The London School of Economics and Political Science

Empirical Essays on the Roles of News Media in an Urban Economy

Hayoung Kim

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Philosophy, London, September 2018.

I dedicate this thesis to my wife Sungmin and my son Jaeik.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 35,000 words.

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by the London School of Economics (LSE) Official Proofreading Service.

Abstract

Economists typically assume perfect information, but households are not always well informed, and face a high degree of uncertainty regarding the quality of goods or value of their assets. As a consequence, information from mass media is a main part of our everyday lives. Coverage in popular media outlets can catch the attention of millions of households and, therefore, news media may influence their decisions in several ways.

This thesis investigates the roles of news media in an urban economy and for housing markets. Each of the three empirical essays provides insights into how information from the media can shape the economic decisions of households as either homebuyers, homeowners, or consumers. While the first and the second chapters estimate the causal effects of publicity, in particular positive information, on house prices and non-housing consumption, respectively, the last chapter explores how media coverage may affect the link between house prices and homeowner spending. This thesis places particular emphasis on understanding the mechanisms through which news media may affect economic outcomes by empirically identifying potential channels.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

CHAPTER 2

CHAPTER 3

Acknowledgements

I am extremely grateful to my supervisors, Christian Hilber, Gabriel Ahlfeldt, and Felipe Carozzi for their advice, encouragement, and inspiration for the last four years. Without their guidance and support, I could not have completed this work.

I am also indebted to Steve Gibbons, Henry Overman, and Olmo Silva for their insightful comments in seminars and in person.

In particular, I would like to thank Keunkwan Ryu. With his support and motivation, I could begin this journey.

My sincere thanks also go to many friends and PhD colleagues, including Ying Chen, Pascal Jaupart, Cong Peng, and Cheng Keat Tang to name just a few. They have nourished my research and also my life in London.

Financial support from the London School of Economics and Political Science via the LSE PhD Studentship is very gratefully acknowledged.

Finally, I am especially thankful to Sungmin, my amazing wife as well as my best friend, for having faith in me and motivating my PhD. I could not have done this without her patience and support.

INTRODUCTION

The roles of news media, information, and quality uncertainty have become increasingly important in recent decades, making the study of media and information interesting both to academics and practitioners/policy makers. While economists typically assume perfect information, households are not always well informed, and face a high degree of uncertainty regarding the quality of goods or value of their assets. As a result, information from mass media is a main part of our everyday lives. Coverage in popular media outlets can catch the attention of millions of households and, therefore, news media may influence their decisions in several ways. News articles often contain rather vague, ambiguous, stale, or false information, but many people still base their decisions on such information (Barber and Loeffler, 1993; Tetlock, 2011; Oliver and Wood, 2014; Silverman and Singer-Vine, 2016; Allcott and Gentzkow, 2017).

By investigating the roles of news media in an urban economy and for housing markets, this thesis provides an interesting and interdisciplinary perspective on urban economies, and broadens the scope of urban and real estate economics. Each of the three empirical essays provides insights into how information from the media can shape the economic decisions of households as either homebuyers, homeowners, or consumers. While the first and the second chapters estimate the causal effects of publicity, in particular positive information, on house prices and non-housing consumption, respectively, the last chapter explores how media coverage may affect the link between house prices and homeowner spending. This thesis places particular emphasis on understanding the mechanisms through which news media may affect economic outcomes by empirically identifying potential channels.

I begin my journey by investigating whether homebuyers are responsive to media information in **Chapter 1**. Media outlets release a variety of best places lists every year. Such media opinions about city quality or livability could affect household location choices and investment decisions, thereby increasing demand for the listed towns. In this chapter, I exploit the *Money* magazine's "*50 Best places to Live in America*" lists to identify the causal impact of the quality information on local housing prices. The empirical results demonstrate that list inclusion has a statistically significant effect on local home prices: a 1-3% or approximately $ 3,000-9,000 increase per housing unit over two years. The finding indicates that third-party media information can affect homebuyers' decisions in the U.S.

housing markets. To the best of my knowledge, this paper is the first to find the evidence that quality information or recommendations on neighborhood quality can affect local housing prices. However, this paper leaves an important question unanswered: What drives such significant effects? Do homebuyers respond to the magazine's positive opinion? Or simply, does an increasing awareness raise demand for the listed towns?

To answer this question, **Chapter 2** examines the potential channels through which publicity can affect consumer demand. As consumers' expectations of product quality are a primary determinant of demand, positive publicity could lead consumers to believe that the product is of high quality, thereby increasing demand (vertical sorting). Meanwhile, media exposure could increase product awareness, which improves a match between heterogeneous consumers and products (horizontal sorting). By exploiting New York Times restaurant reviews, I identify the two potential channels. First, using the number of very localized taxi drop-offs as a proxy of restaurant demand, I find that consumer responses are statistically significant when reviews are positive: a 4.6% increase in taxi passengers, or approximately $1,560 weekly sales growth. By inferring diners' characteristics from destinations of post-dining taxi trips, I also show that demographic characteristics of locations have a significant effect on restaurant choices even when reviews are not positive; in response to non-positive reviews, a 10% larger share of Hispanic residents in a restaurant tract attracts 12.0% more Hispanic diners. The results imply that publicity could boost urban consumption not only by signalling product quality but also by informing consumers about the existence and characteristics of products. Overall, my contribution is twofold: 1) I suggest a novel way to measure real-time local economic activities and to identify socio-economic interactions between locations by utilizing taxi trip records. 2) This chapter makes an important contribution to the literature, as empirical evidence on the effect of awareness is scarce.

Chapter 3 provides further robust evidence on the awareness effect of news media in a different empirical setting. Housing wealth can affect a wide range of economic and social outcomes. Existing literature typically assumes that households make fully informed decisions, but homeowners may not have high level of awareness of their housing wealth. By exploiting local newspaper contents in the US, this chapter finds that more newspaper articles conveying house price information can make homeowner consumption more elastic with respect to regional housing prices. An increase of one standard deviation in the number of housing price news articles is associated with a 0.08 increase in homeowners' consumption elasticity. In contrast with the view that household decisions reflect fully informed and rational behaviors, the result suggests that providing relevant information can alter households' economic decisions by helping them to make more informed choices or making the information salient to the individuals. Thus, this chapter has a potential to dramatically broaden our understanding of the role of news media and information disclosure across a large number of settings.

CHAPTER 1.

Do Homebuyers Pay a Premium for "*Best Places to Live*" Cities?

1.1 Introduction

> *In early fall 2005, Cecile Druzba and her husband, Matt, were looking to make a change in their lives. They lived outside Woodstock, N.Y., but wanted to move to a community where the schools were great, jobs plentiful and their three school-aged kids could hop on their bikes to go to soccer practice, the library and a quaint downtown.*
>
> *So, they began the search by typing the words "Best Places to Live" into an internet search engine. Up came Money magazine's list for 2005 and immediately the couple zeroed in on No. 7 - Middleton. They visited once, fell in love with the place and moved to town that winter.*
>
> *- Milwaukee Wisconsin Journal Sentinel*[1]

Every year American media outlets release a variety of "best places" lists such as best places to live, best places for young couples, best places for retirees, and so forth. Although there is no empirical evidence, a lot of anecdotal stories tell us that the lists influence homebuyers' decisions. Homebuyers could value such lists not only because information about city quality or livability is difficult to obtain, but also because the quality of cities or neighborhoods is hard to judge before experience or residency. Nelson (1970) terms "experience goods" products or services of which quality can be fully evaluated only after consumption or purchase. The paper suggests that quality information is crucial particularly for durable and high-priced experience goods. Indeed, many consumers base their purchase decisions on recommendations from families, friends and consumer magazines/newspapers; in particular, when these decisions could affect their everyday lives in subsequent years. Thus perception of product quality can have substantial effects on consumer behaviors and choices. As Friberg and Grönqvist (2012) point out, lack of information often prompts

[1] The source: http://archive.jsonline.com/news/wisconsin/50987652.html

consumers to choose what others have chosen, potentially leading to herding behavior. In contrast, more information could result in efficiency gain by leading to better sorting between consumers and products. A wide range of methods have been used to convey such quality information to consumers such as ranking/ratings, user reviews, expert opinions or reviews, branding and advertising.

This study identifies an independent effect of neighborhood quality information on housing prices, by exploiting the *Money* magazine's annual reports, *"50 Best places to Live in America"*, which is one of the most popular and influential lists in the United States. The top 50 towns are publicly announced in rank order,[2] and the listed towns receive substantial attention from social media, local newspapers, city governments and homebuyers.[3] The role of the quality information has been little studied in the context of housing or neighborhoods, as experience goods. Studies on hedonic prices, or urban quality of life, typically assume that consumers have perfect information about inter-regional differences in local amenities, so differences in rent or land prices are largely explained by the differences in amenity or quality of life (Roback, 1982; Blomquist et al., 1988; Greenwood et al., 1991; Gabriel et al., 2003; Albouy and Lue, 2015; Albouy, 2016). However, many homebuyers face a high degree of uncertainty regarding the quality of neighborhoods and, therefore, such quality information could influence household location choices by informing buyers about city quality. People may overreact to the magazine lists, as summarized information, since their time and cognitive resources are too limited to process full information such as publicly available market statistics (Hong and Stein, 1999; Dellavigna and Pollet, 2009; Luca, 2016). It is also possible that positive tone of the magazine reports could lead homebuyers to believe that the listed towns are of higher quality than the other towns.

A key empirical challenge in this paper is how to isolate the independent effect of the quality indicator from effects of underlying city characteristics used in the listing procedure. Despite the theoretical potential of quality information to influence consumer choices, only a handful of studies have identified its causal effect on consumer demand for experience goods such as wines (Hilger et al., 2010; Friberg and Grönqvist, 2012), books (Berger et al., 2010), and movies (Eliashberg and Shugan, 1997; Reinstein and Snyder, 2005; Chen et al.,

[2] The magazine listed 100 towns until 2012

[3] For a few examples, see local governments' webpages (https://www.mckinneytexas.org/1017/1-Best-Place-to-Live, http://www.peachtree-city.org/index.aspx?NID=774) and newspaper articles (http://www.philly.com/philly/blogs/phillylists/Local-town-No-2-for-Best-Places-to-Be-Rich-and-Single.html, https://www.bizjournals.com/denver/stories/2010/07/12/daily9.htm)

2012). The limited empirical evidence is mainly because of a potential endogeneity issue. Products that receive positive reviews are likely to experience high sales even in the absence of positive information. For instance, Figlio and Lucas (2004) find that school grades assigned by the state have an independent effect on house prices and residential locations even when school attributes like test scores are controlled for. Kuang (2017) also shows that an effect on house prices of the quality of nearby restaurants is more significant when such information is made easily accessible from Yelp.com, a social network for user reviews on local businesses.

This paper addresses the empirical concern through two different approaches. By relying on difference-in-differences (DiD) estimators combined with propensity score matching, I show that list inclusion has a statistically significant effect on house prices: 1-3% over two years, which approximately corresponds to $3,000-9,000 per housing unit. However, the baseline estimators assume that the magazine lists are not true reflections of city quality. To identify an independent effect of the magazine reports without imposing the assumption, this paper also exploits local newspaper coverage on the magazine lists, and find that the lists have a significant effect on house prices during the one-year post-treatment period only when listed towns are introduced in local newspapers. This result identifies an independent effect of the magazine lists, as one would not expect such insignificant effects of list inclusion if the list inclusion truly reflects the quality of each city and if the underlying city characteristics are the only house price driver. However, a limitation is that the effect of local media coverage itself may not be causally interpreted due to potential endogeneity of newspaper reporting.[4]

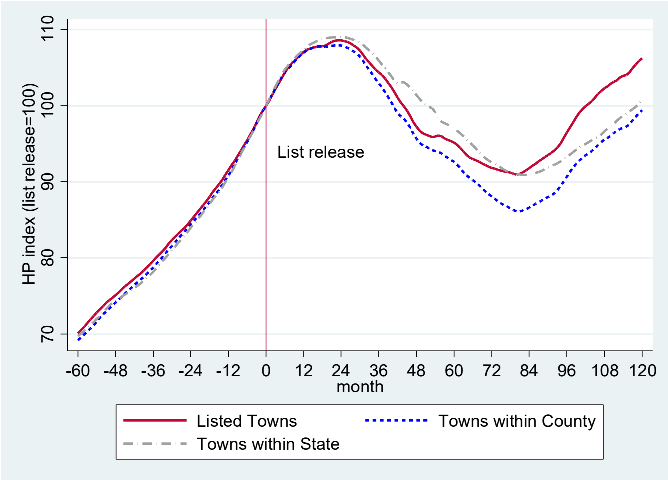
1.2 Backgrounds

Since July 2005 when the magazine released 2005's *100 best places list*, house prices in the listed towns have substantially outperformed those in the other towns within the same county for the following ten years (**Figure 1.1**). However, house price indices of the treated and untreated groups were in nearly parallel trends during the five-year pre-treatment period. This is probably because, by spatial sorting, towns within a county are more likely to have already been in spatial equilibrium and thus have substantially similar prior price trends.

[4] To address the concern, related studies exploit exits/entries of newspapers (Gentzkow et al., 2011; Gao et al., 2018), or reductions in media coverage caused by newspaper strikes (Peress, 2014).

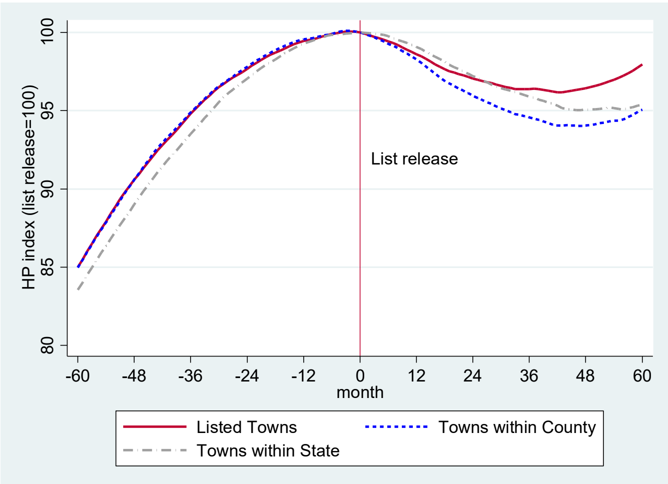
Figure 1.2 shows us consistent and more convincing results by plotting the pre- and post-treatment trends from seven lists (2005-2011). People who have read the reports may believe that the media's goal is just to entertain readers so the lists contain no new information about local fundamentals. In that case, one would not expect to observe any significant association between the list publication and house price growth. However, these analyses provide strong visual support for the link.

Figure 1.1. Pre- and post-treatment house price trends of towns listed in 2005



Notes: Figure 1.1 shows normalized average house price indices for three different groups of towns over the five-year pre-treatment and the ten-year post-treatment periods; each index is equal to 100 when the magazine list of 2005 was published. The first group (red solid line) consists of towns included in the 2005's top 100 list. The second group (blue dot line) compiles the other towns within the same county of each listed town. The last group (grey dash line) is the other towns within the same state of each listed town, including the second group of towns.

Figure 1.2. Pre- and post-treatment house price trends of towns listed in 2005-2011



Notes: Figure 1.2 displays time-series house price trends of the three groups described in Figure 1.2, but the first group consists of the towns included in the magazine's seven lists from 2005 to 2011. As a result of using more lists, the post-treatment period is five years long.

Still, one may be concerned that the listed towns are fundamentally different from the others so the different price trajectories are purely due to different city quality, despite the fact that the two graphs above confirm clearly parallel pre-treatment trends followed by significant post-treatment price deviations. Therefore, a central empirical question to identify is whether the magazine reports affect homebuyers' decisions/choices, or simply reflect local fundamentals that drive house prices up. On the one hand, the information on city quality or livability may facilitate better matches between agents and locations. Informed households may migrate toward towns of higher quality (*vertical sorting*) or to towns of which characteristics best meet their idiosyncratic needs (*horizontal sorting*[5]). As a result, housing markets in listed towns will experience higher demand as displayed in the figures. On the other hand, it is perfectly possible that the list inclusion is simply a proxy for positive past information. Indeed, it is told that the listed towns are selected based on superior city characteristics. If local fundamentals used by the magazine are not yet incorporated into local house prices, then one would also observe a positive relationship between the list publications and house price returns. To investigate this possibility, I begin with looking into the magazine's selection process as the first step.

1.2.1 Listing and Ranking Procedure

Basically, the magazine selects 50 places based on their own editorial constraints and statistics on demographic and socioeconomic characteristics. For the ranking of 2015, the magazine started with data on 3,625 U.S. towns with populations of 10,000 to 50,000, and ranked the towns on factors including job growth, diversity, and ease of living; giving the most weight to economic opportunity, housing affordability, education, and safety. They factored in more data on health, taxes, real estate, culture, and the economy; and limited the results to three places per state, one per county. These places were then sorted to represent all regions evenly. Thus they argue that all the listed towns have an above-average median income, a highly educated and growing population, low crime, good schools, healthy real estate appreciation and a thriving job market.[6]

[5] A seminal paper (Tiebout, 1956) argues that households choose neighborhoods that fit their heterogeneous preferences for public goods. While a primary assumption of the paper is that households have perfect information, they may not be well informed about the quality of multiple neighborhoods. Therefore, the magazine rankings may trigger the *Tiebout sorting*.

[6] *Money* August 2005 issue

Yet their ranking process depends on their subjective judgement about qualitative and intangible city characteristics. The magazine reporters visit the listed towns, interview residents, assess traffic, parks, and gathering places, and consider intangibles like community spirit. Eventually, the magazine determines the rank orders of the finalists and selects a winner based on reporting and the data collected by their reporters[7].

1.2.2 Nature of the rankings

In several respects, the magazine rankings are different from typical quality information from a third-party media outlet. An oft-cited paper, Dranove and Jin (2010), defines quality disclosure as an effort by a certification agency to systematically measure and report product quality for a nontrivial percentage of products in a market. Third-party information senders act as the intermediary of quality disclosure and are believed to report unbiased and accurate information in areas, particularly healthcare, education and finance. However, the magazine lists lack those considerations. The media lists only a tiny percentage of towns in the U.S.; out of thousands of American towns, only 50 towns are selected. More importantly, the rankings are not based on a transparent scoring system, and consequently are not stable over the years. The magazine has used slightly different criteria, methodology and data sets in their listing and ranking procedure over time; although precise variables, weights and functional forms are not disclosed. As Guterbock (1997) points out, indicator variables are added, and some dropped, and factors do not appear to be weighted exactly the same each year. As a result of this inconsistent criteria, methodologies and various editorial constraints, rankings are highly volatile over years.

Table 1.1. Population limits in 2005-2016

2005 : >14,000	2007 : 7,500 - 50,000	2012 : 50,000 - 300,000
2006 : >50,000	2008 : 50,000 - 300,000	2013 : 10,000 – 50,000
	2009 : 8,500 - 50,000	2014 : 50,000 - 300,000
	2010 : 50,000 - 300,000	2015 : 10,000 – 50,000
	2011 : 8,500 - 50,000	2016 : 50,000 - 300,000

Notes: The magazine's population limits differ each year. For the first two years, 2005 and 2006, the press used only lower limits, but since 2007, they have applied both lower and upper limits to narrow down candidate towns.

First of all, target groups differ each year. Recently, they have focused on very small towns in odd-numbered years, and large towns in the other years; a group of towns are listed in alternate years due to the different population limits **(Table 1.1).** Since 2007, population limits have been 50,000-300,000 in even-numbered years, and below 50,000 in the other

[7] For further details, see the webpage (http://time.com/money/3985631/best-places-2015-methodology/)

years. Even when comparing two odd-numbered years, 2013 and 2015, only 21 towns were listed in the both years **(Table 1.2)**. Moreover, McKinney, Texas topped the 2014 list, but dropped out of the 2016 list. Some places are getting better to live whereas others are getting worse, but such changes are very slow at the city level. More than half of the towns included in the top 50 list are not likely to significantly decline within only two years. Indeed, most of the factors used by the magazine are time invariant or only marginally time varying during the short period of time: for example, weather, distance to airport, the number of schools, colleges, universities, movie theaters, etc. Perhaps, most volatile are only some of the economic indicators. Importantly, this great inter-period volatility cannot be explained even by different variables used each year. Any new indicator is likely to covary with others already in use, so it would be highly unusual for the addition of one or two indicators to cause great fluctuations over years (Guterbock, 1997).

More importantly, listed and non-listed towns are not directly comparable across counties or states as a result of some editorial constraints. For instance, the magazine uses state and county quota limits, which slightly change over time. The magazine selects up to only three places per state and one per county in 2015 probably because one of their editorial goals is that the lists represent all U.S regions evenly. However, a plenty of desirable places might be located in a handful of states such as California; it is plausible that the magazine selects three from a cluster of good places in a certain year and another three two years later. Thus listed towns in Iowa are not necessarily of better quality than non-listed towns in California. Editorial reasons play a more crucial role in the listing procedure than local fundamentals do, and therefore the lists are more like the magazine's opinions or recommendations about good places to live rather than credible and reliable city rankings. Provoking controversy among readers and entertaining them may be more important for the for-profit media company than providing more reliable and credible city rankings that are believed to reflect true quality of cities. Luckily, discontinuities from the unstable rankings create an opportunity to test the effects of list inclusion on housing prices. First of all, cities listed in a specific year were not listed in the previous year[8]. Also, nearly half towns were not listed even two years ago. If a group of towns stay listed every year and the rankings stay stable over time, then there would be little time-series variation I can exploit for identification. More detailed research design is discussed in the following section.

[8] By population limits, towns could be listed only in 2005 and 2006 in a row

Table 1.2. 2013-2016 "50 Best Places to Live" rankings by *Money*

Rank	2013	2014	2015	2016
1	Sharon, MA	McKinney, TX	Apex, NC	Columbia, MD
2	Louisville, CO	Maple Grove City, MN	Papillion, NE	Eden Prairie, MN
3	Vienna Town, VA	Carmel, IN	Sharon, MA	Plano, TX
4	Chanhassen, MN	Castle Rock, CO	Louisville, CO	West Des Moines, IA
5	Sherwood, OR	Kirkland City, WA	Snoqualmie, WA	Parsippany/Troy Hills, NJ
6	Berkeley Heights, NJ	Columbia/Ellicott, MD	Sherwood, OR	Highlands Rancho, CO
7	Mason, OH	Clarkstown, NY	Chanhassen, MN	Clarkstown, NY
8	Papillion, NE	Ames, IA	Coppell, TX	Weston, FL
9	Apex, NC	Rochester Hills, MI	Simsbury, CO	Beaverton, OR
10	West Goshen, PA	Reston, VA	Solon, OH	Naperville, IL
11	Westford, MA	Eagan, MN	Acton, MA	Woodbury, MN
12	Parker, CO	Woodbury, MN	Rosemount, MN	Pflugerville, TX
13	Montville, NJ	Centennial, CO	Erie, CO	Centennial, CO
14	Farmington, UT	Irvine, CA	Westborough, MA	Sammamish, WA
15	Shrewsbury, MA	Newton, MA	Edina, MN	West Hartford, CT
16	Hillsborough, NJ	Parsippany/Troy Hills, NJ	Johnston, IA	Nashua, NH
17	Apple Valley, MN	Mansfiled, TX	Mason, OH	Eastvale, CA
18	Westfield, IN	South Jordan, UT	Draper, UT	Euless, TX
19	Newcastle, WA	Cary, NC	Woodbury, NY	Edison, NJ
20	The Colony, TX	Pflugerville, TX	Hewitt, TX	Irvine, CA
21	Savage, MN	Brookline, MA	Bedford, NH	San Ramon, CA
22	Waukee, IA	Gilbert, AZ	Twinsburg, OH	Ashburn, VA
23	Merrimack, NH	Boulder, CO	North Laurel, MD	Franklin, NJ
24	Firestone, CO	Rockville, MD	West Goshen, PA	Appleton, WI
25	Draper, UT	Orem, UT	Wylie, TX	Broomfield, CO
26	Brookfield, CT	Franklin, NJ	Dr. Phillips, FL	Cherry Hill, NJ
27	Farmington, MI	Piscataway, NJ	Nether Providence, PA	Hoffman Estates, IL
28	Menomonee Falls, WI	Bowie, MD	Berkley, MI	Hunter Mill, VA
29	Lindon, UT	Milpitas, CA	Sahuarita, AZ	Overland Park, KS
30	Windham, NH	West Chester, OH	Hillsborough, NJ	Fishers, IN
31	La Palma, CA	Pleasanton, CA	Damascus, MD	Newton, MA
32	Coppell, TX	Pembroke Pines, FL	Menomonee Falls, WI	Novi, MI
33	Suwanee, GA	Naperville, IL	Maryland Heights, MO	Koolaupoko, HI
34	Horsham, PA	Bellevue, NE	Tolland, CT	Oyster Bay, NY
35	Leesburg, VA	Amherst, NY	Urbana, MD	Sioux Fall, SD
36	Mill Creek, WA	Chapel Hill, NC	Springville, UT	Wellington, FL
37	Ankeny, IA	Dale City, VA	Germantown, WI	Cary, NC
38	Twinsburg, OH	Bolingbrook, IL	West Linn, OR	Hamden, CT
39	Cheshire, CT	Overland Park, KS	Mccandless, PA	Huntington NY
40	Ballwin, MO	Johns Creek, GA	Colchester, VT	Greenwich, CT
41	Montgomery Village, MD	O'Fallon, MO	Harrisburg, NC	Levittown, PA
42	Solon, OH	Franklin, TN	Waukee, IA	Matoaca, VA
43	Evans, GA	Ann Arbor, MI	La Palma, CA	Lee's Summit, MO
44	Pflugerville, TX	Fairfield, CT	Heber, UT	Spring, TX
45	Spring Hill, TN	West Hartford, CT	Cheshire, CT	Central Pasco, FL
46	Buffalo Grove, IL	Bensalem, PA	Stallings, NC	Fremont, CA
47	Pelham, AL	St. George, UT	Mukilteo, WA	Ames, IA
48	Peachtree City, GA	White Plains, NY	Vienna, VA	Edmond, OK
49	Walnut, CA	Meridian, ID	Walnut, CA	West Chester, OH
50	Simsbury, CT	Casper, WY	Woodstock, GA	Scottsdale, AZ

Notes: The magazine listed 100 towns until 2012, and has listed 50 best places since 2013. 2013 and 2015 lists include only small towns of population between 10,000 and 50,000, but 2014 and 2016 lists consist of mid-sized cities with 50,000 – 300,000 residents. As a result, none of towns were listed for two consecutive years.

