973, Uses and gratifications research, *The Public Opinion Quarterly* 37, 509–523.

Katz, Elihu, and Paul Felix Lazarsfeld, 1966, *Personal Influence, The part played by people in the flow of mass communications* (Transaction Publishers).

Kaustia, Markku, and Sami Torstila, 2011, Stock market aversion? political preferences and stock market participation, *Journal of Financial Economics* 100, 98–112.

Klibanoff, Peter, Owen Lamont, and Thierry A Wizman, 1998, Investor reaction to salient news in closed-end country funds, *The Journal of Finance* 53, 673–699.

Kross, William, 1981, Earnings and announcement time lags, *Journal of Business Research* 9, 267–281.

Kross, William, and Douglas A Schroeder, 1984, An empirical investigation of the effect of quarterly earnings announcement timing on stock returns, *Journal of Accounting Research* 153–176.

Kuchler, Theresa, Basit Zafar, et al., 2015, Personal experiences and expectations about aggregate outcomes, *Federal Reserve Bank of New York Staff Reports* .

Kuttner, Kenneth N, 2001, Monetary policy surprises and interest rates: Evidence from the fed funds futures market, *Journal of monetary economics* 47, 523–544.

Lazarsfeld, Paul F, Bernard Berelson, and Hazel Gaudet, 1944, The peoples choice: How the voter makes up his mind in a presidential election, *New York: Duell, Sloan and Pearce* .

Ling, David C, Joseph TL Ooi, and Thao TT Le, 2015, Explaining house price dynamics: Isolating the role of nonfundamentals, *Journal of Money, Credit and Banking* 47, 87–125.

Liu, Pu, Stanley D Smith, and Azmat A Syed, 1990, Stock price reactions to the wall street journal's securities recommendations, *Journal of financial and Quantitative Analysis* 25, 399–410.

Livnat, Joshua, and Li Zhang, 2015, Is there news in the timing of earnings announcements?, *The Journal of Investing* 24, 17–26.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? textual analysis, dictionaries, and 10-ks, *The Journal of Finance* 66, 35–65.

McLuhan, Marshall, 1965, Understanding media.

Miller, Gregory S, 2006, The press as a watchdog for accounting fraud, *Journal of Accounting Research* 44, 1001–1033.

Mullainathan, Sendhil, and Andrei Shleifer, 2008, The market for news, in *Economics, Law and Individual Rights*, 90–122 (Routledge).

Nakajima, Makoto, et al., 2011, Understanding house-price dynamics, *Business Review, Federal Reserve Bank of Philadelphia* 2, 20–28.

Nathanson, Charles G, and Eric Zwick, 2014, Arrested development: Theory and evidence of supply-side speculation in the housing market, Technical report, Working paper.

Oikarinen, Elias, and Janne Engblom, 2015, Differences in housing price dynamics across cities: A comparison of different panel model specifications, *Urban Studies* 0042098015589883.

Ouellette, Laurie, 2013, *The media studies reader* (Routledge).

Peress, Joel, 2008, Media coverage and investors attention to earnings announcements .

Piazzesi, Monika, and Martin Schneider, 2009, Momentum traders in the housing market: survey evidence and a search model, Working Paper 14669, National Bureau of Economic Research.

Protess, David L, Donna R Leff, Stephen C Brooks, and Margaret T Gordon, 1985, Uncovering rape: the watchdog press and the limits of agenda setting, *Public Opinion Quarterly* 49, 19–37.

Reuter, Jonathan, and Eric Zitzewitz, 2006, Do ads influence editors? advertising and bias in the financial media, *The Quarterly Journal of Economics* 121, 197–227.

Roberts, Donald F, and Christine M Bachen, 1981, Mass communication effects, *Annual review of psychology* 32, 307–356.

Seasholes, Mark S, and Ning Zhu, 2010, Individual investors and local bias, *The Journal of Finance* 65, 1987–2010.

Sohn, Ardyth Broadrick, 1978, A longitudinal analysis of local non-political agenda-setting effects, *Journalism Quarterly* 55, 325–333.

Soo, Cindy K, 2013, Quantifying animal spirits: news media and sentiment in the housing market, *Ross School of Business Paper* .

Stickel, Scott E, 1995, The anatomy of the performance of buy and sell recommendations, *Financial Analysts Journal* 51, 25–39.

Strömberg, David, 2004, Radio's impact on public spending, *The Quarterly Journal of Economics* 119, 189–221.

Tetlock, Paul C, 2007, Giving content to investor sentiment: The role of media in the stock market, *The Journal of Finance* 62, 1139–1168.

Tetlock, Paul C, 2011, All the news that's fit to reprint: Do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.

Tetlock, Paul C, Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *The Journal of Finance* 63, 1437–1467.

Tipton, Leonard, Roger D Haney, and John R Baseheart, 1975, Media agenda-setting in city and state election campaigns, *Journalism quarterly* 52, 15–22.

Tracy, Joseph S, Henry S Schneider, and Sewin Chan, 1999, Are stocks overtaking real estate in household portfolios?, *Current Issues in Economics and Finance* 5.