1.3 Identification Strategy

In this section, I obtain more robust results by partialling out the effects of some local characteristics that might have predictive power on future house prices. Despite the previous two graphs (**Figures 1.1** and **1.2**) providing a solid support for the relation, identifying the causation is still challenging as the lists are not completely random. How can we tell whether the lists affected the market response, or whether some local fundamentals of listed towns simultaneously drove both the list inclusion and market responses? It may be the case that some pre-determined city characteristics have lagged or anticipatory effects on future house price trends. Most of such fundamentals used in the listing procedures should have already been incorporated into house prices before release of the list. But some may be very slowly capitalized into local house prices. For example, cities with lower unemployment rates last year are more likely to be included in this year's list attracting more households during this year, or even next year, than the other cities. It is also possible that there exists more or less autocorrelation of local fundamentals; places that outperformed neighboring places in economic growth are likely to do so in subsequent years. Thus simply comparing the mean values of outcomes for the listed and non-listed towns without any control variables could seriously overestimate the effect of the magazine reports.

For this reason, this study estimates the effect of list inclusion only. Statistical information and the magazine's various constraints determine which cities are included in a list whereas the ranking procedure relies upon the magazine's self-collected database on qualitative and intangible city characteristics, which are unmeasurable and unobservable. By focusing on the list inclusion and thus by not considering rankings at all, I can rule out any possibility of results being affected by such unmeasurable fundamentals.

1.3.1 Baseline Specification

Most of the studies that attempt to examine a causal impact of quality information face an empirical concern: omitted variable bias (OVB). In order to isolate the effects of list inclusion from those of the underlying statistical data, this study must control for all of the variables used in the magazine's listing procedure. Yet the factors used by the magazine are not fully disclosed. To partially address the issue, I control for county fixed effects. The magazine does use not only city-level characteristics but also county-, MSA[9]- or state-level characteristics such as property tax rates. It is probably because only a few of the statistics are collected and published at the city level in the US; most of the major statistics are compiled at the county or MSA level. Even, a large portion of city-level statistics do not cover all small towns which are the primary focus of the magazine. For this reason, the county-fixed effects are expected to substantially alleviate the empirical issue by controlling for all county-, MSA- and state-level characteristics. Thus city-level pre-treatment characteristics will be a main source of the OVB in this research.

To further overcome the problem, this study employs a difference-in-differences approach, which will cancel out most of factors that are time constant or in parallel trends. Indeed, a great part of house price fluctuations are explained by common factors such as macroeconomic conditions, and federal- or state-level policies, so price trends across towns within a county are highly parallel over time. Moreover, many city characteristics are time invariant or have little time-series variation during short-term periods. Thus those variables are expected to have at most only marginal effects on price differences between treated and control groups in DiD regressions.

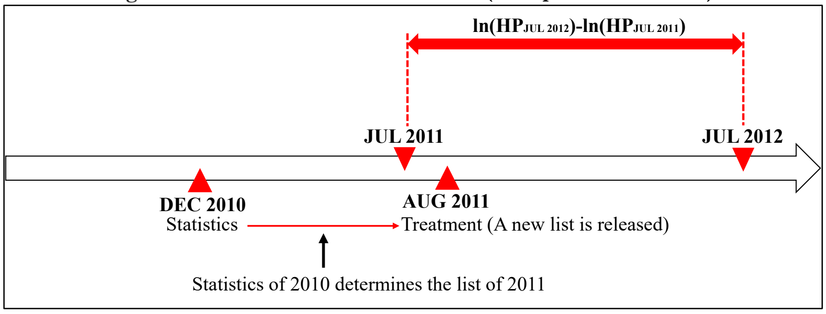
Table 1.3. City-level characteristics used in the magazine's listing procedure

Time varying factors	Factors with no or little time-series variation
median family income, median home price	share of residents completed at least college
population growth, job growth, unemployment rate	share of students attending public schools
family purchasing power	median commute time
violent crime rates, property crime rates	racial diversity, median age of residents
	share of residents married
	number of schools, colleges, universities
	number of movie theaters, restaurants, bars
	number of libraries, museums
	number of public golf courses, ski resorts
	number of doctors and hospitals, cancer mortality
	rainfall, temperature, air quality
	distance to airport

Notes: Table 1.3 presents determinants of the magazine rankings that have been explicitly mentioned in the reports. As you may note, most of them have no or little time-series variation over the short-term period of two years, but the magazine rankings are very unstable.

[9] MSA stands for Metropolitan Statistical Areas MSA, a US geographical core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core (https://www.census.gov/programs-surveys/metro-micro/about.html).

Figure 1.3. Time window for Diff-in-Diff (example of the 2011 list)



Notes: This figure describes the time window for the Difference-in-Differences estimator used in this study. The estimator measures the 1-year relative house price growth (from July 2011 to July 2012 in this example) in response to the list that was determined by statistics of 2010 and published in August 2011. Publicly available statistics are published with a lag of several months, which may affect the release date of the magazine list.

A central consideration in DiD approaches is the length of a time window. As seen in **Figure 1.3**, for example, the list for 2011 is based on statistical information of 2010, and this study measures the house price changes during the one-year period following the list release. However, housing markets will not react to such information immediately. Also, the effect of list inclusion will be further slowly observed due to time lags between list release and transaction closing; since the transaction process of buying a home takes several months, which includes searching for a neighborhood, a home, and then a mortgage lender. Thus this study takes post-treatment periods up to two years. Since 2006, not a single town has been listed for two consecutive years, so the two-year post-treatment period is not likely to cause any serious identification issue.

Now, a central question is whether pre-treatment statistical information has any anticipatory effects on post-treatment house price growth. If prior information has predictive power then the standard two-period DiD model should be as below;

(1.1) $$ln\, HP_{i,t} = \beta_1\, Listed_i + \beta_2\, (Listed_i \times t) + \boldsymbol{X'}_i\, \boldsymbol{\gamma_1} + (\boldsymbol{X'}_i \times t\,)\, \boldsymbol{\gamma_2} \;+\; v_c + (v_c \times t) + \xi_{i,t}$$

where $Listed_i$ is the primary focus, an indicator for whether individual city i is included in a list, and t is equal to 1 for post-treatment and 0 for pre-treatment. X_i is a vector of city-level covariates used in the listing process, and v_c is county fixed effects. Individual city-level differencing will derive the following model from the Equation 1.1;

(1.2) $$ln\, HP_{i,1} - ln\, HP_{i,0} = \beta_2\, Listed_i + \boldsymbol{X'}_i\, \boldsymbol{\gamma_2} + v_c + (\xi_{i,1} - \xi_{i,0})$$

By pooling all available cross-sectional data sets (10 lists from 2005 to 2014), I can obtain the following baseline specification;

(1.3) $ln\, HP_{i,t+1} - ln\, HP_{i,t} = \beta\, Listed_{i,t} + \boldsymbol{X'}_{\boldsymbol{i,t-0.5}}\, \boldsymbol{\gamma} + v_{c,t} + \varepsilon_{i,t}$

where $v_{c,t}$ denotes County × Year fixed effects and $\varepsilon_{i,t} = \xi_{i,t+1} - \xi_{i,t}$. Simply, the list is released in year t, and then regressions estimate the capitalization between year t and $t+1$. The subscript, $t - 0.5$ in $\boldsymbol{X}$ indicates the 7-8 month time gap between underlying statistics and list publications. Therefore, the coefficient β provides a DiD estimate of the causal impact of list inclusion conditional on prior city-level covariates.

To further recover the causal effect, I add two lagged terms, and one lead term on treatment (**Equation 1.4**). Because list inclusion effects may persist, grow or fade as time passes, the lag terms simply capture how long the effects can persist, and also control for previous list inclusion. The lead term is used both for a placebo test and for a reverse causality test. If the treatment causes future outcomes but not vice versa, then the dummy for future treatment should not matter in the equation (Angrist and Pischke, 2009). Thus, an insignificant lead term will be interpreted as evidence for no anticipatory effect and no reverse causation between treatments and outcomes.

(1.4)
$$
\begin{aligned}
(ln\, HP_{i,t+1} - ln\, HP_{i,t}) &= \mu\,(ln\, HP_{i,t} - ln\, HP_{i,t-3}) \\
&\quad + \beta_1\, Listed_{i,t-2} + \beta_2\, Listed_{i,t-1} + \beta_3\, Listed_{i,t} + \beta_4\, Listed_{i,t+1} \\
&\quad + \boldsymbol{X'}_{\boldsymbol{i,t-0.5}}\, \boldsymbol{\gamma} + v_{c,t} + \varepsilon_{i,t}
\end{aligned}
$$

Another concern is serial correlation. House prices are positively autocorrelated over time. The degree and persistence of the momentum in house price changes is one of the housing market's greatest puzzles. Case and Shiller (1989a) conclude that people seem to form their expectations from past price movements rather than knowledge of fundamentals, and the same authors' other paper (Case and Shiller, 1989b) reports that around half of the citywide changes in prices tend to be followed by changes in the same direction in the subsequent year. To avoid the empirical concern, I add one lagged dependent variable, the price growth of the previous three years, which has been widely included across a number of studies particularly in finance literature. It is also because the past house price appreciations are a factor used in the listing procedure. Importantly, adding the term could alleviate a concern that the lists may be more appealing to homebuyers living in places where house prices went up[10].

1.3.2 Matching

A main empirical challenge might be still the omitted variable bias arising from city characteristics that are used in the listing procedure but are missing in my data set. Even after conditioning on observables, there may be systematic differences between treated and untreated outcomes due to the unobserved variables. Most city-level aggregated characteristics except economic indicators do not change drastically over time, especially in the short run of one year or two. Yet the assumption that most omitted variables do not critically bias the empirical results may not be plausible, either. However, it is nearly impossible to address this issue by controlling for all city characteristics. This analysis controls for a range of pre-treatment city characteristics, particularly the most time-varying economic indicators, used in the selection process, but some might still be missing because the precise variables, weights and formula have not been disclosed and also because the magazine's methodology and criteria differ year by year. Moreover, this study could not have secured some data sets provided by commercial database vendors.

Thus the credibility of this analysis depends on how comparable a counterfactual is in the absence of some determinants. Many papers on place-based policies also report that it is challenging to find comparison areas. For this reason, a matching estimator is often combined with a DiD estimator in some recent papers on such place-based policies. Gobillon et al. (2012) take advantage of the French enterprise zone program, which provides wage subsidies for firms to hire local workers. In order to measure the direct effect of the program on unemployment duration, they estimate propensity scores of being designated as a municipality comprising an EZ and then restrict the control group to contain only municipalities whose estimated propensity score belongs to the same support as that of treated municipalities. Busso et al. (2013) compare census tracts in federal urban Empowerment Zones to those in rejected and later round zones with similar characteristics. To identify the causal impacts of EZ designation, they construct a set of control zones based upon data on the rejected tracts and use a DiD estimator adjusted by implicit propensity score weights, Oaxaca-Blinder estimator. More recently, Kline and Moretti (2014) employ

[10] A caveat is that most lagged dependent variable models are subject to endogeneity issues due to a common error term.

the same estimator to examine the effect of a regional development program, the Tennessee Valley Authority (TVA) program. To make treatment and control groups more comparable, they use as controls authorities that were proposed but never approved by Congress and also drop from their models control counties which appear to be substantially different from TVA counties.

Following this literature, I combine the DiD specifications with the propensity score matching (PSM hereafter) to further increase comparability of treated and untreated towns. Basically, this matching technique would make the control group as comparable as possible by omitting from the models most of the untreated observations which are substantially different from listed towns based on their observed city characteristics. It is likely that some towns in the same county are as good as the listed towns but are excluded mainly due to the county quota or another constraints. In this case, PSM will make the treated and untreated groups more similar in the estimated probabilities of list inclusion. More importantly, the matching estimator is one potential solution to the omitted variable issue by making use of information about observables to infer information about omitted variables. Essentially, matching and regressions are identical in terms of the core assumption underlying causal inference, so the differences between the two strategies are unlikely to be of major empirical importance. But if there is a good reason to believe that conditional on the observed, treatment and control groups are similar on unobservables, PSM may help to further isolate the treatment effect (Angrist and Pischke, 2009; Baum-Snow and Ferreira, 2015). Since most of city characteristics are strongly correlated with each other, towns with more similar observed characteristics are likely to feature more similar unobservables.

1.3.3 Data Sources

For empirical tests, I rely on two key data sources: (1) the *Money* magazine's *Best Places to Live* lists for ten years between 2005 and 2014 and (2) the monthly Zillow Home Value Index at the city level from January 2000 to August 2016. The Zillow indices are created from estimated sale prices on every home instead of a repeat sales methodology, and behave quite similarly to the well-known S&P/Case-Shiller indices for most of the historical period[11]. More importantly, the time-series house price indices are available for more than 10,000 U.S. cities or towns, enabling this study to link more than 80% of towns included in the magazine

[11] For further details, see the website (http://www.zillow.com/research/zhvi-methodology-6032/)

rankings to monthly indices during the period. As a result, my sample includes 11,035 U.S. towns in total, and 461 of them were listed at least once over the period.

Table 1.4. Descriptive statistics for listed and non-listed towns (2011-2014)

	Listed towns		Non-listed towns	
	Mean	S.D.	Mean	S.D.
median family income (USD)	98,616	20,198	65,188	28,729
median house value (USD)	319,381	156,125	196,719	161,455
population	53,628	42,320	13,562	31,520
Δln(population) (%)	1.83	5.73	0.84	10.28
unemployment rate (%)	5.88	1.65	8.93	5.11
employers	27,646	22,141	6,265	14,733
Δln(employers) (%)	1.30	5.91	-0.07	12.26
median commute time (mins)	23.61	5.20	21.64	7.15
prior-trends of house prices (t-3, t) (%)	1.09	10.15	-0.45	11.85
share of residents with bachelor or higher (%)	48.71	13.12	24.40	16.05
share of students attending public schools (%)	88.68	6.11	90.88	10.28

Notes: This table summarizes descriptive statistics for variables from the American Community Survey (ACS) data. It presents differences in the city characteristics between the listed towns and the other towns. The sample includes 8,596 towns (Census places) for 2011 to 2014. *prior-trends of house prices* is the house price growth over the 3-year pretreatment period. *Median family income, media house value, population*, and *employers* take logs in the following regressions.

To control for pre-treatment statistics that the magazine used, I use Census place-level demographic and socioeconomic variables from the American Community Survey (ACS). The ACS is a nationwide survey designed to provide communities with reliable and timely demographic, social, economic, and housing data every year. The survey data provides, for the first time, a continuous stream of updated information for local areas[12]. In particular, ACS 5-year estimates cover all areas in the US including very small populations called the Census places. Although the magazine also depends on the data source, the first release was the year 2009. As a result, using the data set allows this study to exploit only five years of the lists between 2010 and 2014, thereby losing half of the observations. My regression models include eight variables from the ACS: median family income, median home price, population growth, job (employer) growth, unemployment rate, median commute time, share of residents with bachelor or higher degrees, and share of students attending public schools[13]. As presented in **Table 1.4**, most of the city quality indicators are superior in listed towns. To construct a better control group, this paper uses only towns with a population below 500,000 and of which the median house price is below $1 million. As discussed earlier, the magazine uses both population limits and house price cap to focus on small affordable

[12] See the website (https://www.census.gov/content/dam/Census/library/publications/2008/acs/ACSGeneralHandbook.pdf)

[13] Most of them are time-varying, but a couple of variables have little time-series variation in them. Such time-invariant variables might have negligible impacts on future house price growth, but will help PSM to pair more comparable control towns to treatment towns.

towns. By adopting the same house price cap, and slightly looser population limits, this paper is able to further increase comparability between treated and untreated towns.

1.4 Results

Before including local fundamentals in regressions, I examine which factors affect the probability of list inclusion using a logistic probability model, which is also used to compute propensity scores; for this analysis, I use the same popuplation limits that the magazine used and differ each year (**Table 1.1**). Model specifications (1) to (4) in **Table 1.5** vary by the scale of fixed effects. All of the variables are used in the listing procedure, but not every characteristic has a significant effect on the probabilty of being listed. As the magazine focuses on affordable places, high incomes and low house prices appear to be consistently significant factors. Above all, this table reveals that socio-demographic characteristics are more crucial determinants than volatile economic characteristics even though the magazine has emphasized the importance of economic growth in their selection process. Towns with more educated residents or with more students attending public schools have a higher probability of being listed, whereas population and employer (or job) growth consistently have no effect on list inclusion; after the following analyses, I drop these two growth indicators in order to benefit from additional one year observations[14].

[14] The magazine uses levels for most variables, but changes for a few variables, such as population growth and job growth, exploiting the ACS data sets of previous two years. Using the change variables allows this study to use only four lists of 2011-2014. Thus I can take advantage of five lists by not using such growth indicators.

Table 1.5. Probability of being listed using logistic regressions (2011-2014)

	Dependent Variable: List inclusion dummy			
	(1)	(2)	(3)	(4)
ln(median family income)	2.886***	3.658***	4.921***	3.979***
	(0.666)	(0.833)	(1.005)	(1.248)
ln(median house price)	-1.458***	-1.091**	-3.165***	-2.272**
	(0.259)	(0.549)	(0.747)	(0.906)
unemployment rate	-0.310***	-0.152**	-0.181**	-0.114
	(0.0498)	(0.0612)	(0.0732)	(0.0895)
% bachelor+	0.0373***	0.0329***	0.0483***	0.0461**
	(0.00939)	(0.0121)	(0.0147)	(0.0189)
% public school	0.0613***	0.0709***	0.0787***	0.0972***
	(0.0137)	(0.0163)	(0.0183)	(0.0239)
median commute time	0.0149	0.0348*	-0.00648	-0.0217
	(0.0160)	(0.0194)	(0.0228)	(0.0302)
Δ ln(population)	5.081	3.188	5.162	2.872
	(3.589)	(3.804)	(4.267)	(4.626)
Δ ln(employer)	-5.346	-3.391	-4.616	-2.136
	(3.473)	(3.661)	(4.085)	(4.480)
Δ ln(house price) btw. t and t-3	0.0603	1.197	0.879	1.172
	(0.782)	(1.383)	(1.684)	(2.251)
Fixed Effects	Year	Year × State	Year × CBSA§	Year × County
Observations	5,805	4,487	2,612	1,193

Notes: This table presents results of logistic regressions with four different fixed effects specifications to infer which city characteristics the magazine has used in their listing procedure. The dependent variable is an indicator capturing whether the town is included in the magazine list of each year, and nine variables from the ACS (American Community Survey) along with the city-level Zillow Home Value Index are used to explain the probability of being listed. *% bachelor+* is the share of residents with a bachelor or higher degree. *% public school* is the share of students attending public schools. Δ *ln(house price) btw. t and t-3* denotes the house price growth over the 3-year pretreatment periods. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

§ CBSA stands for Core-Based Statistical Area, a U.S. geographic area that consists of one or more counties that are socio-economically tied to an urban center of at least 10,000 people, referring to both metropolitan statistical areas and micropolitan areas.

Table 1.6 shows the regression results using all covariates and lagged terms. Coefficients for list inclusion are consistently significant in all specifications during the first year (columns 1-3) and the second year (columns 4-6). Given the statistical significance of the treatment and the lagged treatment (t-1), the list inclusion appears to have a greater contribution to the second year house price appreciation. Many of control variables are somewhat predictive of the post-treatment house price growth at least for a year. Notably, the median house price has negative effects on list inclusion but positive effects on house price changes, so there is little likelihood of mean reversion because there is no tendency that places with lower pre-treatment house prices experience higher increases in post-treatment prices.

Table 1.6. Regression results: DiD (2010-2014)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.		$\ln(HP_{t+1})-\ln(HP_t)$			$\ln(HP_{t+2})-\ln(HP_{t+1})$	
Listed (t)	0.00720***	0.00486***	0.00393*	0.00458***	0.00439***	0.00454***
	(0.00163)	(0.00162)	(0.00209)	(0.00141)	(0.00132)	(0.00155)
Listed (t+1) : Lead			-0.00179			0.00388*
			(0.00178)			(0.00204)
Listed (t-1) : Lag 1			0.00738***			0.00140
			(0.00148)			(0.00177)
Listed (t-2) : Lag 2			0.00222			0.000132
			(0.00185)			(0.00148)
ln(median family income)		0.00423**	0.00419**		-0.00133	-0.00136
		(0.00187)	(0.00187)		(0.00193)	(0.00192)
ln(median house price)		0.00511**	0.00521**		0.00348	0.00356*
		(0.00207)	(0.00207)		(0.00212)	(0.00212)
unemployment rate		-0.000145*	-0.000145*		-0.000160*	-0.000160*
		(8.67e-05)	(8.67e-05)		(8.42e-05)	(8.41e-05)
% bachelor+		0.000151***	0.000146***		3.59e-05	3.21e-05
		(3.73e-05)	(3.74e-05)		(3.76e-05)	(3.78e-05)
% public school		3.97e-05	3.74e-05		1.71e-05	1.56e-05
		(3.08e-05)	(3.08e-05)		(3.15e-05)	(3.15e-05)
median commute time		0.000103**	0.000105**		0.000151***	0.000152***
		(4.92e-05)	(4.92e-05)		(4.82e-05)	(4.82e-05)
Δ ln(house price) between t and t-3		-0.0474***	-0.0475***		-0.0283***	-0.0284***
		(0.00670)	(0.00671)		(0.00613)	(0.00613)
Observations	42,155	40,030	40,030	42,140	39,630	39,630

Notes: This table shows regression results from Equation 1.4. The outcome variable of interest is the log of house price growth during the first posttreatment year (columns 1 to 3) or the second year (columns 4 to 6). *Listed (t)* is an indicator taking the value of 1 if the city is included in the "50 best places to live" list of year t. *Listed (t+1)* is a lead term, capturing whether the city is listed in the following year (t+1), for a reverse causality test. *Listed (t-1)* and *Listed (t-2)* are lagged terms for the two previous years to capture how long the list inclusion effects can persist, and also to control for previous list inclusion. As columns 4, 5, and 6 estimate the second-year effect, *Listed (t-2)*, *Listed (t-1), Listed (t)*, and *Listed (t+1)* serve as *Listed (t-3)*, *Listed (t-2), Listed (t-1)*, and *Listed (t)* in the columns, respectively. *% bachelor+* is the share of residents with a bachelor or higher degree. *% public school* is the share of students attending public schools. All columns control for Year×County fixed effects. Robust standard errors in parentheses are clustered by Year×County. *** p<0.01, ** p<0.05, * p<0.1

Obviously, a picture is worth a thousand words in DiD approaches. As displayed in **Figure 1.4**, listed towns outperformed the other towns in the same county by 1-2% over the two-year pre-treatment period and keep marginally outperforming during the post-treatment years. However, **Figure 1.5** shows that the PSM successfully replicates the pre-treatment house price trends of listed towns[15]. The pre-treatment trends of the treated and the matched control towns are almost parallel, followed by gradual post-treatment price deviations. Thus the list inclusion appears to have a causal effect on the outperforming post-treatment house returns, but the effect is quite lagged. Market responses to the list inclusion become more obvious from around the tenth month perhaps due to the transaction lag. Given these graphical evidences, the OVB does not seem critical[16]. If the treated cities experienced

[15] In order to compute propensity scores, this study exploits logistic regression models with aforementioned covariates. After calculating propensity scores using observations of each year, I match each of the listed towns to towns with the nearest propensity score within the same county.

[16] Nearby towns may be perceived as good as listed towns due to geographical proximity. The listed towns may also have affected nearby towns indirectly by causing households to move from the nearby towns to the listed towns. Either of these effects could lead to underestimation or overestimation of the treatment effect.

fundamenally different pre-treatment conditions in unobservables then the price deviations should have begun before or shortly after list publications.

Figure 1.4. Trends before matching
(2010-2014 Lists)

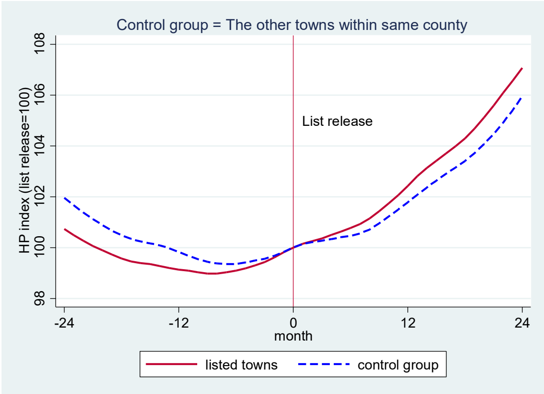


Figure 1.5. Trends after PSM using 6 variables
(2010-2014 Lists)

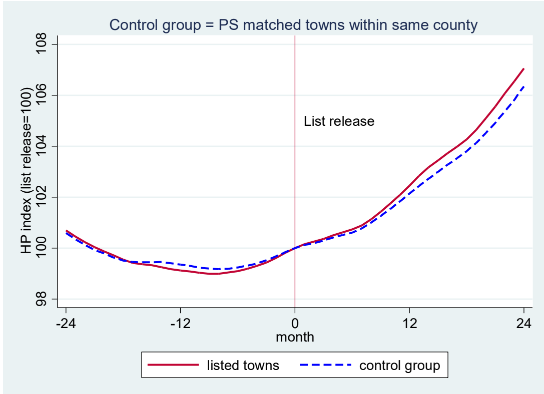


Figure 1.6. Trends after PSM using 2 variables
(2010-2014 Lists)

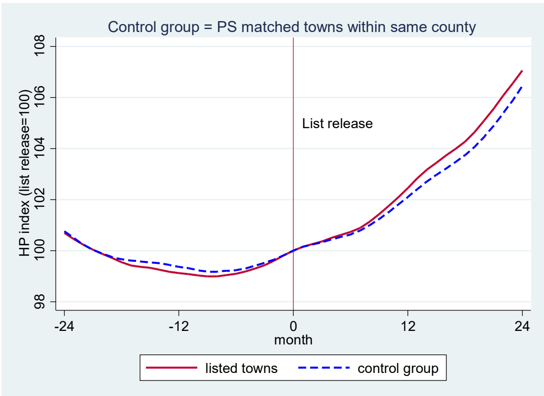
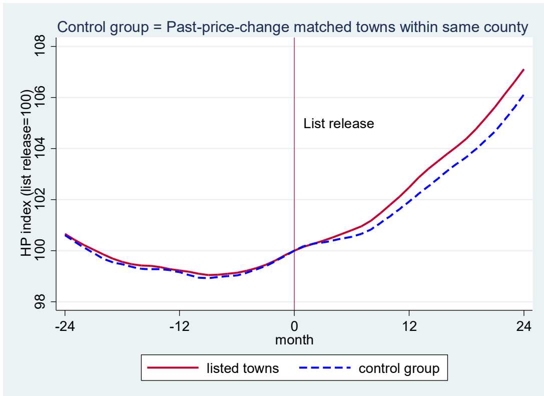


Figure 1.7. Trends after NNM using past price change
(2010-2014 Lists)



Notes: Figures 1.4-1.7 show normalized average house price indices for treatment and control groups over the two-year pre-treatment and post-treatment periods, using the magazine lists of five years from 2010 to 2014; each index is equal to 100 in the month when the magazine list was published. The first group (red solid line) consists of towns nominated as one of the best places to live by the magazine. The second group (blue dash line) compiles towns in the control group: either the other towns within the same county of each listed town (Figure 1.4), the towns selected on the 6-variable propensity score matching (PSM) within the same county (Figure 1.5), the towns selected on the 2-variable PSM (Figure 1.6), or the towns selected on the Nearest Neighbor Matching (NNM) using only the house price apprecition over the past three years (Figure 1.7).

To further clarify to which extent omitted variables can bias the matching estimator, I try another PSM using only two variables, *% bachelor+* and *% public school*, which are consistently and most significant in the previous regressions. Within a county, those two variables might be good enough to picture what each town looks like, but absence of all economic features can possibly lead to a serious mismatch between treated and untreated towns. Yet the control group selected on only the two variables are not substantially different from that using six variables (**Figure 1.6**). It may be the case that, as expected, most of the city characteristics are strongly correlated with each other. For example, towns with more educated residents might achieve higher median income levels and lower unemployment rates. **Table 1.7** confirms that the PSM estimator can substantially reduce unobservable imbalance by reducing observable imbalance, and also that the two-variable PSM does not make any noticeable distinctions compared to the six- variable model: even in regressions (**Table 1.8**). An alternative interpretation for this is that homebuyers are more interested in a town's socio-demographic features, such as education attainment and school quality, than economic indicators (unemployment rate or income levels of a neighborhood). In particular, if most inter-town migrations occur within a metropolitan area, then the town-level economic features would be less important considerations because both departure and destination cities are associated with a common urban core or job center, which mainly determine the residents' economic opportunities. If it is the case, such socio-demographic factors could be main house price drivers within a MSA.

The marginal but distinctive pre-treatment differences in the house price trends motivates an estimation strategy that controls for past house price changes only. The past price change may be one of the most significant factors in predicting house prices in the near future. Thus nearby towns of similar pre-treatment house price trends are more likely to have similar local characteristics or to be under influence of common price determinants. To do so, I employ Nearest Neighbors Matching (NNM) based on the price growth of past three years, and **Figure 1.7** and the specification (6) in **Table 1.8** suggest that the estimator is as good as the PSM.

Table 1.7. Covariate means for each group (2010-2014)

	Treatment Listed towns	Control Gr 1. The other towns in county	Control 2. 6-Var. PSM	Control 3. 2-Var. PSM	Control 4. NNM ($\ln HP_t/HP_{t-3}$)
Panel A - Means					
median family income (USD)	96,863	90,090	99,476	98,016	93,356
median house value (USD)	322,917	342,319	349,000	345,466	326,537
unemployment rate (%)	5.71	7.20	5.30	5.68	6.10
median commute time (mins)	23.52	25.22	24.03	24.08	24.67
share of residents with bachelor or higher	47.99	37.92	46.75	46.60	40.90
share of students attending public schools	88.59	87.32	86.80	87.12	87.11
pre-trends of house prices (t-3, t) (%)	-5.06	-7.19	-5.20	-5.19	-5.19
Panel B - Difference in Means (t-tests: *** p<0.01, ** p<0.05, * p<0.1)					
median family income (USD)		-6773***	2613	1153	-3447
median house value (USD)		19402	26083*	22549	-2984
unemployment rate (%)		1.50***	-0.41***	-0.03	0.30
median commute time (mins)		1.70***	0.51	0.56	1.04**
share of residents with bachelor or higher		-10.06***	-1.24	-1.39	-6.95***
share of students attending public schools		-1.28**	-1.80**	-1.47**	-1.15
pre-trends of house prices (t-3, t) (%)		-2.13*	-0.13	-0.12	0.14

Notes: Panel A shows how average characteristics of towns in each control group vary across the matching techniques. In addition, Panel B shows how each matching estimator can reduce imbalance in the observed city characteristic with t-test results. Control groups 1 to 4 are the results displayed in Figures 1.5 to 1.8, respectively.

Table 1.8. Regression results: Matched DiD (2010-2014)

	Dependent Variable: $\ln(HP_{t+2})-\ln(HP_t)$					
	(1) OLS	(2) OLS	(3) OLS	(4) 6-Var. PSM	(5) 2-Var. PSM	(6) NNM
Listed (t)	0.00932***	0.00714***	0.00674***	0.00713***	0.00693***	0.01110***
	(0.00212)	(0.00212)	(0.00214)	(0.00244)	(0.00262)	(0.00273)
ln(median family income)		-0.0111	-0.0151**			
		(0.00726)	(0.00683)			
ln(median house price)		0.00836	0.00709			
		(0.00885)	(0.00892)			
unemployment rate		-0.00105*	-0.000955			
		(0.000571)	(0.000585)			
% bachelor+		0.000344***	0.000349***			
		(0.000112)	(0.000117)			
% public school		-2.35e-05	-3.50e-05			
		(0.000144)	(0.000145)			
median commute time		-5.52e-05	-3.45e-05			
		(0.000185)	(0.000187)			
Δ ln(house price) btw. t and t-3	0.0813**		0.0454			
	(0.0332)		(0.0327)			
Observations	4,750	4,843	4,750	601	599	589

Notes: This table compares the matched Difference-in-Differences regression results (columns 4 to 6) with the OLS results (columns 1 to 3). The dependent variable is the log of house price growth during the 2-year posttreatment periods. *Listed (t)*, the variable of interest, is an indicator capturing whether the city is included in the "50 best places to live" list of year t. Model 1 controls for house price growth during the past three years, Δ *ln(house price) btw. t and t-3*, and Model 2 controls for basic city characteristics only. Model 3 includes both the city characteristics and the past house price trend. Models 4, 5, and 6 use as control groups towns paired by the six-variable propensity score matching, the two-variable propensity score, and the nearest neighbor matching, respectively. *% bachelor+* is the share of residents with a bachelor or higher degree. *% public school* is the share of students attending public schools. All columns control for Year×County fixed effects. Robust standard errors in parentheses are clustered by Year×County. *** p<0.01, ** p<0.05, * p<0.1

Figure 1.8. Trends before matching (2005-2014 lists)

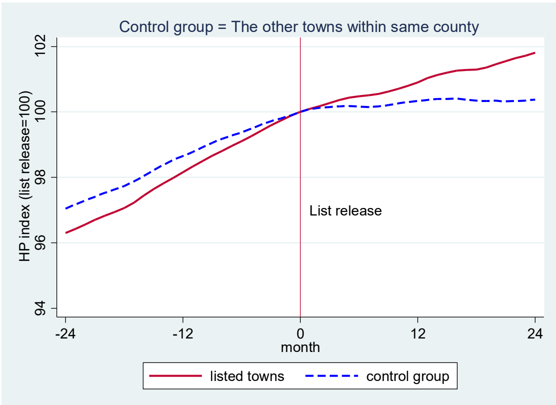


Figure 1.9. Trends after PSM using alternative covariates (2005-2014 lists)

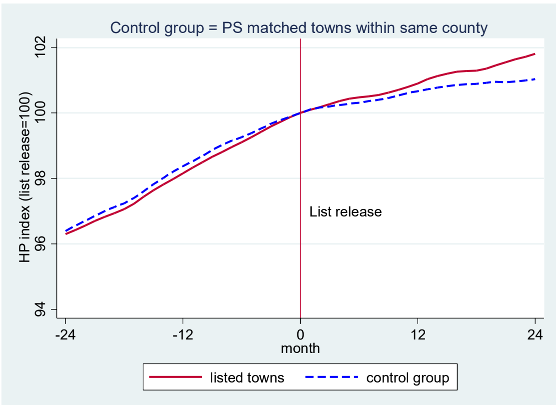
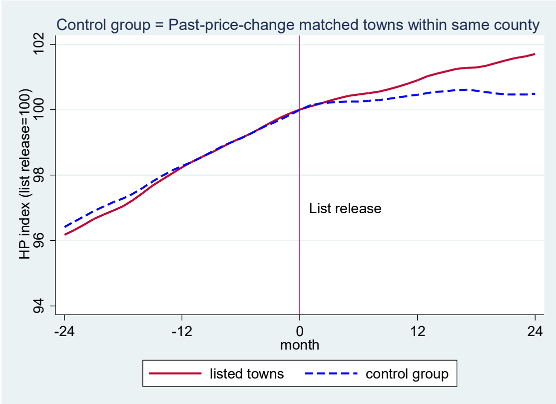


Figure 1.10. Trends after NNM using past price change (2005-2014 lists)



Notes: Each figure shows normalized average house price indices for treatment and control groups over the two-year pre-treatment and post-treatment periods, using the magazine lists of ten years from 2005 to 2014; each index is equal to 100 in the month when the magazine list was published. The first group (red solid line) consists of towns included in the magazine lists. The second group (blue dash line) compiles towns in the control group: either the other towns within the same county of each listed town (Figure 1.8), the towns selected on the propensity score matching using alternative covariables (Figure 1.9), or the towns selected on the Nearest Neighbor Matching using only the house price apprecition over the past three years (Figure 1.10).

Table 1.9. Covariate means for each group (2005-2014)

	Treatment Listed towns	Control 1. The other towns in county	Control 2. PSM	Control 3. NNM (ln HPt/HPt-3)
Panel A - Means				
Δln(population) (%)	1.87	1.15	1.67	1.78
unemployment rate (%)	5.07	6.48	5.38	5.63
Δln(employers) (%)	1.30	0.98	1.41	1.74
pre-trends of house prices (t-3, t) (%)	3.97	5.36	4.03	3.77
Panel B - Difference in Means (*** p<0.01, ** p<0.05, * p<0.1)				
Δln(population) (%)		-0.72***	-0.20	-0.14
unemployment rate (%)		1.41***	0.31**	0.58***
Δln(employers) (%)		-0.32	0.11	0.38
pre-trends of house prices (t-3, t) (%)		1.38	0.06	-0.33

Notes: This table uses alternative control variables from different sources. Panel A shows how average city characteristics vary across the control groups, and Panel B shows whether matching estimators (Control groups 2 and 3) can reduce imbalance in the observed city characteristic with t-test results. Control groups 1 to 3 are the results displayed in Figures 1.9 to 1.11, respectively.

Table 1.10. Regression results: Matched DiD with alternative covariates (2005-2014)

	Dependent Variable.: $\ln(HP_{t+2})-\ln(HP_t)$				
	(1) OLS	(2) OLS	(3) OLS	(4) PSM	(5) NNM
Listed (t)	0.0183***	0.0121***	0.0121***	0.00872***	0.0132***
	(0.00365)	(0.00358)	(0.00341)	(0.00260)	(0.00352)
Δ ln(population)		-0.0635	-0.0602		
		(0.0609)	(0.0632)		
unemployment rate		-0.00626**	-0.00737**		
		(0.00293)	(0.00361)		
Δ ln(employer)		0.0176	-0.00645		
		(0.0432)	(0.0340)		
Δ ln(house price) between t and t-3	-0.0534		-0.0959		
	(0.0825)		(0.0905)		
Observations	2,481	2,499	2,481	672	673

Notes: This table presents how the size and the statistical significance of the key parameters vary across control groups. The dependent variable is the log of house price growth during the 2-year post-treatment periods. *Listed (t)* is an indicator capturing whether the city is included in the magazine list of year t. Model 1 controls for house price growth during the past three years, Δ *ln(house price) btw. t and t-3*, and Model 2 controls for basic city characteristics only. Model 3 includes both the city characteristics and the past house price trend. Models 4 and 5 use as control groups towns paired by the propensity score matching and the nearest neighbor matching, respectively. All columns control for Year×County fixed effects. Robust standard errors in parentheses are clustered by Year×County. *** p<0.01, ** p<0.05, * p<0.

In order to assess the credibility of the research design, I conduct robustness tests by using an alternative set of covariates from different sources. I collected three variables, population from the US Census (2004-2013), and unemployment rate and number of employers from the US Bureau of Labor Statistics (2004-2013). These covariates allow me to use all of the ten lists (2005-2014) but lose nearly 85% observations in my sample because these data sets do not cover all of the small towns[17]. As displayed in **Figures 1.8 – 1.10**, and **Tables 1.9** and **1.10**, the visual illustrations of pre- and post-treatment trends, and regression results are highly consistent with the previous results, thereby confirming the robustness of the finding.

1.5 Role of Local Newspapers

In the previous section, a key assumption for identification is that the magazine rankings are not true reflections of city quality based on the unstable rankings over time. Given the assumption, this paper estimates an independent effect by controlling for underlying, observable city characteristics or by carefully pairing each treated city to a comparable untreated city. However, if the rankings are true reflections of city quality, then none of non-listed towns can be treated as a comparable counterfactual, and therefore the difference-in-differences estimators suffer from an omitted variable bias.

To identify an independent effect of the magazine reports without imposing the assumption, this section exploits local newspaper coverage on the magazine lists. It is hard to assume that most local residents obtain the information directly from the printed magazines or their website. The circulation of the magazine is around two million, which means that only a small share of the population would read the reports in the magazine. Instead, local newspapers may be the main source of local information to those who live in small towns or cities, as the media outlets in both print and electronic versions provide a wide range of local news, such as local jobs, opening of new restaurants, and events for kids in the neighborhood. After a new list is released, some listed towns receive substantial attention from such local media outlets, but the others do not. Thus the magazine effect is likely to be greater in the listed places that are introduced in their local newspapers. When the magazine reports are covered by a local newspaper, local residents within the media coverage

[17] More specifically, the variables cause a loss of more than 80% of non-listed towns but 40% of listed town, so the loss is more critical for control groups rather than treatment groups. As a result, the average number of towns per Year × County group reduces approximately from 19 to 8. In order to avoid additional observation loss, this analysis does not use the house price cap and population limits.

but out of the listed city might more strongly respond to the magazine lists thereby moving into the city. For that reason, I compare listed towns with and without attention from local newspapers, controlling for their ranks.

To do so, I examined whether each of the listed towns was introduced in local media outlets. From *Google News* search, I first collected news articles that report a magazine list between the list publication day and the last day of the year, using a keyword combination including "Money" and "best places to live" in addition to the name of the listed town and the year of each publication. Then I identified the publisher of each article, and included in my sample only articles that were published by local media outlets headquartered in the state where each listed town is located. As displayed in **Table 1.11**, only from 2010 on, more than half of the listed towns were covered by local media outlets. A limitation of this analysis is that I simply depend on *Google News* search results. Recently, almost all local newspapers provide news through both printed papers and websites, but some newspapers might have had only print formats in the early years of analysis. Other than local newspapers, some websites of local real estate or travel agencies also cover the magazine list publication, focusing on their home or neighboring towns included in the lists. But I do not take those company websites into account because people do not visit the websites on a regular basis; in contrast, many people read local newspapers or visit their websites every day[18].

Table 1.11. Local newspaper coverage of listed towns by year

Year	The number of non-listed towns	The number of listed towns		Total
		Covered by Local Newspaper	Not covered by Local Newspaper	
2005	14,767	13	83	14,863
2006	14,772	3	88	14,863
2007	14,766	4	93	14,863
2008	14,767	8	88	14,863
2009	14,769	45	49	14,863
2010	14,762	65	36	14,863
2011	14,771	50	42	14,863
2012	14,764	70	29	14,863
2013	14,814	31	18	14,863
2014	14,813	39	11	14,863
Total	147,765	328	537	148,630

Notes: This table shows variation in local newspaper coverage of the listed towns.

[18] There are more than 1,300 daily newspapers in the United States, and more than a quarter of adults read a newspaper every day (https://www.statista.com/statistics/183408/number-of-us-daily-newspapers-since-1975/)

Table 1.12. Regression results: The effect of local newspaper (2005-2014)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.		$\ln(HP_{t+1}/HP_t)$			$\ln(HP_{t+2}/HP_{t+1})$	
Listed	0.00801***	0.00293	-0.00504	0.0138***	0.0112***	-0.00247
	(0.00307)	(0.00383)	(0.00416)	(0.00273)	(0.00325)	(0.00381)
Listed × Local Newspaper Reporting		0.0103***	0.0158***		0.00531*	0.0122***
		(0.00365)	(0.00400)		(0.00309)	(0.00339)
Listed × Ranking	-0.000110**	-8.86e-05	-7.70e-05	-9.60e-05*	-8.49e-05*	-2.25e-05
	(5.55e-05)	(5.53e-05)	(7.05e-05)	(4.98e-05)	(4.98e-05)	(6.26e-05)
Fixed Effects	Year × State	Year × State	City, Year	Year × State	Year × State	City, Year
Observations	101,558	101,558	101,558	101,089	101,089	101,089

Notes: This table shows how local newspaper reporting affects the link between list inclusion and housing prices. The dependent variable is the natural logarithm of house price growth during the first posttreatment year (columns 1 to 3) or the second year (columns 4 to 6). *Listed* is an indicator capturing whether the city is included in each year's list. *Local Newspaper Reporting* is a dummy variable taking 1 if the listed town received attention from local newspapers between the list release day and the last day of the year. Models 1, 2, 4, and 5 include Year × State fixed effects, but Models 3 and 6 include year and individual city fixed effects. All columns control for house price growth during the past three years. Robust standard errors in parentheses are clustered by Year×County. *** p<0.01, ** p<0.05, * p<0.1

The regression results indicate that the magazine lists have an independent effect through the local newspaper coverage even when the rankings are assumed to truly reflect the order of city quality (**Table 1.12)**. Both list inclusion and rank orders have significant effects on local housing prices in models (1) and (4). However, Model 2 shows that the lists have a significant effect on house prices during the one-year post-treatment period only when listed towns are introduced in local newspapers. It is also notable that homebuyers do not pay more for higher ranked towns conditional on local media coverage. Perhaps, higher ranked towns are more likely to get media attention so could experience higher price increases.

The primary advantage of this analysis is to identify an independent effect of the magazine lists, rather than to identify an effect of local media coverage. If the list inclusion and the rank orders truly reflect the quality of each city and if the underlying city characteristics are the only house price driver, then one would not expect such insignificant effects of list inclusion and rank orders conditional on local newspaper reporting. Thus this finding indicates not only that the magazine lists affect homebuyers by affecting coverage decisions of local newspapers, but also that the results in the previous section do not suffer a serious omitted variable issue.

However, interpreting the effect of local newspapers is subject to an empirical concern, endogeneity of newspaper reporting. Local newspapers could reflect local economic conditions or the preferences of local residents, their readers. The local media outlets do not affect the magazine reports, but it is possible that local newspapers are more willing to deliver such information in MSAs or counties with residents who have a stronger interest in house prices. That is, the local newspapers are likely to reflect local people's characteristics toward investment or asset prices, so the results could be biased by the different characteristics across states or towns. For this reason, I include an individual fixed effects model. As seen in **Table 1.11,** some towns did receive local media attention in some years, but did not in other years. Exploiting this time-series variation eliminates the typical concern about biases arising from different local characteristics across space. The individual fixed effects models (3) and (6) in **Table 1.12** robustly identify the causal effect of the best place lists.

1.6 Conclusion

Most homebuyers base their decisions on information from families, friends, agents, and internet websites due to quality uncertainty of houses and neighborhoods. It is well documented that quality information could facilitate better matches and thus increase consumer welfare. In addition, housing accounts for a large part of household expenditures or family wealth in the US and many other countries (Flavin and Yamashita, 2002; Campbell and Cocco, 2007; Piazzesi et al., 2007; Yamashita, 2007). However, the link between housing demand and information on city quality has not yet been studied. By exploiting the *Money* magazine's annual reports, "*Best places to live in America"* lists, this paper characterizes the relationship. The magazine lists are a potential source that influences local housing markets for a couple of reasons. First of all, the magazine has the largest circulation of any monthly financial magazines in the United States: nearly two million readers. Also, the annual reports are frequently cited by local newspapers and social media, thereby being repeatedly exposed to individual homebuyers. In addition, the magazine is published by the *Time Inc.*, one of the world's biggest media companies, so has an excellent reputation with individuals.

However, identifying the link is challenging. Homebuyers may pay a premium both because the towns are listed and because the listed towns are of better quality than the other towns. To estimate the causal effect, this paper relies on difference-in-differences estimators combined with propensity score matching, and finds that the list inclusion predicts a gradual increase in house prices. For further causal investigations, this paper also exploits variation in local newspaper coverage on the magazine lists, and show that the lists have a significant effect only when the listed towns are introduced in local newspapers, thereby identifying an independent effect of the magazine lists. To the best of my knowledge, this paper is the first to find the evidence that quality information or recommendations from media outlets can affect local housing prices. Despite the fact that the empirical results robustly confirm the causal link, however, the mechanism behind the link is still unclear and unidentified. The magazine lists could make readers believe that the listed towns are far superior to the others and ultimately wish to purchase a home there. People may move to listed towns believing that the places would give them better economic opportunities or quality of life. The third-party information also can increase demand by simply making more people aware of the existence and the basic characteristics of the listed cities, which potentially contributes to better matches between heterogeneous homebuyers and neighborhoods. This paper leaves the question for future research.

CHAPTER 2.

How Does Media Attention Shape Urban Consumption? Evidence from the Restaurant Industry in New York City

2.1 Introduction

Consumers face a high degree of uncertainty regarding the quality of experience goods[19], so quality information such as critical reviews or user ratings is commonly used in relevant industries, and media and experts have played a key role in informing consumers about product quality. A wide range of papers highlight the importance of media reviews in guiding consumer choices, although only a handful of studies have identified its causal effects on consumer demand for experience goods such as wine (Hilger et al., 2010; Friberg and Grönqvist, 2012), books (Berger et al., 2010), and movies (Eliashberg and Shugan, 1997; Reinstein and Snyder, 2005; Chen et al., 2012).

Building upon the existing literature, this paper stresses potential channels through which publicity can affect consumer demand. Information about product quality could increase consumers' likelihood of purchasing the guided goods in two ways. Above all, positive reviews can lead consumers to believe that the product is of high quality,[20] thereby increasing demand (*vertical consumer sorting*). Hence this study emphasizes the role of positive quality indicators. But it is not the only channel. When no one knows all available alternatives, the media attention could increase demand not only by signalling quality of products but also by informing consumers about the existence of the reviewed products. For instance, Berger et al. (2010) documents that a negative review in the *New York Times* (NYT) increased sales of books by relatively unknown authors. As a marginal effect of information is larger when consumers have less prior knowledge or information about

[19] Nelson (1970) introduces the term, 'experience goods' of which quality can be fully judged only after consumption.

[20] In other words, such information increases consumers' expected utility from the good.

certain products, media attention is expected to be able to boost sales simply by making more people aware of the reviewed goods. Yet products or services in many industries are both vertically and horizontally differentiated so that goods differ not only by quality and price but also by characteristics associated with consumers' individual preferences. Simply, such information can help consumers to find products that best meet their heterogeneous needs by reducing search costs (*horizontal consumer sorting*). In this sense, this paper pays particular attention to how publicity can catalyze interactions between characteristics of consumers and products.

For empirical tests, this study relies upon the restaurant industry in New York City (NYC). Numerous traditional media outlets and internet blogs publish professional restaurant reviews every day, but not a single paper has reported any empirical evidence on the role of media reviews in this industry, to the best of my knowledge. First of all, this paper starts by identifying a causal effect of the *Michelin guide*, one of the most popular and iconic dining guidebooks in this industry. By using NYC's yellow taxi trip record data to construct a weekly panel of measures for restaurant demand, I find that newly Michelin-starred restaurants experience a 3.2% increase in the number of taxi passengers dropped off within a 100 feet radius from each restaurant, which approximately corresponds to weekly sales growth of $936. Yet an empirical limitation with the guidebook is that only starred restaurants receive media attention although hundreds of non-starred restaurants are included. Thus it is not possible to disentangle the effect of the quality indicator, Michelin stars, from the media effect, which increases restaurant awareness. To overcome this issue, I utilize another popular and influential information source, NYT restaurant reviews. Specifically, all of the restaurant reviews are published in the press but only 60-70% of them are awarded positive quality indicators by NYT critics. Empirical results demonstrate that consumer responses to the reviews are significant when restaurants win the media's positive indicators. Positively reviewed restaurants experience around 4.6% more taxi traffic or $1,560 more weekly revenues. This finding is consistent with existing literature showing that, controlling for price and quantity, positively reviewed products are more demanded.

Then this paper focuses on locational factors that possibly affect consumers' restaurant choices in order to identify the awareness effect. Since NYT is one of the US's most influential newspapers with a large readership, the reviews could substantially improve awareness of the existence and characteristics of each reviewed restaurant. Considering that restaurants differ not only in quality but also in many other characteristics, increasing awareness might be able to contribute to better matches between restaurants and consumers. Among many characteristics of restaurants, location is one of the most important considerations in restaurant choices. A consumer may visit a reviewed restaurant regardless of quality indicators if the venue is located in her favorite neighborhood. In particular, Davis et al. (Forthcoming) demonstrate that demographic similarity between restaurant locations and either home locations or diners' racial/ethnical identity has a large impact on diner decisions. By exploiting each taxi rider's neighborhood characteristics inferred from destinations of post-dining taxi trips during evening peak hours, I present two main findings: 1) Black and Hispanic diners have stronger preferences for demographic similarity than Asians. In response to NYT reviews, a 10% larger share of black or Hispanic residents in a restaurant tract is associated with 6.0% more black or 5.8% more Hispanic consumers, respectively. 2) Even when the reviews are not positive, black and Hispanic diners significantly respond to demographic characteristics of restaurant locations: a 6.7% or 12.0% increase in black or Hispanic diners, respectively. Thus the findings suggest that demographic characteristics of locations are important considerations for urban retailers targeting a specific demographic group of customers.

Given that around one thousand restaurants open each year and about 80% of them go out of business within the next five years in New York City [21], the findings of this paper have several implications. First of all, professional critics have played a critical role in deciding who survives and who fails in the competitive industry. Thus this research sheds light on the roles of food writers and media outlets by answering two questions; how much economic value the professional critics provide, and why many of them use star-rating systems in addition to informative and detailed reviews. Second, restaurants are both vertically and horizontally differentiated, so media information can improve matches between heterogeneous consumers and products even in the absence of quality indicators. That is, such information can increase demand both through quality indicators and through awareness, resulting in both vertical and horizontal consumer sorting. Last, not only centrality but also demographic characteristics are important locational forces positively affecting urban consumers' decisions[22].

[21] Source: a documentary film, Eat This New York (2004)

[22] Davis et al. (2016) report that restaurant consumption in NYC is only half as segregated as residence, but publicity appears to make urban consumption more segregated.

Section 2.2 further discusses the role of publicity in the restaurant industry, and the following section 2.3 details the explanatory and dependent variables of this research, and describes how a weekly panel is constructed. After identifying the causal effect of the Michelin guide in section 2.4, this paper carefully disentangles the effects of quality indicators and awareness in section 2.5, exploiting NYT reviews.

2.2 Media Attention in the Restaurant Industry

Considering the size of the restaurant industry, particularly in the U.S.[23], the potential economic impacts of popular restaurants are not negligible. Increasing consumer flows to top restaurants benefit nearby retailers such as cafés, pubs, bars, or alternative restaurants. Thus if a restaurant is successful, others tend to follow in nearby locations, and the area possibly becomes a popular destination for dining and, sometimes, also for housing. Indeed, many successful retail districts are anchored by popular restaurants. The presence of such tenants is also a plus in leasing adjacent space, so developers or investors always wish to attract popular restaurants to their properties. In the meantime, such popular restaurants could trigger urban retail gentrification. Rents in the hot neighborhood will go up as a result of increasing consumer flows, so more profitable retailers will displace early settlers including even landmark restaurants. That is, the composition of nearby retailers keeps changing to convert more traffic into sales or profits. A well-known example is *Union Square Café* in New York City. This iconic restaurant lost its space in 2015 although the anchor tenant was believed to have helped to revive its neighborhood since opening in 1986. It is simply because more profitable retailers such as international chain stores, banks and drugstores are willing to pay higher rents,[24] ultimately benefiting landlords and real estate investors.

Publicity has played a crucial role in the early prosperity of such top restaurants. As quality of food cannot be fully evaluated before experience, many diners base their choices on recommendations from various sources like the *Michelin Guide*. The French tyre company *Michelin* has awarded Michelin stars to a list of selected restaurants and published their

[23] The restaurant industry is the second-largest private sector employer in the U.S., and industry sales are projected to reach at $798.7 billion with 14.7 million employees in 2017 (National Restaurant Association, http://www.restaurant.org).

[24] The newly proposed rent for the space was $650,000 per year, and a Japanese chain restaurant moved into the location in 2016. For more information, see the following articles : New York Times (www.nytimes.com/2014/06/24/dining/ union-square-cafe-joins-other-victims-of-new-york-citys-rising-rents.html?_r=0), and Urban Land Institute (urbanland.uli.org/ sustainability/food-adds-flavor-value-real-estate-agrihoods-food-halls-food-based-concepts/).

guidebooks across the globe for more than a century, initially, to boost the demand for their car tires. There is no empirical evidence, but some anecdotal stories tell us that such third-party judgement has strong influences on diners' choice. For instance, a café in France was overwhelmed with phone calls for a table reservation after it won a Michelin star by mistake.[25] Even, the Michelin star has been alleged to boost nearby house prices,[26] as top restaurants are frequently considered as local amenities or attractions; a recent paper (Kuang, 2017) demonstrates that both quantity and quality of restaurants are capitalized in the value of nearby housing.

A wide range of methods have been used to convey such quality information to diners, but the internet user reviews on social media are recently becoming important in informing consumers about restaurant quality. In response, a couple of papers estimate a causal effect of Yelp's user ratings[27]. Since users assign a rating from 1 to 5 stars but the social media displays the average rating after rounding off to the nearest half star, researchers can take advantage of the display system by employing regression discontinuity design. Anderson and Magruder (2012) find that an extra half-star rating causes restaurants to sell out their prime time tables 19% points more frequently, using restaurant reservation availability as an outcome variable, and Luca (2016) reports that a one-star increase in rating leads to a 5-9 % increase in revenue.

Yelp might be crucial for established restaurants, and owners certainly pay a lot of attention to user ratings on social media. However, many newly opened retailers still depend on the traditional mass media to grab attention. When they cannot afford a prime location with high foot traffic, the new restaurateurs use various marketing strategies to attract traffic to their locations. In particular, quality information from seemingly neutral and credible third-party critics has the potential to have a great impact on demand (Nelson, 1974). Positive publicity by an influential newspaper draws substantial attention from potential customers, stimulating their appetites. Therefore, retailers have strong incentives to draw attention from news media. For this reason, sponsored reviews have been one of the most controversial ethical topics among food journalists. However, the economic

[25] For more information, see the Telegraph article (www.telegraph.co.uk/news/2017/02/18/workmens-cafe-overwhelmed-customers-accidentally-given-michelin/)

[26] See Financial Times (www.ft.com/content/f72f5962-0522-11e5-8612-00144feabdc0), Telegraph (www.telegraph.co.uk/property/house-prices/how-the-uks-best-restaurants-are-driving-up-property-prices/), and Country & Town House (www.countryandtownhouse.co.uk/property/michelin-stars-good-property-prices/).

[27] A user review website for local businesses (www.yelp.com/)

impacts or values of such media attention have not been studied yet in this business, to my knowledge, despite the fact that professional media critics have existed for a long time in the food and beverage market.

2.3 Data

For empirical tests, I rely upon 807 restaurants located in New York City. Specifically, 135 restaurants won a Michelin star or more at least once between 2006 and 2017. The number of starred restaurants in NYC has increased from 40 in 2006 to 99 in 2017, and, on average, about 10 restaurants were newly starred every year. In the meantime, the New York Times reviews two restaurants nearly every week, so the sample includes 755 restaurants reviewed by the newspaper from 2008 to 2016. Out of the 807 venues, 78 were both NYT-reviewed and Michelin-starred restaurants.

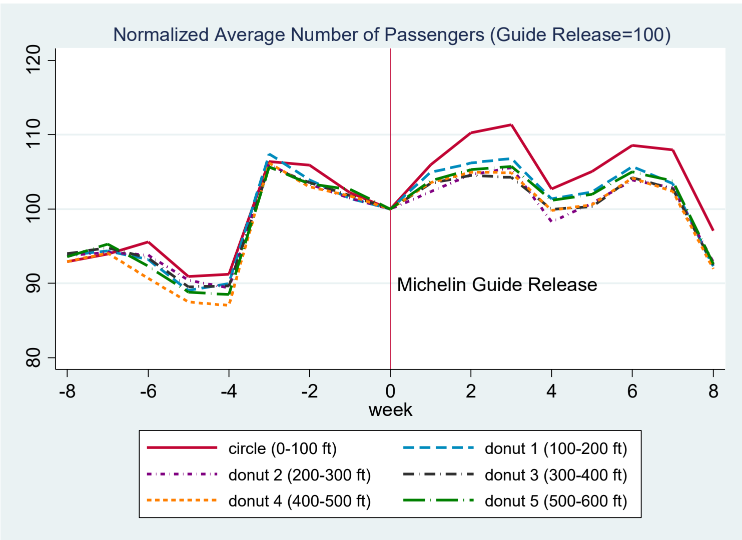
Notably, this study uses the number of taxi drop-offs or passengers as a proxy of demand for each restaurant. One of the key empirical issues in this kind of research is availability of individual restaurant-level sales data, so it is not surprising that there exist only a couple of empirical papers in the restaurant industry given that micro-data for restaurant revenues is rare or very inaccessible to researchers. To overcome the limitation, I utilize NYC's Yellow Taxi Trip Record Data collected and provided to the NYC Taxi and Limousine Commission (TLC) through the Taxi Passenger Enhancement Program (TPEP)[28]. The records cover all trips completed in yellow taxis from January 2009 to June 2016, including dates, times, locations (longitudes and latitudes) for both pick-ups and drop-offs in addition to passenger counts, trip distances, itemized fares, etc. More than 160 million yellow taxi trips per year, and around 1.2 billion trips in total occurred during the period. To construct a panel of restaurant demand measures, I identified taxi drop-offs within a 100 feet radius from each restaurant's centroid[29], and then counted each week's total number of drop-offs. By doing so, my data set consists of weekly taxi drop-offs for the 807 restaurants over 391 weeks; in addition to drop-offs, I also use the total number of passengers as a dependent variable. This indirect measure is expected to effectively capture consumer attention because taxis are a major mode to travel to restaurants in the crowded

[28] For more information on the data, see the website (www.nyc.gov/html/tlc/html/about/trip_record_data.shtml)

[29] More precisely, the radius is based on each restaurant's coordinates (longitude and latitude) extracted from Google Map, using a Stata package, *geocode3*.

city. Consumers visiting a restaurant after release of a review might have visited the same venue by another transport mode. But the change of transport mode may be independent of the change of consumer choices. Also, increases in taxi riding customers are likely to be correlated with increases in ones relying upon other transportation modes. In some senses, the number of taxi rides is a more accurate measure for consumer attention because revenues reflect only traded services, but taxi rides capture both traded and untraded demand; when there is no available table in destination restaurants, people often choose nearby alternative restaurants, so realized revenues underestimate actual demand shocks under supply constraints of restaurants. *Uber*, the ride-sharing company, has already taken advantage of the link between taxi rides and restaurant demand, releasing multiple types of restaurant lists ranked by the total number of drop-offs. For example, its '*Up-and-coming*' restaurants have the greatest increase in drop-offs, and '*Brunch spots*' list is determined by the number of drop-offs during weekend brunch hours[30].

Figure 2.1. Pre-and post-treatment trends of passenger counts at newly Michelin-starred restaurants



Notes: I investigate whether newly Michelin-starred restaurants experience higher taxi traffic than before. The Circle area is within a 100 feet radius from the centroid of each restaurant. Donut 1 is between two circles of 100 and 200 feet radiuses from the same restaurant. Likewise, Donut areas 2, 3, 4, and 5 are 200-300, 300-400, 400-500, and 500-600 feet far from each restaurant centroid, respectively. By identifying the taxis arrived at each restaurant, this figure shows the normalized average number of taxi passengers dropped off inside each of six areas over 8-week pre- and post-treatment periods. Each passenger count index is equal to 100 in the week when a new Michelin guide was published (week 0). This result is from 77 restaurants that newly won a Michelin star or more between 2009 and 2015.

More importantly, this real-time measure enables this study to conduct short-term causal investigations. As quality, price and quantity aspects of food are all time-varying but not

[30] For further details, visit Uber's website (www.uber.com/info/restaurant-guide/)

very observable, mid- or long-term analysis is subject to serious omitted variable biases. It is possible that restaurants raise prices or fail in quality control in response to increasing demand since a media review is published. If it is the case, then the treatment effects are underestimated. For this reason, this study primarily focuses on short-term analyses, taking full advantage of the proxy. To investigate whether and how much the outcome variable can capture the treatment effect in the short run, I counted the number of passengers dropped off inside areas of one circle and five donuts centered at each newly Michelin-starred restaurant: 0-100, 100-200, 200-300, 300-400, 400-500, 500-600 feet radiuses, respectively. The following visual illustration shows strongly parallel pretreatment trends, but distinctive deviations after treatment between the circle and donuts, suggesting that the treatment has a significant effect on the number of taxi riders dropped off inside the circle area (**Figure 2.1**).

An empirical challenge is that each restaurant has a different life span. For instance, many restaurants are in operation only for a few years. However, the accurate dates when each restaurant began to operate and closed permanently are missing[31]. To approximate the open date of each restaurant, I collected the date of the first user review on Yelp.com, and computed the date three months before each treatment (the publication date of a Michelin guide or a NYT review), assuming that each restaurant was in operation at least during the three-month pretreatment period[32]. Then I treated an earlier one of the two dates as the open date. Likewise, I treated a later one of the last user review date and the date six-month later as the close date. As a consequence, each restaurant's operation periods vary so the panel data set used in this study is highly unbalanced.

2.4 Impact of the Michelin Guide

A variety of quality indicators on restaurants have been favorites of diners, and one of the most iconic indicators is *the Michelin Guide*. In this section, I investigate whether a new star in the guidebook can help restaurants attract more customers. The Michelin star, as a quality indicator, could have a significant influence on consumer choices and thus on demand for the starred restaurants. To identify the link, this study estimates the impact

[31] I treat relocated restaurants as permanently closed to construct a consistent panel data.

[32] Compared to NYT reviews, the Michelin guide reportedly takes longer time to award a star, so I assume that newly Michelin-starred restaurants were in operation for at least six months before treatment.

only on newly Michelin-starred restaurants to exploit time-series variation in treatment. In addition, comparing Michelin-starred restaurants to non-starred or non-NYT-reviewed restaurants is subject to a potential endogeneity issue because quality of restaurants is unobservable. Consumers could visit a restaurant not only because it has won the star but also because it serves or is told to serve food of high quality even in the absence of such treatment. To alleviate the bias, my data set includes only Michelin-starred or NYT-reviewed restaurants because ever-treated restaurants are the best counterfactual in terms of quality. Thus treatment groups are newly starred restaurants each year, and control groups are the restaurants reviewed by NYT or starred by the Michelin guide in earlier or later years than the treated.

As the firm announces a new list of starred restaurants once a year (between September and November), I restrict the time window of each year to 17 weeks (8 pre-treatment weeks, 1 treatment week, and 8 post-treatment weeks[33]), thereby involving 117 weeks in total (17 weeks × 7 years). As discussed earlier, the main reason for this short-term analysis is because information on changes in prices or food quality is missing. Sellers can capitalize such information not only through quantity but also through price. That is, supply-inelastic restaurants are more likely to accommodate a demand shock by raising prices rather than quantities. Other than price changes, quality of food, the single most important omitted variable, is also time varying in the long run. Although some restaurants succeed in serving higher quality food over time, many fail in quality control as demand increases. In this case, the coefficient would badly underestimate the effect of the quality information. Therefore, following equation measures the relative demand increase in the posttreatment weeks, assuming that prices and food quality stay unchanged during the short-term time window.

(2.1) $Y_{it}^{C} = \alpha_i + \lambda_t + \beta\ Post\_Michelin_{it} + \gamma_1 Y_{it}^{D} + \gamma_2\ NYT_{it} + \varepsilon_{it}$

where,

Y_{it}^{C} : the natural log of the numbers of taxi drop-offs or passengers dropped off inside a circle area within a 100 feet (30.48 m) radius from restaurant i in week t.

[33] Technically, they are 8-week pre-treatment and 9-week post-treatment periods.

Y^D_{it} : the logarithm of drop-off or passenger count inside the donut area between 300 and 400 feet[34] radiuses from restaurant i in week t. This spatially lagged dependent variable is mainly used to control for regional taxi traffic volumes or trends. Some of the variation in taxi trips is due to variation in a measurement error. In addition, some unobservables may drive both the treatment and the demand for each restaurant. Thus the treatment is possibly correlated with some omitted variables. For example, as the number of taxi trips is a measure for consumer flows, restaurants in more intensely developed or popular retail areas are more likely to serve high quality food and to have more nearby taxi trips. Also, emerging retail districts or neighborhoods could be closely associated with the likelihood of getting Michelin-starred or NYT-reviewed, becoming increasingly popular destinations for taxi riders. Such spatial omitted variables can be a source of serious endogeneity. Thus this control variable is expected to substantially address the concern.

$Post\_Michelin_{it}$: the primary focus of this section taking the value 1 if restaurant i is newly Michelin-starred and week t is after restaurant i is treated, and 0 otherwise.

NYT_{it} : an indicator that is equal to 1 if week t is after restaurant i is reviewed by the New York Times. I include this term to control for prior awareness or perception of restaurant quality. As mentioned earlier, hundreds of restaurants open and other hundreds close every year, so it is not possible that people know about many of them. When NYT, an influential newspaper, reviewed a restaurant, the restaurant is more likely to get Michelin-starred because the review makes more people aware of the existence and quality of the restaurant, including the Michelin guide inspectors. Thus adding this term could help further recover the treatment effect of a Michelin star conditional on prior NYT reviews; if a restaurant was NYT reviewed before winning a Michelin star, the marginal effect of a Michelin star could be smaller because people are already aware of the existence and quality of the restaurant.

As selection is based on the quality of each restaurant, a main source of OVB (omitted variable bias) is at an individual level, but restaurant characteristics reflecting quality of food are unmeasurable and unobservable in this analysis. My data set contains some restaurant characteristics such as zip code, neighbourhood, price range, and cuisine type. They could partially affect consumer flows, but might not be significantly correlated with

[34] Most of avenues and streets in New York City are 100 feet and 60 feet wide, respectively. Some exceptional avenues are between 60 and 150 feet wide, while some streets are either 80 or 100 feet in width. If some taxis drop passengers off across the street or avenue, then a donut area between 100 and 300 feet radius could be partially affected by the treatment. For this reason, I use the donut area between 300 and 400 feet as a spatial covariate.

the treatment propensity. Moreover, the number of taxi trips is affected not only by restaurant characteristics but also by various location-specific components that are potentially associated with the treatment but cannot be explained by restaurant characteristics. In this sense, control restaurants may not provide a good measure of counterfactual taxi trips in the absence of the treatment even when the quality of each restaurant is perfectly observable.

For these reasons, I control for individual fixed effects (α_i), which can capture both restaurant- and location-specific omitted variables, exploiting temporal variation. In order to control for time trends and seasonal variation in restaurant demand, I also include week fixed effects (λ_t: week dummies for each of the 391 weeks represented in the sample), thus the number of taxi trips or passengers is determined mainly by the sum of a time-invariant individual effect and a week effect that is common across restaurants or locations.

Table 2.1.The effect of a new Michelin star by time window

	(1)	(2)	(3)	(4)
Pre- and Post-treatment Periods	2 weeks	4 weeks	8 weeks	12 weeks
Dependent Var.	ln (Dropoff)	ln (Dropoff)	ln (Dropoff)	ln (Dropoff)
Post_Michelin	0.0355**	0.0221*	0.0328**	0.0336***
	(0.0156)	(0.0133)	(0.0131)	(0.0123)
NYT	0.0524***	0.0515***	0.0408***	0.0305**
	(0.0161)	(0.0147)	(0.0134)	(0.0132)
ln($Dropoff^D$)	0.319***	0.360***	0.394***	0.412***
	(0.0538)	(0.0414)	(0.0391)	(0.0372)
Observations	16,013	28,347	52,581	75,799
Adj. R-squared	0.975	0.974	0.973	0.973

Notes: Based on regression results from Equation 2.1, this table shows how the results vary by time window. The dependent variable is the log of the number of taxi drop-offs inside the circle area of each newly Michelin-starred restaurant. *Post_Michelin* is an interaction of two indicators capturing whether the restaurant is newly Michelin-starred, and whether the week is a post-treatment period. *NYT* is a dummy that is equal to 1 if the restaurant is reviewed by New York Times before Michelin-starred. This term controls for prior awareness or perception of restaurant quality. *ln($Dropoff^D$)*, a spatially lagged dependent variable, is the log of the weekly number of taxi drop-offs inside the donut area between a 300 and 400 feet radiuses from each restaurant, and controls for local taxi traffic trends. All models include week and restaurant fixed effects. Robust and neighborhood-level clustered standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.1 presents that length of the pre- or post-treatment periods does not make any noticeable distinction in estimating the causal effect. The numbers of drop-offs and passengers also generate fairly similar results (**Table 2.2**). In response to the treatment, newly Michelin-starred restaurants see more than 3% increases in taxi drop-offs or passengers, or an increase of \$ 935.7[35] in weekly revenues during the 8-week posttreatment period. Interestingly, NYT reviews, as a control variable, have a larger effect. The

[35] For convenient interpretation, I use approximate sales as a dependent variable. On Yelp.com, each restaurant is assigned 1-4 \$ sign symbols, and the cost per person for a meal is displayed in **Table 2.3**. To approximate marginal sales in dollar terms, I take the median cost, and then restaurant sales are equal to (number of passengers) × (median cost per person). Hence estimated coefficients can be interpreted as the approximate marginal revenue of information. As the estimated sales growth is only a portion measured by increases in taxi-riding consumers, actual monetary growth could be greater.

newspaper has more circulations, so reviewed restaurants may receive more attention. Also, Michelin-starred restaurants may have inelastic supply due to their kitchen capacity and quality control. That is, they are not likely to strongly respond to demand shocks. Although the estimated coefficient is more significant on the trip count, I use the passenger count as a base dependent variable in the following regressions as it is a more accurate measurement for demand or consumer attention.

Table 2.2. The effect of a new Michelin star by dependent variable

Dependent Var.	(1) ln(Dropoff)	(2) ln(Passenger)	(3) Passenger
Post_Michelin	0.0328**	0.0313**	10.18
	(0.0131)	(0.0142)	(9.625)
NYT	0.0408***	0.0463***	24.99***
	(0.0134)	(0.0151)	(9.053)
ln(Dropoff^D)	0.394***		
	(0.0391)		
ln(Passenger^D)		0.334***	
		(0.0370)	
Passenger^D			0.0932***
			(0.0116)
Observations	52,581	52,580	52,581
Adj. R-squared	0.972	0.964	0.960

Notes: This table shows regression results from Equation 2.1 using three different dependent variables: the log of the number of taxi drop-offs inside the circle area of each newly Michelin-starred restaurant in Model 1, the log of the total number of taxi passengers (= the number of taxi drop-offs × the number of passengers of each taxi trip) in Model 2, and the number of passengers in Model 3. *Post_Michelin* is an interaction of two indicators capturing whether the restaurant is newly Michelin-starred, and whether the week is a post-treatment period. *NYT* is a dummy that is equal to 1 if the restaurant is reviewed by the New York Times before Michelin-starred. *ln(Dropoff^D)*, *ln(Passenger^D)*, *Passenger^D* are spatially lagged dependent variables to controls for local taxi traffic trends. All models include week and restaurant fixed effects. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.3.Yelp.com price information

Yelp Price Sign	*$*	*$$*	*$$$*	*$$$$*
Cost Per Person	under $10	$11-30	$31-60	above $61
Approximate Median Cost Per Person	$ 7	$20	$45	$80

Notes: Based on the cost per person estimated by Yelp.com (Row 1), this study applies the approximate median cost per person in Row 2 to calculate approximate increases in restaurant sales (=approximate median cost per person × increase in taxi passengers).

2.5 Impacts of NYT reviews

The Michelin guide does not enable this study to disentangle effects of quality indicators and publicity. This is because only starred restaurants receive substantial media attention although the guide book includes hundreds of non-starred restaurants in NYC. To overcome this obstacle, I use NYT reviews since all of the reviewed restaurants are published in the mass media but not all of them get starred by NYT critics. Yet relying upon the media reviews is subject to another empirical challenge arising from the unbalanced panel. As each reviewed restaurant has a different publication date and a different life span, the composition of treated and untreated groups change every week as a result of new publications. Thus estimates based on variation across restaurants could be inconsistent varying by each weeks' treatment-control composition. To relieve this concern, the baseline equation for NYT reviews depends largely on short-term time-series variation by replacing individual fixed effects (α_i) with restaurant-specific time window fixed effects (w_i). Many of the reviewed restaurants have more post-treatment than pre-treatment weeks. By using the same 17-week time window[36], I can exploit time variation within the symmetric time window of each restaurant, estimating the short term response to the treatment. Thus only within-window variation over time is used to estimate the treatment effect.

(2.2) $Y_{it}^{C} = w_i + \lambda_{tzcp} + \beta\ (PostNYT_{it}) + \gamma_1\, Y_{it}^{D} + \gamma_2\, Michelin_{\text{it}} + \varepsilon_{itzcp}$

where $PostNYT_{it}$, the key variable of this section, is an indicator taking the value 1 if t is a post-treatment week for restaurant i, whereas $Michelin_{\text{it}}$ is a control variable indicating whether or not entity i is a Michelin-starred restaurant in week t. In addition, I include week × zipcode × cuisine × price[37] fixed effects (λ_{tzcp}) instead of week fixed effects (λ_t) to further recover causal effects by controlling for heterogeneous demand shocks. In the previous section, effects of group-specific time-invariant unobservables could average out to zero over the weeks because the panel is perfectly balanced. But compositions of restaurants vary each week in this specification, so certain shocks are not likely eliminated. To avoid any biases associated with the group-specific demand, I exploit variation within week-zipcode-cuisine-price groups. It is also to more strongly restrict control groups or an alternative set of restaurants to more comparable restaurants that are expected to be under the influence of common demand shocks each week. In other words, restaurants in each group can be considered as a group of fairly good substitutes in the same market; when we decide which restaurants to go to, we comprehensively consider time, locations, cuisine types, and price ranges. Thus this specification could capture whether quality indicators significantly affect consumer decisions, controlling for combinations of time, location, cuisine, and price level. This is also a more conservative approach as the

[36] Again, 8 pre-treatment weeks, 1 treatment week, and 8 post-treatment weeks

[37] This study categorizes cuisine types into six groups (African, American, Asian, European, Spanish/Mexican/Latin American, and the others), and prices into four group ($, $$, $$$, and $$$$ in Table 2.3).

specification generates a larger robust standard error; **Table 2.4** shows that standard errors increase by about 64% when moving from column 1 (λ_t) to column 5 (λ_{tzcp}).

Table 2.4. The effect of New York Times reviews by fixed effects specification

	(1)	(2)	(3)	(4)	(5)
Dependent Var.	ln(Passenger)	ln(Passenger)	ln(Passenger)	ln(Passenger)	ln(Passenger)
PostNYT	0.0467***	0.0431***	0.0377***	0.0431**	0.0416**
	(0.0103)	(0.0121)	(0.0110)	(0.0168)	(0.0169)
Fixed Effects					
time window	YES	YES	YES	YES	YES
week	YES				
week×zipcode		YES			
week×zipcode×price			YES		
week×zipcode×cuisine				YES	
week×zipcode×cuisine×price					YES
Observations	183,696	183,696	183,696	183,696	183,696
Adj. R-squared	0.966	0.975	0.977	0.977	0.979

Notes: This table shows regression results from Equation 2.2 with various fixed effects specifications. In this table, I investigate how the standard error and the size of coefficients vary by fixed effects specification. The dependent variable is the log of the number of taxi passengers. *PostNYT* is an interaction term of two indicators capturing whether the restaurant is reviewed by the New York Times (NYT), and whether the week is a post-treatment period. All models rely on 17-week time window which consists of 8 pre-treatment weeks, 1 treatment week, and 8 post-treatment weeks. By controlling for individual-restaurant-specific *time window* fixed effects, the models exploit within-time-window variation over time. The *cuisine* fixed effects consists of six dummies for African, American, Asian, European, Spanish/Mexican/Latin American, and the others, and the *price* fixed effects employ the four price groups in Table 2.3 ($, $$, $$ or $$$$). All models include two control variables, a Michelin star indicator and a spatial lag. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

NYT reviews increase taxi drop-offs or passengers by about 4% (**Table 2.5**), or weekly sales by $1,294, which amounts to more than $10,000 in total for the 8-week post-treatment period. The marginal revenue also can be interpreted as the monetary value of the incentive for restaurants to sponsor the media critics; total marginal sales including all transit modes in longer terms could be larger.

Table 2.5. The effect of New York Times reviews by dependent variable

	(1)	(2)	(3)
Dependent Var.	ln (Dropoff)	ln(Passenger)	Passenger
PostNYT	0.0362**	0.0416**	30.26**
	(0.0166)	(0.0169)	(12.53)
Observations	183,699	183,696	184,011
R-squared	0.984	0.979	0.952

Notes: This table shows regression results from Equation 2.2 with three different dependent variables: the log of the number of taxi drop-offs in Model 1, the log of the total number of taxi passengers in Model 2, and the number of passengers in Model 3. *PostNYT* is an interaction term of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. Relying on a 17-week time window, all models include two control variables, a Michelin star indicator and a spatial lag, and two fixed effects, individual-restaurant-specific time window and week-zipcode-cuisine-price. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

2.5.1 Quality Indicators and Consumer Responses

NYT reviews *per se* do not ensure that reviewed restaurants are of higher quality than the non-reviewed. Just as the Michelin guide uses star-ratings, the newspaper uses two kinds of quality indicators to summarize their opinion: *NYT Critics' Pick* (CP) and stars. As displayed in **Figure 2.2**, CP is a binary indicator (Yes or No). In addition to that, the chief restaurant critic can assign star ratings from zero to four[38]. Out of 755 restaurants reviewed from January 2008 to December 2016, 425 won the CP, and 304 obtained at least one star (**Table 2.6**).

Figure 2.2. NYT restaurant reviews

Panel A. Positive review

Panel B. Non-positive review

Notes: This figure shows examples of positive and non-positive reviews by the New York Times (NYT). NYT uses two types of quality indicators, *NYT Critics' Pick* (yes or no) and star ratings (0-4 stars). Panel A displays a positive review with both quality indicators, whereas Panel B shows a non-positive review without any of them.

Table 2.6.Frequency distribution of each quality indicator

Panel A. Individual frequency distribution

	YES	425		0	451
Critics' Pick	NO	330	***Star***	1	124
	Total	*755*		2	136
				3	40
				4	4
				Total	*755*

Panel B. Joint frequency distribution

		Star					
		0	1	2	3	4	*Total*
	YES	196	61	125	39	4	*425*
Critics' Pick	NO	255	63	11	1	0	*330*
	Total	*451*	*124*	*136*	*40*	*4*	*755*

Notes: This table shows the frequency distribution of the two quality indicators for 755 restaurants reviewed by NYT from 2008 to 2016.

[38] Star ratings range from zero to four stars and reflect the reviewer's reaction primarily to food, with ambiance, service and price taken into consideration zero is poor, fair or satisfactory—One star, good. Two stars, very good. There stars, excellent. Four stars, extraordinary—according to the NYT critics (https://archive.nytimes.com/www.nytimes.com/packages/html/dining/info/ratings.html and https://dinersjournal.blogs.nytimes.com/2012/03/13/why-our-reviews-have-stars/). Thus reviews without any of their indicators can be considered as negative or at most neutral.

Table 2.7.Heterogeneous effects of New York Times reviews

Dependent Variable: ln(Passenger)	(1)	(2)	(3)	(4)
PostNYT × Pick (Yes)	0.0581***			
	(0.0212)			
PostNYT × Pick (No)	0.00851			
	(0.0241)			
PostNYT × Star (4)		0.281***		
		(0.0480)		
PostNYT × Star (3)		0.0532		
		(0.0540)		
PostNYT × Star (2)		0.0580*		
		(0.0310)		
PostNYT × Star (1)		-0.0103		
		(0.0262)		
PostNYT × Star (0)		0.0377*		
		(0.0206)		
PostNYT × Pick (Yes) × Star (Yes)			0.0625**	
			(0.0292)	
PostNYT × Pick (Yes) × Star (No)			0.0490**	
			(0.0247)	
PostNYT × Pick (No) × Star (Yes)			-0.0257	
			(0.0352)	
PostNYT × Pick (No) × Star (No)			0.0282	
			(0.0327)	
PostNYT × Positive				0.0452**
				(0.0191)
PostNYT × Nonpositive				0.0282
				(0.0337)
Observations	183,696	183,696	183,696	183,696
Adj. R-squared	0.979	0.979	0.979	0.979

Notes: This table shows heterogeneous effects by quality indicator. The dependent variable is the log of the number of taxi passengers. *PostNYT*, the variable of interest, is an interaction term of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. *Pick (Yes)* and *Pick (No)* are indicators capturing whether the restaurant received *NYT Critics' Pick*. *Star (0), Star (1), Star (2), Star (3),* and *Star (4)* in Model 2 are dummy variables indicating the number of stars that the restaurant received. *Star (Yes)* and *Star (No)* in Model 3 are dummies capturing whether the restaurant received any stars. *Positive* and *Nonpositive* in Model 4 are indicators capturing whether the review is positive or non-positive. In this study, positive reviews are with any of the two quality indicators (*NYT Critics' Pick* or stars), and negative reviews are without any of them. Relying on a 17-week time window, all models include two control variables, a Michelin star indicator and a spatial lag, and two fixed effects, individual-restaurant-specific time window and week-zipcode-cuisine-price. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

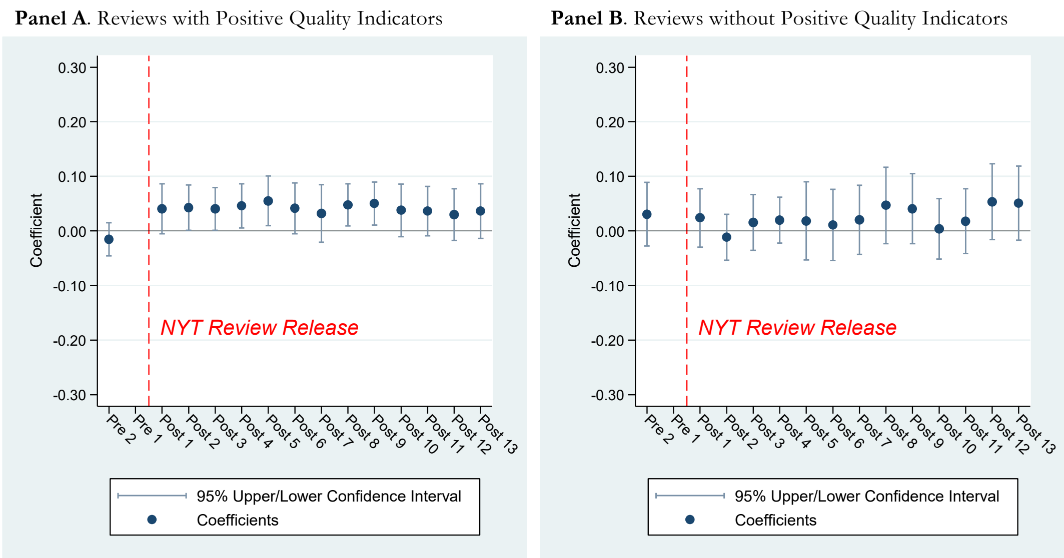
To investigate whether the marginal impact of information varies according to various quality groups, the key treatment dummy is interacted with each indicator dummy. Relevant literature, particularly in finance, points out that such indicators, as summarized information, could play a crucial role in consumer responses; as they have a limited amount of time and cognitive resources to process information, people underreact to full information (Hong and Stein, 1999; Dellavigna and Pollet, 2009; Luca, 2016). Indeed, **Table 2.7** tells us that the positive quality indicators have a significantly positive impact on restaurant sales. Specifically, restaurants that have won any positive indicator significantly outperform the others, generating 4.6%[39] more taxi riders (**column 4**) or

[39] exp(0.0452)-1=0.0462

approximately \$1,560 additional weekly sales. The empirical results indicate that consumers selectively respond to the reviewed restaurants, confirming that the indicators play a critical role in boosting demand. Positive indicators may strongly influence consumers' expected utility or product evaluations, thereby leading to vertical consumer sorting. Importantly, selection or being reviewed itself could be an outcome of gaining popularity, but positive indicators are the critics' subjective judgement, so this result robustly confirm the causal link between quality information and restaurant demand.

As mentioned earlier, even negative reviews could have positive effects. Given that many of NYT reviewed restaurants are relatively new and thus few people know about the restaurants, it is not surprising that reviewed restaurants experience an increase in sales even when reviews are not positive, as seen in column (4). Yet the link between awareness and demand appears weak, as the increase is statistically insignificant. To further examine the effect of awareness, I conduct a mid-term analysis using the following equation with an extended time window of 8-week pretreatment and 52-week posttreatment periods.

Figure 2.3. Mid-term estimates of the NYT review effect (8 pre-treatment and 52 post-treatment weeks)



Notes: The figures display estimated coefficients and 95 percent confidence intervals from Equation 2.3 during two pre-treatment and 13 post-treatment periods; each period is 4-week long, and thus the time window is 60 weeks in total. The missing *Pre_1* is the baseline period. Panels A and B show the effects of a positive review and a non-positive review, respectively, over the periods. All models include two control variables, a Michelin star indicator and a spatial lag, and two fixed effects, individual-restaurant-specific time window and week-zipcode-cuisine-price. Standard errors are robust and neighborhood-level clustered.

$$(2.3)\quad Y_{it}^{C} = \beta_{-2}\, Pre\_2_{it} \times Positive_i + \textstyle\sum_{j=1}^{13} \beta_j\, Post\_j_{it} \times Positive_i$$
$$+\, \delta_{-2}\, Pre\_2_{it} \times NonPositive_i + \textstyle\sum_{j=1}^{13} \delta_j\, Post\_j_{it} \times NonPositive_i$$
$$+\, w_i + \lambda_{tzcp} + \gamma_1\, Y_{it}^{D} + \gamma_2\, Michelin_{it} + \varepsilon_{itzcp}$$

where each Pre or $Post$ dummies indicates a 4-week long pre- or post-treatment time period[40], and the missing dummy, Pre_1 (a period from the 4th to the 1st pre-treatment week), is the reference period. Thus this model estimates each coefficient as the effect of NYT reviews during the corresponding period relative to the baseline period. Pre_2 is included not only to illustrate pretreatment trends but also to conduct reverse causality and time placebo tests. If the review is simply in response to positive demand shocks or if unobserved factors influence both the treatment and demand for the restaurants, then this term is expected to be significantly negative. **Figure 2.3** visually illustrates the regression results, showing that the pattern persists over the period; demand is significantly strong only for restaurants with positive reviews, and non-positive reviews see mostly positive but statistically insignificant sales growth with much fluctuation.

2.5.2 Location and Horizontal Consumer Sorting

There is no doubt that high quality products are more demanded than low quality ones, *ceteris paribus*, as a result of vertical product differentiations. But restaurants compete not only through quality differentiation but also in other dimensions[41] as individual diners have different tastes depending on their backgrounds such as race, ethnicity, income and education levels, and home or work locations. In reality, no one has perfect information on all available restaurants and their observed characteristics at any given point in time, and the media news informs millions of readers about the guided restaurants[42]. In this context, the media reviews could play a critical role in not only signalling restaurant quality but also providing consumers with details about each reviewed business. Thus publicity might be able to contribute to a better match between heterogeneous diners and restaurants even when the review is not positive.

From the consumers' viewpoint, one of the most important considerations in restaurant choices is location. Indeed, we choose a restaurant not only because of food quality or popularity of the venue but also because of its location, in particular where each neighborhood features different social and cultural characteristics. Customers may visit a reviewed restaurant that has no quality indicator but is located near their home/workplace

[40] For example, Pre_2_{it} is equal to 1 if restaurant i is reviewed by NYT and week t is between the 8th and the 5th pre-treatment week, and 0 otherwise. Likewise, $Post\_1_{it}$ is a dummy taking the value 1 if week t is between the 1st and the 4th post-treatment week of newly reviewed venue i, and $Post\_13_{it}$ is from the 49th to the 52nd post-treatment week of restaurant i.

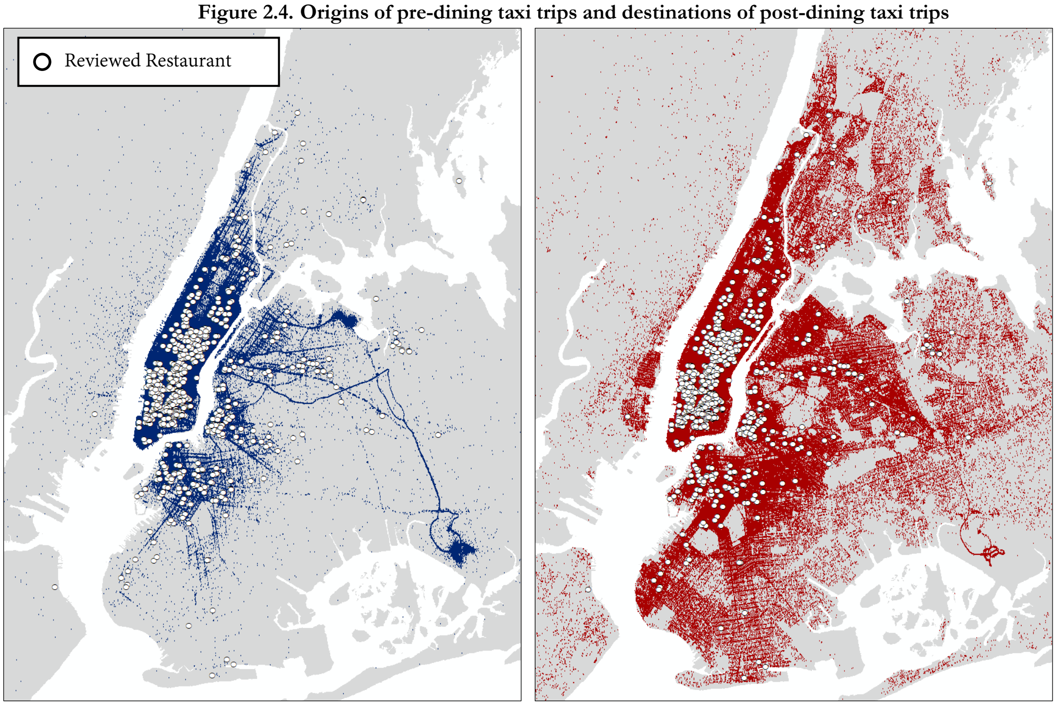
[41] In other words, services are both vertically and horizontally differentiated in the restaurant industry.

[42] The *New York Times* is one of the top 3 U.S. newspapers in circulation, according to Alliance for Audited Media (auditedmedia.com)

or in one of their favorite hangouts. In the meantime, they may be reluctant to visit a restaurant located in neighborhoods that they rarely travel to, even if the restaurant has got four stars. Consumer perceptions of restaurant quality may be determined not only by how many stars the restaurant was awarded, but also by to what extent observed characteristics satisfy their idiosyncratic multi-dimensional tastes. Hence it is plausible that zero-star restaurants in a popular hangout are more demanded than four-star restaurants in another in response to NYT reviews. For this reason, this section investigates whether media attention can result in restaurant-diner matches, with a focus on locational forces possibly affecting consumer restaurant choices.

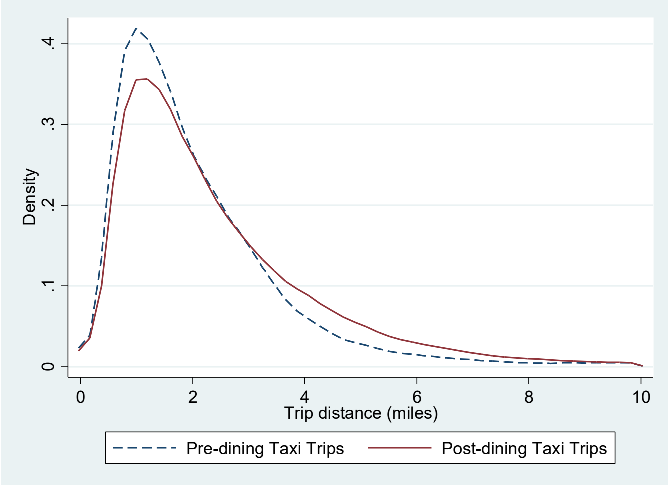
To do so, this study examines roles played by demographic characteristics of restaurant locations. As documented in Davis et al. (2018), diners are more likely to choose restaurants in neighborhoods that are demographically similar to their own. Hence a restaurant's demand could be driven by demographic similarities between its location and diners' homes or between that location and the ethnic identity of diners. However, it does not necessarily mean that diner decisions are directly influenced by interactions between restaurant and home locations. It is more likely that individuals' unobservable preferences for particular sociocultural environments determine both their home and preferred hangout locations.

Figure 2.4. Origins of pre-dining taxi trips and destinations of post-dining taxi trips



Notes: The figures show the geographical distributions of the origins of 14,855,705 taxi trips that arrived at the NYT-reviewed restaurants between 18:00 and 22:00 (Left panel), and the destinations of 15,224,286 taxi trips that departed from the same restaurants between 20:00 and 24:00 (Right panel) from January 2009 to June 2016.

Figure 2.5. Kernel density estimates of pre- and post-dining taxi trips



Notes: This figure shows kernel density estimates of trip distances of the pre- and post-dining taxi trips displayed in Figure 2.4, using only taxis of which trip distances are within 10 miles. Pre-dining taxi trips (blue dash line) include 14,855,705 taxis that arrived at the NYT-reviewed restaurants between 18:00 and 22:00, and Post-dining trips (red solid line) include 15,224,286 taxis that departed from the same restaurants between 20:00 and 24:00 from January 2009 to June 2016.

Table 2.8.Summary statistics of pre- and post-dining taxi trips

	Pre-dining trips		Post-dining trips	
	Mean	SD	Mean	SD
Trip distance (miles)	2.1222	1.7278	2.4947	1.9588
Fare (USD)	9.7851	5.4808	10.4339	48.1491

Notes: This table shows summary statistics of the pre- and post-dining taxi trips described in Figures 2.4 and 2.5.

I investigate locational interactions empirically by exploiting diners' neighborhood characteristics as inferred from the destinations of taxis that departed NYT-reviewed restaurants during evening peak hours. The left and right panels of **Figure 2.4** display origins of pre-dining taxi trips which occurred during the peak hours between 6pm and 10pm, and destinations of post-dining trips from 8pm to 12am, respectively. The locations of origins and destinations may contain useful information on taxi riders. Most of the origins are clustered in Manhattan or near reviewed restaurants, but post-dining taxi trips are spread out over the larger areas including some parts out of New York City. Considering that a lot of consumers visit restaurants after work and take a taxi to get back home after dining, the origins and destinations are likely to partially reflect locations of work and home, respectively[43]. Based on this pattern, I match each destination location to a census tract in order to infer consumer characteristics from the tract-level demographic

[43] It appears that the restaurant choices are more affected by workplaces rather than residential locations.

characteristics[44], as actual individual-level consumer characteristics are not available in this study. Then I relate the inferred information to locational characteristics of the restaurants that, presumably, they visited. Although the inferred information would not precisely nor fully reflect true interactions between consumers and restaurants, the difference-in-differences estimates would be valid unless the reviews influence diners' choice of transport mode or post-dining destinations in the short run.

Table 2.9.Summary statistics: Census tracts of Manhattan and New York City

	Census Tract level					
	Manhattan		New York City		Manhattan	New York City
	Mean	S.D.	Mean	S.D.		
Share of white	56.9%	27.0%	43.6%	29.4%	57.4%	44.0%
Share of Asian	11.9%	13.8%	12.6%	15.5%	11.3%	12.7%
Share of black	17.1%	21.5%	27.2%	31.0%	15.6%	25.5%
Share of Hispanic	23.4%	23.1%	26.7%	22.5%	25.4%	28.6%

Notes: This table shows summary statistics of demographics in New York City. The first four columns are Census-tract-level means and standard deviations, and the last two columns are Manhattan and NYC regional statistics.

Table 2.10. Summary statistics: Census tracts of restaurants and inferred individual homes

	Nonpositive Reviews				Positive Reviews			
	Restaurant tracts		Home tracts		Restaurant tracts		Home tracts	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Share of white	66.7%	21.6%	67.2%	10.7%	67.9%	21.1%	69.0%	9.1%
Share of Asian	15.1%	15.9%	13.4%	5.4%	15.8%	16.1%	13.8%	4.5%
Share of black	8.3%	13.9%	10.2%	9.3%	7.2%	12.5%	8.9%	7.8%
Share of Hispanic	16.2%	15.3%	14.6%	7.4%	14.6%	14.7%	13.4%	6.1%
price signs ($-$$$$)	1.98	0.79			2.59	0.97		

Notes: This table summarizes Census-tract-level demographic statistics for restaurant locations and diners' residential locations inferred from the destinations of post-dining taxi trips. It presents demographic similarity between restaurant locations and diners' home locations.

Demographic characteristics of the inferred home tracts are strongly correlated with those of restaurant tracts. **Table 2.10** summarizes tract-level demographics of the 807 restaurants included in my sample and those of the inferred home locations. Compared to both Manhattan and NYC averages (**Table 2.9**), the reviewed restaurants are located in tracts with substantially more white and Asian but less black and Hispanic residents. The tract-level average shares of white residents are 56.9% and 43.6% in Manhattan and New York City, respectively. But the white population is overrepresented in the tracts of reviewed restaurants. Likewise, Asian residents account for only 12.7% of the NYC population, but more than 15% in the restaurant locations. The strong similarity between restaurant and home tracts is more obvious when comparing the positive and nonpositive review groups. Positively reviewed restaurants are in tracts with marginally more white and Asian but less black and Hispanic populations than nonpositively reviewed ones. Noticeably, consumers who visited restaurants with a positive review also come from

[44] The US Census Bureau provides tract-level demographic information, and this study relies upon the Census 2010 data.

tracts with more white and Asian but less black and Hispanic populations than those who visited venues with a nonpositive review. While Davis et al. (2018) suggest that consumption segregation might be associated with residential segregation, it appears that this demographic similarity is a consequence of short trip distances seen in **Figure 2.5** and **Table 2.8**, as adjacent two tracts are likely to feature similar demographic characteristics.

Table 2.11.The effect of New York Times reviews by fixed effects specification

	Dependent Variable: ln(Passenger)			
	(1)	(2)	(3)	(4)
PostNYT	0.0573***	0.0780***		
	(0.0200)	(0.0147)		
PostNYT × Positive			0.0590***	0.0673***
			(0.0221)	(0.0163)
PostNYT × Nonpositive			0.0504	0.113***
			(0.0392)	(0.0329)
Fixed Effects				
time window	YES	YES	YES	YES
week × zipcode × cuisine × price	YES		YES	
week × cuisine × price		YES		YES
Observations	165,536	165,536	165,536	165,536
Adj. R-squared	0.947	0.930	0.947	0.930

Notes: This table shows regression results of Model 4 specification in Table 2.7. Columns 1 and 3 control for the same week-zipcode-cuisine-price fixed effects of the previous specifications, but columns 2 and 4 include week-cuisine-price fixed effects to estimate the effects of locational factors. The dependent variable is the log of the number of taxi passengers. *PostNYT* is an interaction term of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. *Positive* and *Nonpositive* are indicators capturing whether the review is positive or non-positive. All models rely on a 17-week time window, and control for a spatially lagged dependent variable and individual-restaurant-specific time window fixed effects. Robust and neighborhood-level clustered standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

To estimate an effect of locations, this section assumes that diners choose a restaurant across zip codes, not within a zip code. To do so, following models rely on cross-sectional variation within week-cuisine-price groups. By using the number of passengers who left reviewed restaurants between 20:00 and 24:00, **Table 2.11** shows that NYT reviews have a larger impact with the week-cuisine-price fixed effects (column 2) than with the week-zipcode-cuisine-price fixed effects (column 1). Also, the coefficient on nonpositive reviews becomes more sizable than that on positive reviews (column 4). These results tell us that nonpositively reviewed restaurants are not significantly more demanded than the other comparables within the same zip code, but that they are more demanded than similar restaurants in another zip codes. Obviously, quality indicators are significant demand drivers when control groups are restricted to alternative restaurants within a zip code. However, more factors, particularly locational characteristics, are expected to play a key role in diners' restaurant choice problems when across-zip code restaurants are considered as control groups. An empirical concern arising from the absence of location (zip code) in the week fixed effects might be non-parallel trends as hot or hip neighborhoods are likely to have more restaurants reviewed and also to see increasing taxi trips. But a donut control, or a spatially lagged dependent variable, could substantially relieve the concern.

To identify heterogeneity in consumer tastes for demographics of restaurant locations, as observed product characteristics, the usual approach is conditional logit discrete choice models which can consider both consumer tastes and product characteristics and also their interactions. However, the approach is not computationally feasible given millions of taxi trips and hundreds of alternative restaurants in my sample. Instead of the choice models, I use segmented aggregate demand to achieve similar results at an aggregate level rather than choice data at the individual level. Let S_{ij}^{d} be a fraction of demographic group d in the destination tract of taxi trip j that departed from restaurant i during the peak hours. Then N_i^d, the inferred total number of diners in demographic segment d who choose restaurant i, is obtained by summing the shares of the racial group:

$$\text{(2.4)} \quad N_i^d = \sum_j n_{ij} \, S_{ij}^d$$

where n_{ij} is the number of passengers of trip j that departed from restaurant i. This segmented demand would reflect a specific racial/ethnical group's preference for each restaurant. Hence the following equation would capture demographic interactions between locations of homes and restaurants.

$$\text{(2.5)} \quad ln\, N_{it}^d = w_i + \lambda_{tcp} + \beta_1 \left(PostNYT_{it}\right) + \beta_2 \left(PostNYT_{it} \times Share_{it}^d\right) + \boldsymbol{x'}_{\boldsymbol{i,t}} \boldsymbol{\gamma} \; + \varepsilon_{itcp}$$

where $Share_{it}^d$ is the share of demographic segment d in the census tract of restaurant i. Notably, Equation 2.5 interprets, for example, $N_{it}^{Hispanic}$ as the probabilistic total number of Hispanic diners who visited restaurant i in week t, but a precise interpretation is not whether Hispanic diners are more likely to visit a restaurant in a Hispanic neighborhood but whether diners residing in a Hispanic area are more likely to visit a restaurant in another predominantly Hispanic neighborhood. It is possible that diners choose reviewed restaurants not because of demographic similarity between the two locations but because the retailers are geographically near their post-dining destinations. Also, the specification may capture inter-neighborhood economic interactions rather than demographic similarity, as racial wealth divide is reportedly growing. For example, diners in a poor neighborhood are more likely to visit a restaurant in another poor neighborhood, and they are more likely black diners than white diners. For these reasons, this specification controls for a vector of covariates ($\boldsymbol{x_{i,t}}$) including the log of average trip distance and the log of average income of diners[45] in addition to a Michelin-star indicator and a spatially-lagged dependent variable.

Table 2.12.The effect of New York Times reviews: Demographic interaction 1

Dependent Var.	(1) ln(Asian)	(2) ln(Asian)	(3) ln(Black)	(4) ln(Black)	(5) ln(Hispanic)	(6) ln(Hispanic)
PostNYT	0.0860***	0.0803***	0.0990***	0.0640***	0.0797***	0.00711
	(0.0170)	(0.0228)	(0.0175)	(0.0190)	(0.0159)	(0.0232)
PostNYT × %Asian		0.0366				
		(0.0858)				
PostNYT × %Black				0.601***		
				(0.149)		
PostNYT × %Hispanic						0.576***
						(0.109)
Observations	152,795	152,795	152,798	152,798	152,802	152,802
Adjusted R-squared	0.911	0.911	0.853	0.853	0.894	0.895

Notes: This table shows regression results of Equation 2.5. The dependent variables are the log of the number of Asian diners (columns 1 & 2), Black diners (columns 3 & 4), and Hispanic diners (columns 5 & 6) inferred from destinations of post-dining taxi trips; for more details, see text. *%Asian*, *%Black*, and *%Hispanic* are the shares of Asian, Black, and Hispanic residents in the restaurant's census tract, respectively. *PostNYT* is an interaction of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. All models rely on a 17-week time window, and control for a spatially-lagged dependent variable, the log of average trip distance, the log of average (inferred) income of diners, and two fixed effects, restaurant × time window and week × cuisine × price. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.12 uses the natural logarithms of the segmented aggregate demand of three groups, Asian, black, and Hispanic, as dependent variables, and interacts the variable of interest (PostNYT) with the share of each corresponding demographic group in a restaurant tract. When a reviewed restaurant is located in a neighborhood demographically more similar to diners', diners are more likely to visit the venue than another restaurant of the same cuisine and price ranges but located in a demographically less similar location. Thus estimated coefficients in columns (2), (4), and (6) are expected to capture heterogeneous treatment effects caused by demographic similarity of two locations. The results indicate that customers' post-dining destinations are significantly different by demographic characteristics of restaurant locations, so the marginal utility of location may vary according to demographic groups. Notably, black and Hispanic diners have stronger preference for demographic similarity than Asians. Specifically, 10% more shares of black or Hispanic population in a restaurant tract attract 6.0% more black or 5.8% more Hispanic diners, respectively (columns 4 and 6). This finding is fairly consistent with what existing literature on segregation reports; African Americans are the most segregated

[45] This study infers the aggregate income of diners by summing the tract-level incomes:

$$INCOME_i = \sum_j n_{ij}\ income_{ij}$$

where $INCOME_i$ is the aggregate income of diners who visited restaurant i. n_{ij} is the number of passengers of taxi trip j that departed from restaurant i during the peak hours, and $income_{ij}$ is a median household income of the destination tract of the post-dining trip j.

minority, and Asians are the least segregated in term of both residence (Ihlanfeldt and Scafidi, 2002; Flores and Lobo, 2013) and restaurant choices (Davis et al., Forthcoming); Davis et al. (Forthcoming) identify racial and ethnic consumption segregation for the NYC residents from user reviews on Yelp.com. As Yelp users very rarely review a restaurant twice, the authors' baseline model, a discrete-choice model of restaurant visits, suffers from a selection bias, and therefore assumes that individuals' unobserved restaurant preferences are independent over time. However, this study is not likely to suffer from the same bias unless the restaurant review publications alter diners' transportation mode choices or have a more significant effect on diners who depend on taxis.

Table 2.13.The effect of New York Times reviews : Demographic interaction 2

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.	ln(Asian)	ln(Asian)	ln(Black)	ln(Black)	ln(Hispanic)	ln(Hispanic)
PostNYT × Positive	0.0731***	0.0486**	0.0860***	0.0536***	0.0688***	0.0493
	(0.0182)	(0.0243)	(0.0196)	(0.0198)	(0.0183)	(0.0325)
PostNYT × Nonpositive	0.126***	0.169***	0.139***	0.0970*	0.113***	-0.0607
	(0.0365)	(0.0583)	(0.0417)	(0.0491)	(0.0342)	(0.0768)
PostNYT × Positive × %Asian		0.158*				
		(0.0929)				
PostNYT × Nonpositive × %Asian		-0.279				
		(0.215)				
PostNYT × Positive × %Black				0.571***		
				(0.130)		
PostNYT × Nonpositive × %Black				0.673*		
				(0.348)		
PostNYT × Positive × %Hispanic						0.155
						(0.263)
PostNYT × Nonpositive × %Hispanic						1.198***
						(0.447)
Observations	165,462	165,462	165,465	165,465	165,486	165,486
Adjusted R-squared	0.911	0.911	0.842	0.842	0.889	0.889

Notes: Based on Equation 2.5, this table interacts each of the variables of interest with two dummies, *Positive* and *Nonpositive*, that capture whether the review is positive or non-positive. The dependent variables are the log of the number of Asian (columns 1 & 2), Black (columns 3 & 4), and Hispanic diners (columns 5 & 6). *%Asian*, *%Black*, and *%Hispanic* are the shares of Asian, Black, and Hispanic residents in the restaurant's census tract, respectively. *PostNYT* is an interaction of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. All models rely on a 17-week time window, and control for a spatially-lagged dependent variable, the log of average trip distance, the log of average (inferred) income of diners, and two fixed effects, restaurant × time window and week × cuisine × price. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.13 provides further support to the results. Even when a restaurant achieves no positive indicator, a 10 percentage point increase in the fraction of black (Hispanic) residents in a restaurant tract is significantly associated with a 6.7% (12.0%) increase in black (Hispanic) diners, as seen in columns 4 and 6. Interestingly, Hispanic diners do not appear to care about the demographical environment of a restaurant when the business is positively reviewed. In contrast, Asian consumers consider the locational characteristics only when the retailer wins positive indicators.

Table 2.14.The effect of New York Times reviews by trip distance : Post-dining taxi trips

	Dependent Variable: ln(Passenger)					
	(1)	(2)	(3)	(4)	(5)	(6)
Trip Distance (miles)	0 to1	1 to 2	2 to 3	3 to 4	4 to 5	above 5
PostNYT × Positive	0.0709***	0.0611***	0.0548***	0.0439**	0.0502**	0.0808***
	(0.0172)	(0.0196)	(0.0166)	(0.0192)	(0.0208)	(0.0253)
PostNYT × Nonpositive	0.123***	0.0950**	0.0786**	0.119***	0.103**	0.0523
	(0.0347)	(0.0420)	(0.0336)	(0.0385)	(0.0395)	(0.0429)
Observations	152,949	135,294	140,690	135,354	130,589	122,885
Adj. R-squared	0.929	0.834	0.879	0.824	0.747	0.677

Notes: This table shows whether geographical proximity has an effect on the coefficients of interest. The dependent variable is the log of the number of taxi passengers, but each model column uses only taxis of which trip distances are within the correspondingly specified range; for example, Model 1 counts the number of passengers whose trip distances are between 0 and 1 mile. *PostNYT* is an interaction term of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. *Positive* and *Nonpositive* are indicators capturing whether the review is positive or non-positive. All models rely on a 17-week time window, and control for a spatially-lagged dependent variable, and two fixed effects, restaurant × time window and week × cuisine × price. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.15.The effect of New York Times reviews : Income interaction

	Dependent Variable: ln(Aggregate Income)			
	(1)	(2)	(3)	(4)
PostNYT	0.0873***	1.004**		
	(0.0160)	(0.474)		
PostNYT × ln(Income)		-0.0801*		
		(0.0409)		
PostNYT × Positive			0.0728***	0.728
			(0.0173)	(0.470)
PostNYT × Positive × ln(Income)				-0.0573
				(0.0410)
PostNYT × Nonpositive			0.133***	1.518
			(0.0377)	(1.333)
PostNYT × Nonpositive × ln(Income)				-0.122
				(0.116)
Observations	152,781	145,557	152,781	145,557
Adjusted R-squared	0.927	0.928	0.927	0.928

Notes: To capture economic interactions between locations of the restaurant and diners' homes, this table uses the log of the aggregate income of diners as the dependent variable. *ln(Income)* is the log of median household income of the restaurant tract. *PostNYT* is an interaction of two indicators capturing whether the restaurant is reviewed by NYT, and whether the week is a post-treatment period. *Positive* and *Nonpositive* are indicators capturing whether the review is positive or non-positive. All models rely on a 17-week time window, and control for the log of spatially-lagged dependent variable, and the log of average trip distance, and include two fixed effects, restaurant × time window and week × cuisine × price. Robust and neighborhood-level clustered standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Additionally, **Table 2.14** tells us that the geographical proximity also has an effect. Diners might be willing to visit nearby restaurants in spite of non-positive reviews. To examine the possibility, I break the number of passengers down into six groups by taxi trip distance: 0-1, 1-2, 2-3, 3-4, 4-5, above 5 miles. If proximity is a main driver, the number of taxi trips from nearer origins would increase more significantly in response to publicity. The size and statistical significance of coefficients decrease with taxi trip distance from column 1 to 3 (or from 0 to 3 miles) even when the reviews are not positive. Spatial proximity may be an even more important consideration for diners who visit restaurants on foot or by public transport. However, the income interaction has no significant effect (**Table 2.15**); each model uses tract-level incomes instead of shares of a demographic group in a restaurant tract, and the dependent variable is (inferred) aggregate income of diners. Overall, socio-demographic characteristics of restaurant locations can have a substantial impact on urban consumers' decisions with geographical proximity also having an effect. The results suggest that media attention influences not only vertical but also horizontal consumer sorting, thereby possibly resulting in more demographically segregated consumption.

2.6 Conclusion

In this internet era, media information could influence our decisions in several ways. A wide range of studies have emphasized its importance in social and economic contexts, but only a limited amount of empirical evidence has been documented. To add to existing literature, this paper characterizes how media attention could affect consumer choices. Before identifying potential channels, I begin with examining whether publicity has a causal effect on demand for restaurants, relying upon two information sources, the *Michelin Guide* and NYT restaurant reviews, and also the number of taxi drop-offs as a proxy of restaurant demand. The link between quality information and taxi trips is surprisingly strong and statistically significant.

Then what drives such significant effects? Consumers' expectations of product quality are a primary determinant of demand, and thus positive indicators can shape their expectations. At the same time, media information has a potential to improve a match between heterogeneous consumers and products by increasing product awareness. The regression results demonstrate that positive quality indicators, like stars, are strong demand drivers. Most of the existing studies have concentrated only on this mechanism mainly due to lack of individual-level characteristics. By inferring individual diners' characteristics from destinations of post-dining taxi trips, this study overcomes the empirical obstacle, and empirically shows that demographic characteristics of locations play a key role in informed consumers' restaurant choices. Thus this paper concludes from the findings that media information could affect economic outcomes of urban retailers not only by signaling product quality but also by catalyzing interactions between consumer and product characteristics.

CHAPTER 3.

Housing Prices and Consumption: The Role of News Media

3.1 Introduction

Housing prices can affect a wide range of socio-economic outcomes, particularly homeowner non-housing consumption.[46] It is commonly assumed that lowering interest rates can induce an increase in asset prices including house prices, which can boost consumer spending and ultimately the economy. Therefore, a key question in macroeconomics and monetary policy is how strong is the link between asset prices and aggregate consumption. As a consequence, there exists a large literature on the wealth effect, with housing prices reportedly having larger and more important impacts on spending than do stock prices (Case et al., 2005; Bostic et al., 2009; Carroll et al., 2011; Calomiris et al., 2012; Case et al., 2012). However, the existence of and mechanisms for the effects have been controversial with somewhat conflicting theoretical predictions.[47] Recent research provides strong empirical evidence from clean identification (Mian et al., 2013; Aladangady, 2017), but observed effects are still inconsistent over time. [48]

To provide an explanation to the heterogeneity in observed housing wealth effects, this paper examines whether information interventions may alter homeowners' consumption decisions in response to house price fluctuations. The traditional view in the wealth effect literature is to assume that homeowners are making fully informed utility-maximizing

[46] A few more examples are fertility rates (Lovenheim and Mumford, 2013; Dettling and Kearney, 2014), college entrance (Lovenheim, 2011), entrepreneurship (Corradin and Popov, 2015), and portfolio choice (Chetty and Szeidl, 2017).

[47] The standard channel relaxes the household lifetime resource constraint. That is, higher home values have a positive endowment effect, and thus rational households maximize their utility by consuming more. Another channel is collateralized lending. Rising housing prices allow households to borrow more by providing additional collateral. As housing is not only an investment asset but also a consumption good, however, rising home values could raise the future cost of living, and this negative effect could offset the positive wealth effect. Therefore, aggregate housing wealth effects should be small for aggregate non-housing consumption (Sinai and Souleles, 2005; Buiter, 2008; Calomiris et al., 2012).

[48] For example, housing prices have already recovered to the pre-crisis level, but economic growth is slowing in the US and many other countries; see the newspaper article (https://www.cnbc.com/2016/08/01/debt-is-holding-back-the-global-economic-recovery-say-central-bankers-dudley-rajan-and-zeti.html).

choices, so one would not expect provision of additional information about housing prices to change homeowners' consumption behaviors. However, this paper takes a different view motivated by the behavioral economics literature that suggests that any forms of information disclosure can have a significant impact on household economic decisions (Choi et al., 2010; Bertrand and Morse, 2011; Beshears et al., 2018). If homeowners do not have high level of awareness of their home values, a sizable housing wealth effect can hardly be expected, given the considerable house price fluctuations that occur over time. In reality, many homeowners do not actively seek housing price information in their daily lives. Also, individuals' cognitive resources are very limited, so it is possible that homeowners might be making a cognitive lapse. In this sense, media coverage on housing prices may affect homeowners' consumption decisions by increasing homeowners' awareness about their housing wealth and also by de-biasing the cognitive lapse; one might expect their consumption decisions to respond to how much or how frequently information on housing price growth is being disclosed.

To allow the size of the housing wealth effect to vary according to local media coverage on house prices, I exploit local newspapers, one of the main information sources for homeowners.[49] People often base their economic, financial, and political decisions on the news that they read in newspapers or watch on television. For example, related studies suggest that geographic areas with reduced local media coverage see less stock trading volumes (Engelberg and Parsons 2011) and lower voter turnouts (Gentzkow, et al. 2011).[50] Therefore, information from local media outlets may increase people's awareness on the given topic, and then the increasing awareness may shape individuals' behaviors and decisions. Literature on the relation between news media and real estate/housing markets is particularly scarce, but several recent studies identify the causal effects of local newspapers and internet social media on homebuyer decisions. By employing textual analysis on local U.S. newspaper articles, Soo (2018) demonstrates that the qualitative tone of local housing news can predict future house prices. Bailey et al. (2018) also show that potential homebuyers rely on information from Facebook. These findings imply that the information channels of both print and online media may play a key role in informing homebuyers and shaping their economic decisions. Therefore, media coverage may make homeowners more

[49] More than a quarter of all adults in the US read at least one newspaper every day (https://www.statista.com/statistics/183408/number-of-us-daily-newspapers-since-1975/)

[50] Engelberg and Parsons (2011) exploit the daily newspaper of each major U.S. city, and find that local newspaper coverage of earnings announcements significantly increases local trading volumes. Gentzkow et al. (2011) document that reading a newspaper increases the probability of voting by 4 percentage points.

aware of their housing wealth, and lead them to further learn their home values from previous sales or current listings by visiting real estate websites such as Zillow.com.

This paper relates the local media coverage on housing prices to the housing wealth effects by regressing household-level consumption on the MSA-level housing price index interacted with the number of newspaper articles conveying house price information.[51] First, household-level expenditures are from the public-use microdata of the quarterly Consumer Expenditure Survey (CEX) across 22 MSAs over 11 years from 2006 to 2016. Second, I use the MSA-level house price index provided by the Federal Housing Finance Agency (FHFA). I assume that homeowners learn their home values in part from the regional price growth covered by local newspapers. Last, I exploit local newspaper contents in the US in order to quantify media coverage on housing prices. From three newspaper databases, I collected newspaper articles including any of the search queries, "home price," "house price," and "housing price", and then I counted the total number of news articles published by major local newspapers during each quarter in each city.

Regression results robustly show that more newspaper articles conveying housing price information can make homeowner consumption more elastic with respect to regional housing prices. In other words, relative to less informed homeowners, the more informed consume more in response to high housing prices and consume less in response to low housing prices. By stratifying the analysis for homeowners and renters, I find that the regression results are statistically significant only for homeowners; an increase of one standard deviation in the number of housing price news articles is associated with a 0.08 increase in homeowners' consumption elasticity and only with an insignificant increase in renters' elasticity. Since rising home values do not clearly benefit renters, this result alleviates the empirical concern regarding unobservable common factors, in particular, expected future income growth.

For further causal investigations, I also identify the headline effect. Specifically, only housing news articles that include housing or real estate terms in their headline—headlined housing news—have a significant impact on the wealth effect. However, articles that convey housing price information in their body but do not include any housing or real estate terms in their

[51] MSA stands for Metropolitan Statistical Area, a U.S. geographical core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core (https://www.census.gov/programs-surveys/metro-micro/about.html).

headline—non-headlined housing news—have no significant effect, thereby providing further support for the causality. It may be random whether or not each housing news article includes housing terms in its headline. That is, similar unobservable factors may influence supply of both headlined and non-headlined housing news articles, but demand sides, homeowner expenditures, respond differently to the exogenous changes in the headlines.

Still, one may be concerned about omitted macroeconomic factors that may drive all of the variables of interest. Also, the effect of media coverage may be asymmetric in response to housing booms and busts. To address the concerns, this study employs two regression models. The first baseline model exploits both cross-sectional and time-series variation, but the second exploits only cross-sectional variation in media coverage by taking 11-year averages for each city. Estimates from the two specifications are consistently significant, confirming robustness of the baseline specification.

To the best of my knowledge, this is the first to relate information interventions to the housing wealth effect. Arguably, the dependence of the housing wealth effect on media reporting has a couple of important implications. First of all, this idea can shed light on inconsistent housing wealth effects, given the substantial variation in media reporting across cities and over time. Information interventions may make the aggregate consumption or economy more responsive to asset prices, impacting the effectiveness of related policies. Second, this paper shows that information disclosure has a nontrivial effect on homeowner consumption decisions. In contrast with the view that household decisions reflect fully informed and rational behaviors, the result suggests that providing relevant information can alter household economic decisions by helping them to make more informed choices and/or by making the information salient to the individuals. Thus, the finding has a potential to dramatically broaden our understanding of the role of news media and information disclosure across a large number of settings.

3.2 Data

3.2.1 Data Sources

For empirical tests, I construct a quarterly dataset for 22 U.S. metropolitan areas by exploiting the Bureau of Labor Statistics' (BLS) Consumer Expenditure Survey (CEX) public-use microdata from 2006 to 2016.[52] The CEX data provides a good measure of household consumption and consists of two surveys: the Diary survey and the Interview survey. While the Diary survey is designed to capture small expenditures on frequently purchased items such as food over a two-week period, the Interview survey is conducted quarterly for major expenses that occur on a regular basis. Although the surveys are reportedly subject to an underreporting issue, this paper nevertheless uses the Interview survey data[53], following two papers (Bostic et al., 2009; Aladangady, 2017) that find significant wealth effects using the data. The Interview surveys collect detailed household-level information on expenditures, incomes, and household characteristics. As each household is interviewed for only four consecutive quarters, with new families entering each quarter, the dataset is basically repeated cross-sections. [54] A limitation is that detailed geographical information has been de-identified. Only about 20 major metropolitan areas are identified in the public-use data.

The household-level datasets enable me to compare different responses between homeowners and renters. **Table 3.1** reports summary statistics of major economic and socio-demographic characteristics for each family or reference person included in the controls.[55] For example, I control for household-level income levels in my models because both expenditures and incomes are noticeably higher for homeowners; all of the variables of interest, consumption, homeownership decisions and housing prices may be affected by income levels. In addition, the majority of homeowners are married, white, and hold a

[52] The CEX data is available from 1986, but the public-use microdata contains MSA-level geographical identifiers from 2006 onward.

[53] This paper assumes that the underreporting is independent of media coverage.

[54] The observation unit in the CEX is called the consumer unit (CU), which consists of any of the following: (1) all members of a particular household who are related by blood, marriage, adoption, or other legal arrangements; (2) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in permanent living quarters in a hotel or motel, but who is financially independent; or (3) two or more persons living together who use their incomes to make joint expenditure decisions. However, the terms *consumer unit*, *family*, and *household* are used interchangeably. For further details, see the BLS webpage https://www.bls.gov/cex/csxfaqs.htm.

[55] According to the BLS webpage (https://www.bls.gov/cex/csxfaqs.htm#PUMD), the reference person for the consumer unit is the first member mentioned by the respondent when asked "What are the names of all the persons living or staying here? Start with the name of the person or one of the persons who owns or rents the home." It is with respect to this person that the relationship of the other consumer unit members is determined. Thus, the two terms *reference person* and *head of household* are used interchangeably in related studies.

bachelor's degree or higher, while renters are more likely to be single and black, with only 29% holding a bachelor's degree.

Table 3.1. Summary statistics of Consumer Expenditure Survey (2006-2016)

	Home owners (59.5%)		Renters (40.5%)	
	Mean	S.D	Mean	S.D.
Total Expenditure (USD, quarterly)	16,579	14,666	9,814	7,667
Income (USD, annual)	79,403	90,477	37,775	46,012
Family size	2.68	1.51	2.34	1.53
Age	53.55	15.48	43.68	17.24
Public sector employer	0.42	0.49	0.37	0.48
Private sector employer	0.49	0.50	0.57	0.49
Married	0.63	0.48	0.32	0.47
Never married	0.13	0.34	0.39	0.49
Separated/divorced/widowed	0.24	0.43	0.28	0.45
White	0.81	0.39	0.68	0.47
Black	0.10	0.30	0.20	0.40
Asian	0.07	0.26	0.09	0.29
High school or lower	0.29	0.45	0.42	0.49
Some college or Associate's degree	0.27	0.45	0.29	0.45
Bachelor or higher	0.44	0.50	0.29	0.45

Notes: This table summarizes descriptive statistics for variables from the Consumer Expenditure Survey (CEX) data. It presents differences in family characteristics between homeowners and renters. As each household is interviewed for at most four consecutive quarters, the quarterly dataset is repeated cross-sections, and the sample includes 38,938 households or 110,677 observations from Q1.2006 to Q4.2016.

To measure the housing wealth effects, I link the household-level spending data from the CEX to the MSA-level house price index provided by the Federal Housing Finance Agency (FHFA).[56] Many empirical papers rely upon household balance sheet information to test whether household spending responds to the self-reported market value of their home. Instead of the household-level housing wealth measure, I exploit the local housing price index because this is the most common housing price information that homeowners can obtain from media outlets, assuming that they learn their home values in part from the regional price growth covered by local newspapers. By doing so, this study can identify the differential wealth effects on homeowners and renters.

Lastly, I depend on three newspaper databases—Factiva, Nexis, and Newslibrary.com—to quantify the media coverage on housing price. First, I identified the dominant local newspapers for 22 cities (**Table 3.2**). U.S. newspapers have historically been local in nature. The number of US cities that can support multiple daily newspapers is fast declining (Chandra and Kaiser, 2015), and the median newspaper sells more than 90 percent of its copies in the county in which it is headquartered (Gentzkow and Shapiro, 2010). When multiple newspapers are presented for one city in **Table 3.2**, this is mostly due to the

[56] The FHFA All Transactions House Price Index is used in this study.

geographical boundaries of the metropolitan areas in the CEX data. For example, one sampling unit of the CEX consists of San Francisco, Oakland, and San Jose. As a consequence, in addition to the *San Francisco Chronicle*, both the *Oakland Tribune* and the *San Jose Mercury News* are also included in my sample; importantly, this unique local monopoly (or duopoly) characteristic of newspapers enables researchers to exploit discontinuous cross-sectional variations in media coverage for identification.[57]

Then I collected news articles published by each newspaper, mainly from Newslibrary.com, and used the other two databases, Factiva and Nexis, to complement the datasets. I detail how I construct my measure of media coverage from this collection of newspaper articles in the following subsection.

Table 3.2. List of newspapers in the sample

Metropolitan Area	Newspapers Included	Missing Major Newspapers
Atlanta	Atlanta Journal-Constitution	
Baltimore	Baltimore Sun	
Boston	Boston Globe, Boston Herald	
Chicago	Chicago Sun-Times, Daily Herald, Courier-News	Chicago Tribune
Cleveland	Plain Dealer, Akron Beacon Journal	
Dallas	Dallas Morning News, Star-Telegram	
Denver	Denver Post	
Detroit	Detroit News	Detroit Free Press
Honolulu	Honolulu Star-Advertiser	
Houston	Houston Chronicle	
Los Angeles	Daily News Of Los Angeles, Orange County Register, Long Beach Press-Telegram	Los Angeles Times
Miami	Miami Herald	
Minneapolis-St. Paul	Star Tribune, Pioneer Press	
New York City	New York Times, New York Post, Daily News of New York, Star-Ledger	
Philadelphia	Philadelphia Inquirer, Philadelphia Daily News	
Riverside-San Bernardino	Press-Enterprise, San Bernardino Sun	
San Diego	San Diego Union-Tribune	
San Francisco	San Francisco Chronicle, San Jose Mercury News, Oakland Tribune	
Seattle	Seattle Times, Seattle Post-Intelligencer, News Tribune	
St. Louis	St. Louis Post-Dispatch	
Tampa	Tampa Bay Times (St. Petersburg Times), Tampa Tribune	
Washington	Washington Post, Washington Times	

Notes: Table 3.2 lists 22 metropolitan cities and corresponding local newspapers in my sample covering from January 2006 to December 2016. Most US cities have only one daily newspaper with some exceptions for that I include two major local newspapers (Boston, Philadelphia, Seattle, Tampa, and Washington) or three (New York City). The other cities with multiple newspapers presented in this table are due to the large geographical boundaries of the metropolitan areas in the Consumer Expenditure Survey (CEX) data. In addition, three major newspapers (Chicago Tribune, Detroit Free Press, and Los Angeles Times) are missing in my sample.

[57] For instance, Engelberg and Parsons (2011) take advantage of the fact that local media outlets often differ in their coverage of the same underlying information events. As another example, Gentzkow et al. (2011) and Peress (2014) utilize exits and entries of newspapers and reductions in media coverage caused by newspaper strikes, respectively.

3.2.2 Measuring Media Coverage

The single most important variable in this study is a measure of local media coverage on housing prices. A wide range of literature identifies causal impacts of news media, suggesting that information environments influence economic agents' behaviors and decisions.[58] Media news conveying house price information can not only help readers form expectations about future housing price growth, but they can also inform readers about the realized housing price growth. I place particular emphasis on the second role of news media, while most related papers focus on the first role. In other words, this paper is mainly interested in how media reporting can increase awareness about realized housing wealth gains rather than how media can help form future price expectations, as news articles are essentially informative. The quantity of news articles about housing prices may be positively correlated with how well local homeowners are informed about or aware of the values of their homes. In the absence of media reporting, household spending might be less responsive to housing price fluctuations, while it may become more responsive when homeowners are more frequently informed.

Table 3.3. An example of newspaper articles in the key independent variable

The Philadelphia Inquirer

August 26, 2009

Region's Home Prices Gain In Second Quarter

Driven by sales in the city, home prices in the Philadelphia region rose by an average 3.8 percent in the second quarter over the previous quarter - the first increase in two years. Data compiled by Kevin Gillen of Econsult Corp. in Philadelphia showed that even with sales volume nearly 50 percent below normal in the second quarter, homes that did sell sold for more. In the city, prices were 6.8 percent higher, while suburban counties showed an average price increase of 2.7 percent, Gillen said...

Notes: This is an example of housing news articles used to generate the key independent variable (I call it "the number of housing news articles") in my sample. The example article includes "home" in its headline, and "home price" in its body, conveying explicit information about local house price growth to readers.

To quantify the housing price information from local papers, I counted the total number of articles including specific keywords by city–quarter, using the articles collected from the three newspaper archives. More specifically, each article must include "home price," "house price," or "housing price" in its headline or body to ensure that it provides readers with explicit information on housing prices. In addition, each article also includes "home," "house," "housing," "real estate," or "property" in its headline. The additional queries further narrow the search results down to articles supposedly written for homeowners or homebuyers. For example, *The Philadelphia Inquirer*'s article in **Table 3.3** includes "home" in

[58] Largely in finance (Chan, 2003; Frazzini, 2006; Tetlock, 2007; Tetlock et al., 2008; Fang and Peress, 2009; Tetlock, 2010; Engelberg and Parsons, 2011; Tetlock, 2011; Da et al., 2014) and in political science (Gentzkow, 2006; Gerber et al., 2009; Gentzkow et al., 2011).

its headline and also "home price" in its body. This report provides readers with information on local housing price trends, and more importantly, the headline obviously indicates that the following body of the article is about housing markets and should therefore draw the attention of existing homeowners or potential homebuyers. Notably, the search query restrictions exclude articles that do not include the price keywords even though they discuss some aspects of housing markets or appear in the housing/real estate section. For instance, some housing news articles report on local building permits or new constructions, an important housing price determinant. Such information may help readers form expectations regarding their future house price growth but does not directly inform them about the realized housing wealth growth.

An empirical concern is that the search results for these keywords might possibly capture not only the trends of local housing markets but also those of national or international housing markets that might be less relevant to local housing prices. However, attention to readers' own properties is often triggered by spatially distant markets. A homeowner may search for the list prices of nearby similar properties after reading an article about a housing boom in another city. Bailey et al. (2018) also find that both local and out-of-state friends' experiences affect housing investments. This finding implies that all news articles containing housing price information have the potential to influence households' economic decisions, probably by making readers more and better aware of their housing wealth in any context.

3.3 Empirical Analysis

3.3.1 Housing Wealth Effect

Before including the MSA-level quarterly number of housing articles as a key independent variable in the econometric models, I first test a simple wealth effect model. The standard empirical functional form to measure the housing wealth effect is given in **Equation (3.1)**:

(3.1) $y_{i,c,t} = \beta\, \rho_{c,t-1} + \boldsymbol{x}'_{i,c,t}\, \boldsymbol{\gamma} + \lambda_t + \phi_{c,j} + \varepsilon_{i,c,t}$

where $y_{i,c,t} = ln\, Expenditure_{i,\,t}$ is the natural logarithm of the total expenditures of household i in city c in year-quarter t, and $\rho_{c,t-1} = ln(HPI_{c,\,t-1})$ is the natural logarithm of the house price index for city c in year-quarter $t-1$.[59] This model controls for a vector of family characteristics ($\boldsymbol{x}_{i,c,t}$: income, family size, age of household head), major factors affecting both consumption and housing demand. Also, I include year-quarter fixed effects (λ_t) and cohort fixed effects determined along seven dimensions ($\phi_{c,j}$: city × housing tenure × employer × occupation × marital status × education × race).[60] As discussed in the previous section, the dataset in this study is not a panel, but repeated cross-sections. In this context, Campbell and Cocco (2007) use sample cohort means from a time series of cross-sections to construct pseudo-panel data, following the methodology suggested by Deaton (1985). In doing this, Campbell and Cocco can exploit both time-series and cross-sectional variation to identify wealth effects. To obtain similar results, this study employs the cohort fixed effects that would capture average differences across cohorts in omitted variables; in other words, the regression coefficients of interst are driven by the variation over time within each cohort. Lastly, standard errors are clustered by city × housing tenure to allow for correlation over time within each group in all of the following models. With this log-log specification,[61] the estimated coefficient β can be directly interpreted as the elasticity of consumption with respect to housing wealth.

Then I split the variable of interest into two groups: owners and renters, as in **Equation (3.2)**:

$$\textbf{(3.2)} \quad y_{i,c,t} = \beta_1 \rho_{c,t-1} \times Owner_i + \beta_2 \rho_{c,t-1} \times Renter_i + \boldsymbol{x'}_{i,c,t}\,\boldsymbol{\gamma} + \lambda_t + \phi_{c,j} + \varepsilon_{i,c,t}$$

Renters do not clearly benefit from rising house prices. Thus, if this paper finds significantly positive relationship between housing wealth and the spending of renters, the estimated coefficients might be suggestive of an omitted variable bias. An empirical concern arising from splitting the independent variable is that home ownership decisions are endogenous (Campbell and Cocco, 2007). However, the cohort fixed effects can

[59] The main assumption of this specification is that consumption responds to home price information conveyed by local newspapers. Such information is usually a quarter lagged because home price indices are published with a few months' lag. Therefore, I let the housing price term refer to the previous quarter (*t*-1) instead of the current quarter (*t*).

[60] Individuals sharing the following seven characteristics are grouped into cohorts: 1) MSA; 2) Housing Tenure, with six categories: Owned with mortgage, Owned without mortgage, Owned mortgage not reported, Rented, Occupied without payment of cash rent, or Student housing; 3) Employer, with three categories: Private company, Government, or Self-employed/Family business; 4) Occupation, with four categories: Manager/Professional, Admin/Sales/Retail/Technician, Service, or Laborer/Production/Farming/Armed Forces; 5) Marital status, with five categories: Married, Widowed, Divorced, Separated, or Never married; 6) Education, with four categories: Master/Professional/Doctorate degree, Bachelor's degree, College/Associate degree, or High School/Less; and 7) Race, with three categories: White, Black, or Others.

[61] One is added to the relevant variables to avoid logarithms of zero.

address this concern by exploiting variations within the homeowners or within the renters over time.

Table 3.4 reports the estimated wealth effects for homeowners and renters. A well-known empirical issue in the related literature is that both consumption and housing prices can be driven by common factors, in particular, expectations about permanent income growth or economic prospects. Another issue of concern is reverse causality: higher consumption may increase local employment and thus lead to higher home values (Aladangady, 2017). Therefore, running the simple OLS models is likely to result in a biased estimate of the housing wealth effects. Indeed, the coefficients for renters are positive and significant across models, confirming this concern.[62] Overall, the results suggest that while common factors may play a role across specifications, there seems to be a causal effect of housing wealth on consumption given that the wealth effects are strong for owners and relatively weak for renters.

Table 3.4. Housing wealth effects

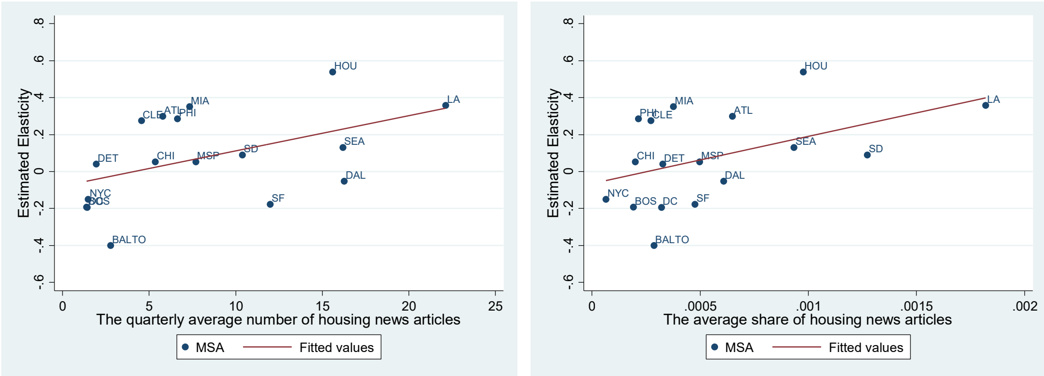
Dep. Var.: ln(Total Expenditure)	(1)	(2)	(3)
Owner × ln(HPI)	0.284***	0.267***	0.224***
	(0.0740)	(0.0546)	(0.0695)
Renter × ln(HPI)	0.166*	0.165**	0.113*
	(0.0967)	(0.0639)	(0.0621)
ln(Income)		0.247***	0.167***
		(0.00721)	(0.00684)
Age		-0.00195***	0.000304
		(0.000428)	(0.000385)
ln(Family Size)		0.244***	0.258***
		(0.00961)	(0.0102)
Fixed Effects			
Year-Quarter	YES	YES	YES
MSA × Housing Tenure	YES	YES	
Cohort			YES
Observations	110,622	96,776	96,776
Adjusted R-squared	0.225	0.477	0.617

Notes: This table presents regression results based on Equation 3.2. The dependent variable is the log of quarterly non-housing total expenditures of each individual household, and variables of interest are the logarithms of the Federal Housing Finance Agency (FHFA) MSA-level quarterly house price index (HPI) interacted with either a homeowner or a renter dummy. As the dataset is not a balanced panel but repeated cross-section survey data, Model 3 includes the cohort fixed effects defined by City × Housing tenure × Employer × Occupation × Marital status × Education × Race. By doing so, Model 3 can exploit within-cohort temporal variation; the other specifications control for MSA-housing tenure instead of the cohort FEs. Additionally, all models include year–quarter fixed effects. Robust standard errors in parentheses are clustered by MSA-housing tenure. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

[62] To address the endogeneity concern, recent empirical work uses the interaction between an exogenous demand shock and housing supply elasticity measured by Saiz (2010) as an instrument for house price growth, and popular demand shifters are national average house prices (Dettling and Kearney, 2014; Chetty and Szeidl, 2017) and long-term interest rates (Chaney et al., 2012; Aladangady, 2017).

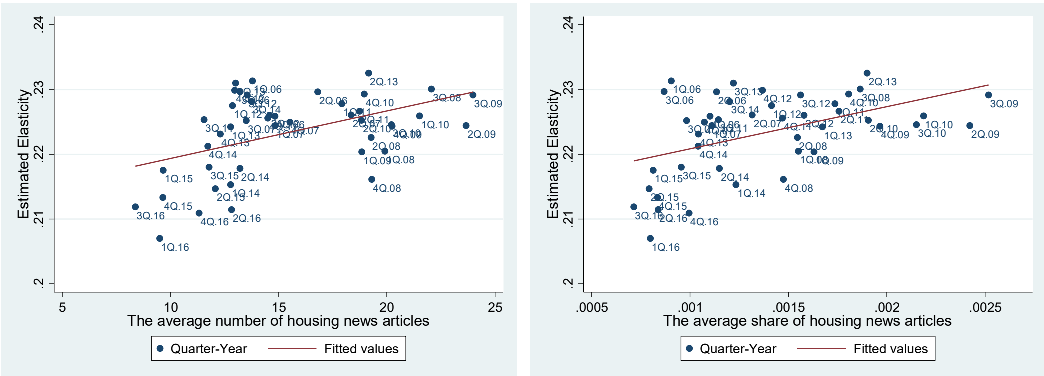
Figure 3.1. Correlations between wealth effect elasticities and housing news volumes by MSA and by quarter-year

Panel A: By MSA



Notes: Each data point reflects estimated elasticity and a housing news quantity measure for each MSA. The MSA-level elasticity estimates are based on Model 3 of Table 3.4. By interacting the housing wealth measure with a group of MSA dummies in Equation 3.2, the estimated parameter on each interaction term can be interpreted as a MSA-specific consumption elasticity with respect to the local housing price index. *The quarterly average number of housing news articles* (Left) and *The average share of housing news articles* (Right) reflect averages over the period from Q1.2006 to Q4.2016. The share is a ratio of the number of housing news articles to the number of all news articles published in the corresponding city.

Panel B: By quarter-year



Notes: All data points reflect elasticities and housing news quantity measures for each quarter-year period. To calculate period-specific elasticities, I interact the housing price variable with a group of quarter-year dummies in Equation 3.2. *The average number of housing news articles* (Left) and *The average share of housing news articles* (Right) are averages across MSAs during the corresponding period.

Figure 3.1 visually presents the relationship between the housing wealth effect estimates and the media coverage measure. First, I regress the household expenditures on the housing wealth interacted with a group of MSA dummies. By doing so, I can estimate housing wealth elasticity over time for each MSA. Then the estimated elasticity is plotted against the average quarterly number of housing price news articles for the corresponding city (the left graph) or against the average ratio of housing news articles to the total news articles for each city (the right graph) in Panel A. Instead of absolute quantity measures, I also use shares in the right graph because some major newspapers are missing in my sample, and thus differences in absolute quantities may not be comparable across cities. Likewise, Panel B plots the estimated quarterly elasticity against each quarter's average number of housing reports and its share of housing reports in the left and right graphs, respectively.

The elasticity appears to be correlated with the housing news volumes both across cities and over time. For example, although New York City and Los Angeles are both major American metropolitan cities, New York City sees much less housing news and lower elasticity, while Los Angeles sees more housing news and higher elasticity (Panel A). The housing wealth effects are also larger when more housing reports are published (Panel B). For example, the effect was relatively large during the housing market bust period from 2008 to 2010, when housing markets received substantial media attention. However, the elasticity estimates declined in the recent years 2014–2016, with less housing news published.

3.3.2 Baseline Specification

If consumption responses to housing wealth shocks vary systematically with the number of housing news articles, **Equation (3.3)** can provide a valid identification by characterizing the link graphically displayed in **Figure 3.1**:

$$\text{(3.3)}\quad y_{i,c,t} = \beta_1\, \rho_{c,t-1}\, z_{c,t} + \beta_2\, z_{c,t} + \beta_3 \rho_{c,t-1} + \boldsymbol{x'}_{i,c,t}\, \boldsymbol{\gamma} + \lambda_t + \phi_{c,j} + \varepsilon_{i,c,t}$$

where $\rho_{c,t-1}$ is the house price index for the previous term $(t-1)$ and is interacted with $z_{c,t}$, the variable of interest in this study, which is the standardized number of housing news articles in city c in year-quarter t.[63] In addition to the same covariates and fixed effects that are used in **Equation (3.1)**, this specification also controls for a square term of housing prices ($\rho^2_{c,t}$), which could capture the correlation between the housing prices and the number of housing news. Unless publication of housing news (z) is determined by common factors such as expected future income growth,[64] the coefficients for housing prices and squared housing prices would capture the impacts of the omitted variable, thereby recovering an independent effect of media reporting on the elasticity (β_1).

Just like in the previous subsection, this study interacts the three key independent variables with homeowner and renter dummies as in **Equation (3.4)**. By doing so, I can disentangle the effects on owners and renters:

$$\begin{aligned}\text{(3.4)}\quad y_{i,c,t} = {} & Owner_i\, [\, \beta_1 \rho_{c,t-1}\, z_{c,t} + \beta_2\, z_{c,t} + \beta_3\, \rho_{c,t-1}] \\ & + Renter_i\, [\, \beta_4 \rho_{c,t-1}\, z_{c,t} + \beta_5\, z_{c,t} + \beta_6 \rho_{c,t-1}] \\ & + \boldsymbol{x'}_{i,c,t}\, \boldsymbol{\gamma} + \lambda_t + \phi_{c,j} + \varepsilon_{i,c,t}\end{aligned}$$

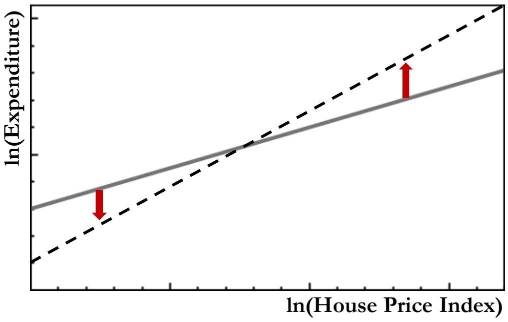
[63] This study standardized the key variable mainly for easier interpretation. Using the logarithm of the number of articles does not make any noticeable differences.

[64] This is not very likely, given the low correlation between housing prices and the news volume.

This baseline specification is motivated by **Figure 3.2**. For simplicity, suppose that the log consumption is linear to the log housing wealth. Then dependence of the housing wealth effect on newspaper reports could be visually described as in the figure. The dashed line represents the optimal consumption set of the fully informed homeowners, characterizing an information environment with more housing news, and the solid line represents less informed homeowners in an environment with less housing news. The fully informed utility-maximizing homeowners should consume more than the less-informed in response to high housing prices. But the fully-informed should consume less than the less-informed when housing prices are very low. In other words, more news about housing prices has a positive impact on spending when house prices are high, but a negative impact when prices are low. Therefore, the wealth-spending elasticity should be a function of the volume of housing news, and the function $(\beta_1 z_{c,\,t} + \beta_3)$ can be obtained by simply reordering terms (**Equation 3.5**).[65]

Figure 3.2. Expected effect of increasing newspaper reports on the link between house price index and spending

When the number of housing reports increase, the relationship between housing wealth and consumption is expected to shift from the solid line to the dash line. The slope of the equation should go up with the y-intercept decreasing.



$$\textbf{(3.5)}\quad y_{i,c,t} = \alpha + \beta_1 \rho_{c,t-1}\, z_{c,t} + \beta_2 z_{c,t} + \beta_3\, \rho_{c,t-1} + \varepsilon_{i,c,t}$$
$$= \left[\alpha + \beta_2\, z_{c,t}\right] + \left[\beta_1 z_{c,t} + \beta_3\right] \rho_{c,t-1} + \varepsilon_{i,c,t}$$

Importantly, the validity of this specification depends not only on β_1, but also on β_2. In **Figure 3.2**, the slope of the equation should go up, with the y-intercept decreasing when the number of housing reports increases. Thus, β_1 and β_2, respectively, are expected to be positive and negative.

$$\textbf{(3.6)}\quad y = \alpha + \beta_1 \rho z + \beta_2 z + \beta_3\, \rho + \varepsilon \;\; \Rightarrow \partial e = \underbrace{\partial \rho\, (\beta_1 z + \beta_3)}_{Wealth\ Effect} + \underbrace{\partial z\, (\beta_1 \rho + \beta_2)}_{Media\ Effect}$$

[65] This specification assumes that the elasticity of household expenditures with respect to housing wealth is linear to the housing news volumes.

To see this more clearly, **Equation (3.6)** disentangles housing wealth effects from media effects by partial differentiation.[66] Regardless of the volume of housing price news (z), the wealth effect is widely believed to be positive, but more housing reports combined with a positive β_1 could make the consumption response more elastic. That is, an increase of one standard deviation in the number of housing price articles is associated with a β_1 increase in the elasticity. In the meantime, the media effect ($\beta_1 \rho + \beta_2$) is negative when housing prices (ρ) are low enough. Hence, β_2 is expected to be negative. For example, more media reporting about distressed housing markets further discourages household consumption. As the housing wealth increases, the media effect becomes positive, and thus β_1 should be positive.

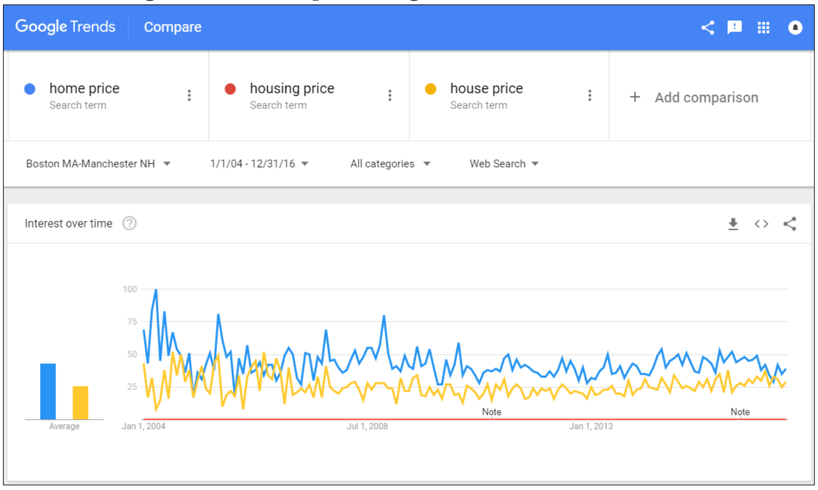
An empirical concern is unobservable common shocks driving up both housing news publications and the wealth-consumption elasticity. Even when the underlying information is fixed, a local paper's reporting decisions may be correlated to unobserved determinants of the elasticity. For example, a local media outlet is more likely to report housing price dynamics in MSAs where local residents are more interested in the housing market and thus their consumption is more responsive to housing wealth. Therefore, the number of housing news articles may simply reflect local residents' interest in or prior knowledge about housing prices, so it may be endogenous. Basically, this paper can substantially address this concern by including the cohort fixed effects that capture any time-invariant location- or group-specific heterogeneity. However, it is not impossible that the local citizens' interest in housing prices is time-varying for some reason or other. To further alleviate this concern, I include two additional control variables.

The first control variable is the Google Trends Search Volume Index (GSVI), used as a measure of internet users' attention to certain topics such as housing prices. Google Trends allows users to obtain a query index for a specific phrase. Attention to housing prices may be correlated with Google search volumes, as Google is the most commonly used online information source. Several recent works make use of the search query data to measure internet users' attention. For example, Da et al. (2011) show that an increase in the search volume index predicts an increase in stock prices during the next two weeks, and more recently, Chauvet et al. (2016) develop a mortgage default risk index using the Google data

[66] It is challenging to empirically disentangle the effect of media reporting from the effect of the underlying information reported. Thus, this breakdown simply assumes that each article may be informative in several ways, but the MSA-level housing price index is the only information related to consumption consequences.

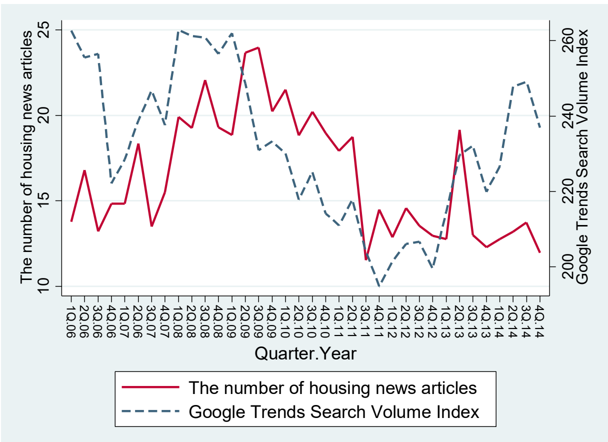
and find that their index is also predictive of housing market sentiment and performance. In this paper, I use "home price," "housing price," and "house price" as search queries on the Google Trends website in order to capture local residents' general interest in housing prices by MSA-quarter; **Figure 3.3** shows an example for Boston. I then add the indices for the three search terms to construct a quarterly panel for the MSA-level search index. **Figure 3.4** shows that local searches for housing prices on Google appear to be somewhat predictive of the number of housing news articles published during the U.S. subprime mortgage crisis period and the following Great Recession from 2007 to 2011.

Figure 3.3. An example of Google Trends Search Index for Boston



Notes: This figure shows a screen shot of the monthly Google Trends Search Indices for the search queries, "home price" (blue line), "housing price" (red line), and "house price" (yellow line) for the metropolitan Boston area. All three indices are normalized so that the period with the maximum search volume takes 100, which is assigned to March 2004 of the "home price" index. As a result, the other values of indices represent relative search volumes instead of absolute volumes. For example, the term "housing price" has relatively zero search volume over the period compared to the maximum value (March 2004, "home price").

Figure 3.4. Time-series trends of the number of housing articles and the Google search index



Notes: The red solid line denotes the quarterly time-series trend of the MSA-level average numbers of housing news articles published by all newspapers in the sample during each quarter (left axis). The blue dash line presents the average Google Trends Search Volume Index across MSAs in the sample (right axis). To calculate the Google index, I use three search queries, "home price", "housing price", and "house price", for each MSA, and add up the indices from the three search terms to construct a quarterly MSA-level index.

The other control variable is the share of construction and real estate industries in the local GDP. Unobservable industry-specific shocks may exert an influence on media reporting as well as the consumption elasticity for the following two reasons.[67] In MSAs where construction and real estate industries account for a larger fraction of the city economy, residents' consumption may be more responsive to housing prices, not only because their current or future incomes are highly correlated with the housing market performance but also because they are likely to obtain more and better house price information from their workplace or peers. At the same time, the large contribution of related industries to the local GDP can possibly affect the number of housing price reports by contributing to the advertising revenues of local newspapers. When newspapers sell more advertising to these industries, the media outlets are more likely to decide to cover topics that are relevant to their clients. To address this issue, I add the share of construction and real estate industries in the state-level GDP, using the Gross-Domestic-Product-by-Industry Data provided by the Bureau of Economic Analysis (BEA). As MSA-level data is available only on a yearly frequency, I instead use the state-level quarterly GDP shares of the industries as a control variable.

Table 3.5 presents the baseline regression results. Model 1 is a standard specification for the wealth effect, and Model 3 allows the consumption elasticity to vary according to the number of news articles conveying house price information by including the interaction term. The number of housing news articles is strongly correlated with the consumption elasticity. Specifically, an increase of one standard deviation in the number of housing price reports raises homeowners' consumption elasticity by 0.0813 (Model 3). Importantly, in the absence of the interaction term, housing prices have a significant effect on homeowner consumption while the housing news effect is insignificant (Model 2). However, adding the interaction term makes the housing wealth effect disappear, but the media effect becomes significant with a negative sign, as expected (Model 3). The findings indicate that the quantity of housing price news reports may be a main driver of the observed housing wealth effects, although interpretation of the coefficient for housing prices has strong limitations due to the omitted common shocks. A concern is that the effect of the interaction term is also

[67] A typical concern in the housing wealth literature is that the fall in housing wealth and consumption simply reflects the decline in the recession-prone construction industry (Mian et al., 2013).

sizable and significant for renters, implying that common factors may be driving up both the news publications and the consumption elasticity.[68]

Table 3.5. Results: baseline specifications

Dep. Var.: ln(Total Expenditure)	(1)	(2)	(3)	(4)	(5)	(6)
Owner × ln(HPI)	0.224***	0.220***	-0.299	-1.268	-0.970	-0.134
	(0.0695)	(0.0796)	(0.837)	(0.920)	(1.111)	(1.208)
Owner × z(#Housing News)		0.00576	-0.438***			-0.459***
		(0.00937)	(0.154)			(0.153)
Owner × ln(HPI) × z(#Housing News)			0.0813***			0.0852***
			(0.0280)			(0.0279)
Owner × ln(Google Index)				0.293		0.386
				(0.514)		(0.571)
Owner × ln(HPI) × ln(Google Index)				-0.0579		-0.0742
				(0.0964)		(0.107)
Owner × GDP Share RE-Const.					-1.823	3.980
					(8.506)	(10.64)
Owner × ln(HPI) × GDP Share RE-Const.					0.478	-0.605
					(1.551)	(1.957)
Renter × ln(HPI)	0.113*	0.0828	-0.460	-0.285	-1.115	0.549
	(0.0621)	(0.0553)	(0.865)	(1.049)	(1.085)	(1.367)
Renter × z(#Housing News)		0.0179**	-0.292*			-0.350**
		(0.00892)	(0.154)			(0.164)
Renter × ln(HPI) × z(#Housing News)			0.0564**			0.0672**
			(0.0271)			(0.0290)
Renter × ln(Google Index)				1.404**		1.276*
				(0.696)		(0.747)
Renter × ln(HPI) × ln(Google Index)				-0.265**		-0.240*
				(0.126)		(0.134)
Renter × GDP Share RE-Const.					-2.302	2.336
					(8.667)	(8.919)
Renter × ln(HPI) × GDP Share RE-Const.					0.559	-0.311
					(1.614)	(1.680)
Observations	96,776	90,052	90,052	96,199	96,776	89,475
Adjusted R-squared	0.617	0.617	0.617	0.617	0.617	0.618

Notes: This table presents regression results from Equation 3.4, the baseline specification. The dependent variable is the log of quarterly non-housing total expenditures of each individual household. ln(HPI) is the logarithm of the Federal Housing Finance Agency (FHFA)'s MSA-level house price index. Owner and Renter are indicators capturing whether the household is a homeowner or a renter. z(#Housing News) is the standardized number of housing news articles varying by MSA and quarter-year. Google Index denotes Google Trends Search Indices for three search queries, "home price", "housing price", and "house price". GDP Share RE-Const. is the share of construction and real estate industries in the state-level quarterly GDP; I use the state-level GDP, as the MSA-level GDP is available only at yearly frequency. As the dataset is repeated cross-sections, all models include the cohort fixed effects defined by City × Housing tenure × Employer × Occupation × Marital status × Education × Race in order to exploit within-cohort temporal variation for identification. All models also control for time-varying household characteristics and year-quarter fixed effects. Robust standard errors in parentheses are clustered by MSA-housing tenure. *** p<0.01, ** p<0.05, * p<0.1

Interestingly, the Google search index has a significantly negative effect on renters' elasticity and an insignificant but negative effect on that of owners. While the number of newspaper articles may be a measure of information that is passively received, the Google search index may reflect the attention of active information seekers in housing markets, who are mostly renters or potential homebuyers. By googling, they become better informed about housing prices than homeowners, and therefore the search index volumes should have a negative impact on renters' consumption responses to rising house prices, as seen in column 4. Some

[68] Despite the fact that rising house prices could have negative effects on renters because they are potential homebuyers, the estimated coefficients for renters are consistently positive.

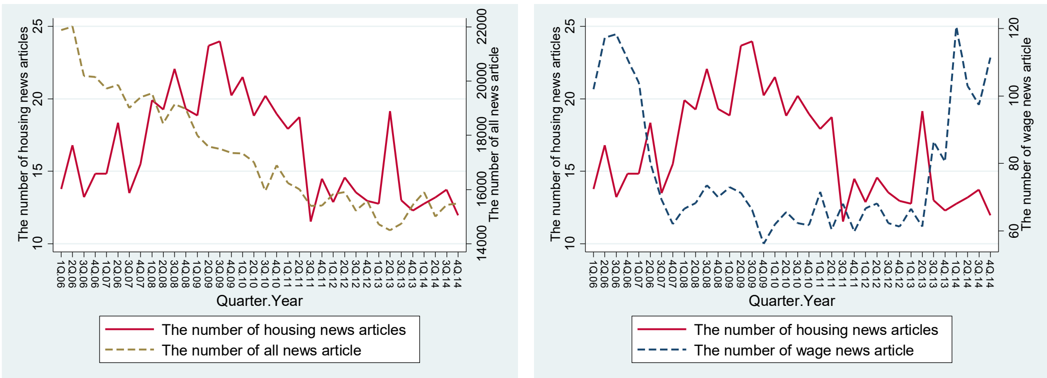
homeowners may plan to move and therefore probably actively search for housing prices on the internet. That might be the reason for the negative coefficient for homeowners.

3.3.3 Placebo Tests

The findings of **Table 3.5** may be due to random chance as opposed to a true causal effect. To relieve the concern, this paper conducts a couple of placebo tests.

The first placebo treatment is the number of wage news articles. It is plausible that housing price news contains information about incomes or wages that can affect consumption. When house prices go up, wage expectations may go up concurrently. Wage news is identified as newspaper articles that include the keyword "wage" but exclude "home," "housing," and "house". The housing terms are excluded to isolate the effect of information on wages or future income expectations from that of housing price information. If this study finds similar regression results for the wage news conveying no housing market information, then the estimated impact of housing news can be considered badly biased.

Figure 3.5. Time-series trends of the number of housing news, all news, and wage news articles



Notes: These figures show time-series relationships between the variable of interest (the number of housing news article) and two control variables (the numbers of all news and wage news articles). The left graph plots the quarterly time-series trends of the MSA-level average numbers of housing news articles (red solid line, left axis) and all news articles (yellow dash line, right axis) published by all newspapers in the sample during each quarter. In the right graph, the blue solid line denotes the MSA-level average number of wage news articles published by all the sample newspapers during each quarter. A wage news includes a keyword, "wage", but excludes all the three queries, "home", "housing", and "house", in its headline and body.

The second is the number of all news articles. The number of housing price reports may simply capture general trends in the media markets. The growth of local media markets or advertising revenues is possibly associated with household consumption, not only because more advertising sales are affected by a strong local economy but also because more advertising may boost consumption. Generally speaking, advertising revenues for U.S. newspapers have dropped since 2004, largely due to declining circulation and the growing domination of online advertising options (Chandra and Kaiser, 2015). As a result, the total number of newspaper reports have also declined, as shown in **Figure 3.5**.

Table 3.6. Results: placebo tests

Dep. Var.: ln(Total Expenditure)	(1)	(2)	(3)	(4)
Owner × ln(HPI)	-0.134	-2.311*	-2.576**	-1.328
	(1.208)	(1.219)	(1.190)	(1.203)
Owner × z(#Housing News)	-0.459***			-0.368**
	(0.153)			(0.151)
Owner × ln(HPI) × z(#Housing News)	0.0852***			0.0681**
	(0.0279)			(0.0276)
Owner × z(#All News)		0.389*		0.0803
		(0.199)		(0.192)
Owner × ln(HPI) × z(#All News)		-0.0718*		-0.0123
		(0.0381)		(0.0365)
Owner × z(#Wage News)			0.378***	0.325**
			(0.144)	(0.157)
Owner × ln(HPI) × z(#Wage News)			-0.0704***	-0.0609**
			(0.0269)	(0.0293)
Renter × ln(HPI)	0.549	-1.638	-1.765	-0.338
	(1.367)	(1.183)	(1.251)	(1.478)
Renter × z(#Housing News)	-0.350**			-0.269
	(0.164)			(0.216)
Renter × ln(HPI) × z(#Housing News)	0.0672**			0.0524
	(0.0290)			(0.0385)
Renter × z(#All News)		0.179		0.0978
		(0.142)		(0.191)
Renter × ln(HPI) × z(#All News)		-0.0388		-0.0227
		(0.0271)		(0.0351)
Renter × z(#Wage News)			0.132	-0.0136
			(0.196)	(0.216)
Renter × ln(HPI) × z(#Wage News)			-0.0253	0.00238
			(0.0368)	(0.0402)
Observations	89,475	91,734	91,734	88,572
Adjusted R-squared	0.618	0.619	0.619	0.618

Notes: The dependent variable is the log of quarterly non-housing total expenditures of each individual household. *ln(HPI)* stands for the log of the FHFA's MSA-level house price index. *Owner* and *Renter* are indicators capturing whether the household is a homeowner or a renter. *z(#Housing News)* is the standardized number of housing news articles varying by MSA and quarter-year. *z(#All News)* and *z(#Wage News)* are the standardized numbers of all news and wage news articles, respectively. As the dataset is repeated cross-sections, all models include the cohort fixed effects. All models also control for time-varying household characteristics, the *Google Trends Search Index,* the share of construction and real estate industries in the state-level quarterly GDP, and year-quarter fixed effects. Robust standard errors in parentheses are clustered by MSA-housing tenure. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

However, both all news and wage news articles present negative impacts (Models 2 and 3 of **Table 3.6**). A plausible reason for the negative effect of all news is that more information simply means more divided attention to news about a specific topic such as housing prices, considering that people have only limited amounts of time and cognitive resources to process information. As a consequence, household consumption may become less responsive to housing wealth. Also, it may be a spurious correlation. In recent years, housing prices have recovered, whereas the number of newspaper articles has drastically declined. Yet there are no causal relationships between the two trends. The negative coefficient for wage news might be due to macroeconomic policies. Wage-related policies are strongly determined by present economic conditions. During a recession, politicians usually call for an increase in the minimum wage to stimulate consumption, but their concern is the negative impact of high labor costs during the next phases of a business cycle. Thus, such discussions in newspapers may have a countereffect on the consumption elasticity. Media coverage of policy discussions on an increase in the minimum wage could encourage households to consume more even during an economic downturn, when housing prices usually plummet. Notably, using both placebo news articles as additional control variables helps to further recover the causal effect of media coverage (Model 4), implying that the wage information may be to some extent correlated with the housing price information. As the effects of the interaction and housing price news for renters become insignificant, the endogeneity issue is considerably alleviated.

3.3.4 Headline Effect

For further causal investigations, this subsection identifies the headline effect. The variable of interest in this study is the number of newspaper articles for which the headline includes "home," "house," "housing," "property," or "real estate," and the headline or the body includes "housing price," "house price," or "home price." If an article conveying house price information captures local readers' dynamic interest in housing prices or any other omitted variables that can affect local elasticity, then the article should do so even without housing terms in its headline. **Table 3.7** presents an example. The left panel shows one of the articles included in the key variable; it has "home" in its headline and "home price" in its body. The article on the right also provides readers with similar information about housing prices, but its headline does not include any housing term. Considering the underlined phrases, "the first quarter-over-quarter improvement in three years" (left) and "its first quarterly increase in three years" (right), a common information source seems to have influenced both of these articles published on the same date by two neighboring local papers.[69] The article on the left is obviously about housing market trends, and the one on the right is more about general economic conditions. However, both of the articles appear to reflect local readers' interest in housing prices or similar supply-side factors affecting media reporting.

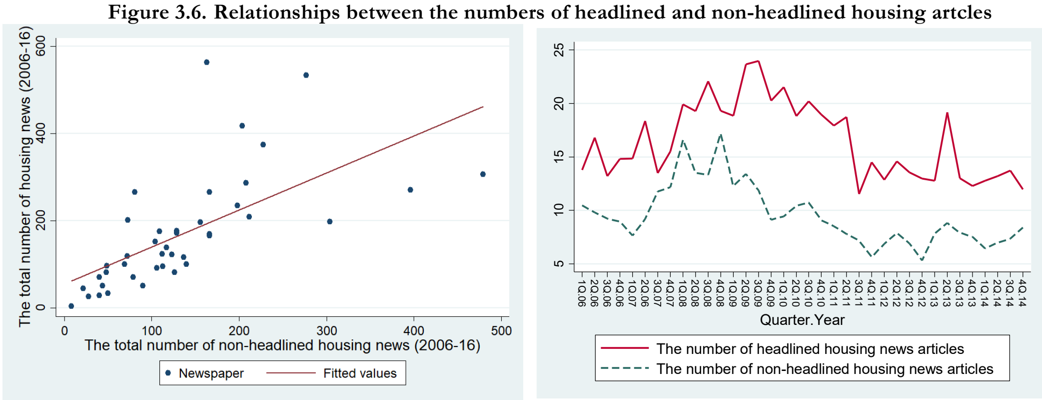
[69] It is possible that housing report volumes simply capture time-varying local residents' interests in housing markets or prices that potentially are influenced by other information channels rather than local newspapers in the sample. Therefore, coverage of other media outlets could jointly determine the newspaper reporting and the consumption responses to housing wealth.

Table 3.7. Examples: headlined and non-headlined housing news articles

	Headlined housing news article	Non-headlined housing news article
Newspaper	Chicago Sun-Times (Chicago, IL)	Daily Herald (Arlington Heights, IL)
Date	Aug 25, 2009	Aug 25, 2009
Headline	Chicago home prices rise in June	Consumer sentiment improves more than expected
Body	Home prices in the Chicago metropolitan area rose 1.1 percent in June over May, but were down 16.7 percent from a year earlier according to the latest Standard & Poor's Case-Shiller home price index. Nationally, prices rose 2.9 percent in the second quarter from the first quarter, the first quarter-over-quarter improvement in three years. But prices were down 14.9 percent from a year earlier. "For the second month in a row, we're seeing some positive signs," David Blitzer, chairman...	Consumer sentiment rose more than expected in August and expectations hit the highest level since the recession began, indications that Americans' pessimism about the economy may be lifting. The housing sector also showed signs of life as a national measure of home prices posted its first quarterly increase in three years. The New York-based Conference Board said today its Consumer Confidence index rose to 54.1 from an upwardly revised 47.4 in July. Economists surveyed...

Notes: Both the articles were published on the same date by two local newspapers within a metropolitan area. The left article is included in the key explanatory variable whereas the right is included in the control variable. In this study, a housing news article includes "home", "house", "housing", "property" or "real estate" in its headline, and also includes "housing price", "house price" or "home price" in its headline or body. In contrast, a non-housing-headlined news article includes "housing price", "house price" or "home price" in its body, but does not include any of the five terms, "home", "house", "housing", "property", and "real estate" in its headline. As examples, the left article (a housing news article) includes "home" in its headline, as well as "Home prices" in its body. The right (a non-housing-headline news article) includes "home prices" in its body without any housing terms in its headline. Given the underlined phrases, both may be influenced by an unobservable common factor, but may have different effects on readers.

Figure 3.6. Relationships between the numbers of headlined and non-headlined housing artcles



Notes: These figures show relationships between the number of headlined housing news articles and the number of non-headlined housing news articles by newspaper (left) and over time (right). Each data point in the left panel reflects the total number of housing news articles and that of non-headlined news articles published by a local newspaper from 2006 to 2016. The right graph plots the quarterly time-series trends of the MSA-level average numbers of housing (red solid line) and non-housing-headlined news articles (green dash line) published by all newspapers in the sample during each quarter.

By taking advantage of this pattern, I estimate the effect of the number of articles in which the body includes "housing price," "house price," or "home price" but whose headline does not include any of the five terms "home," "house," "housing," "property," and "real estate"; I call these "non-headlined housing news" or "non-headlined news" hereafter. The non-headlined news articles largely consist of two groups. Most of them focus on non-housing economic issues but more or less relate the main issues to housing prices. Thus, the number of such reports could capture the extent to which local readers are interested in housing markets. The other group includes reports about the housing market with a less straightforward headline. In this case, I assume that whether the headline includes any of the housing terms is random, so this group can partially capture the treatment effect but does not bias the estimated treatment effect. **Figure 3.6** displays strong correlations between the headlined and non-headlined housing news articles across newspapers and over time.

Table 3.8. Results: headline effect

Dep. Var.: ln(Total Expenditure)	(1)	(2)	(3)
Owner × ln(HPI)	-1.328	-2.439**	-1.537
	(1.203)	(1.202)	(1.212)
Owner × z(#Housing News)	-0.368**		-0.418***
	(0.151)		(0.156)
Owner × ln(HPI) × z(#Housing News)	0.0681**		0.0780***
	(0.0276)		(0.0286)
Owner × z(#Non-Headlined)		-0.104	0.0718
		(0.124)	(0.121)
Owner × ln(HPI) × z(#Non-Headlined)		0.0173	-0.0152
		(0.0229)	(0.0221)
Renter × ln(HPI)	-0.338	-1.546	-0.579
	(1.478)	(1.491)	(1.503)
Renter × z(#Housing News)	-0.269		-0.286
	(0.216)		(0.269)
Renter × ln(HPI) × z(#Housing News)	0.0524		0.0558
	(0.0385)		(0.0479)
Renter × z(#Non-Headlined)		-0.0584	0.0449
		(0.189)	(0.215)
Renter × ln(HPI) × z(#Non-Headlined)		0.0107	-0.00903
		(0.0345)	(0.0392)
Observations	88,572	88,572	88,572
Adjusted R-squared	0.618	0.618	0.618

Notes: The dependent variable is the log of quarterly non-housing total expenditures of each individual household. *ln(HPI)* stands for the log of the FHFA's MSA-level house price index. *Owner* and *Renter* are indicators capturing whether the household is a homeowner or a renter. *z(#Housing News)* and *z(#Non-Headlined)* are the standardized numbers of respectively headlined and non-headlined housing news articles varying by MSA and quarter-year. As the dataset is repeated cross-sections, all models include the cohort fixed effects. All models also control for time-varying household characteristics, the *Google Trends Search Index,* the share of construction and real estate industries in the state-level quarterly GDP, the standardized numbers of all news and wage news articles, and year-quarter fixed effects. Robust standard errors in parentheses are clustered by MSA-housing tenure.
*** p<0.01, ** p<0.05, * p<0.1

However, only headlined housing news has a statistically significant effect on the wealth effect elasticity (**Table 3.8**). If the number of housing news is endogenous, then the non-headlined news also should have a significant effect, but it does not.[70] Thus this result substantially alleviates the endogeneity issue. Controlling for the number of non-headlined news articles, the treatment effect becomes statistically more significant for homeowners (Model 3). Due to the non-headlined news, the treatment effect may largely capture the effect of headlining, and this result is consistent with a recent study's finding that about 60% of news articles are shared on Twitter without even being read (Gabielkov et al., 2016).

3.3.5 Alternative Specification

There is no doubt that media reporting, housing prices, and consumer spending are all strongly influenced by macroeconomic fluctuations. Given what housing and financial markets have gone through during the recent decade, unobservable macroeconomic factors

[70] I assume that it is random whether or not each housing news article includes housing terms in its headline.

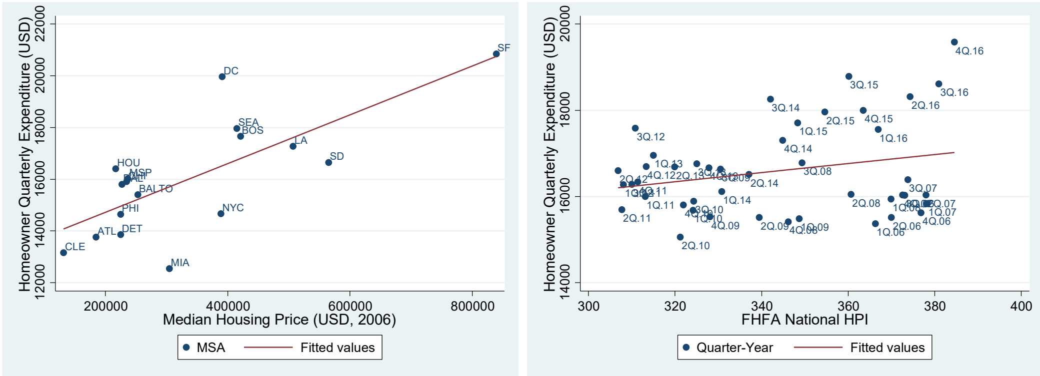
are likely to have exerted a large influence on both housing price news and household consumption elasticity. After the U.S. subprime mortgage crisis of 2007–2009, the following Great Recession saw a collapse in housing prices and household consumption. Such economic events strongly impact expectations of permanent income, a common shock affecting both housing prices and spending, by drawing substantial media attention to housing markets. Thus, time-varying unobservable macro factors may have a great potential to bias empirical results by driving up both the wealth-consumption elasticity and media reporting over the boom and bust. **Figure 3.7** presents relationships among the three key variables in this study. The relations between housing prices and consumption are consistently positive, both across cities and over time (Panel A). However, the link between housing prices and the volume of housing price news appears to be positive across cities but negative over time (Panel B). Media outlets tend to publish more articles conveying housing price information in the MSAs where house prices are higher, but they seem to provide readers with more housing price news when housing prices are lower over time. At the least, this conflicting pattern confirms that housing prices and housing price news reporting are not systematically correlated. More importantly, the visual depiction may highlight the strong influence of time-series macro factors; because it is widely believed that time-series correlations are more vulnerable to omitted variables. Homeowners may simply consume more during housing booms and less during busts, with both extreme cases endogenously seeing more media news about housing prices than usual.

To address this issue arising from time-varying unobservable factors, this section exploits only cross-sectional variation in media reporting by averaging the variable of interest over the entire period from 2006 to 2016 for each MSA. In principle, the panel estimation employed in the previous section enables a more powerful test by taking full advantage of variation across cities and over time. However, the estimates could be biased because the media decision to report may heavily depend on the macroeconomic fundamentals. Hence, I identify the media effects from the interaction of local housing price swings and presumably less correlated time-invariant local housing price news volumes. Exploiting only cross-sectional variation in housing news lends additional credibility that the estimates are less biased by common factors, as city-level heterogeneity in housing price news volumes might be less correlated with the information events being reported over time and therefore, more exogenous to time-series fluctuations of the elasticity.

Figure 3.7. Correlations among median housing prices, consumption, and housing news volumes

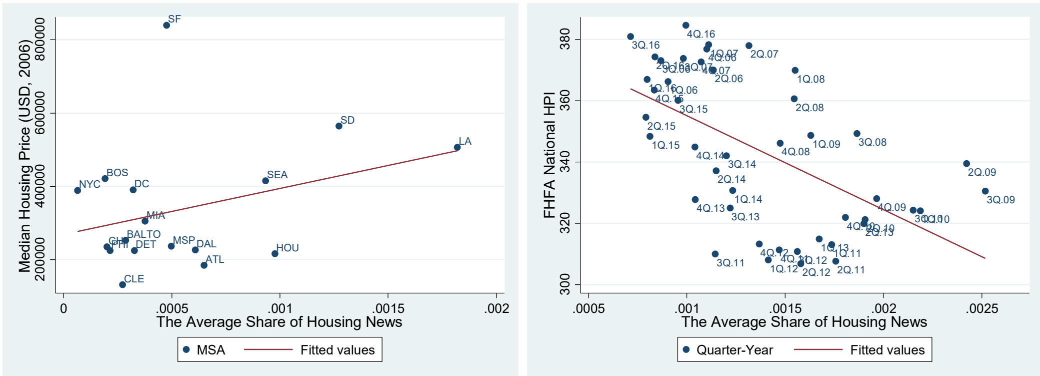
These figures show cross-sectional (left graphs) and time-series (right graphs) correlations among housing prices, household consumption, and housing news volumes, the three key variables in this study.

Panel A: Median housing prices (USD, 2006) – Average quarterly expenditures by MSA (Left) and by quarter (Right)



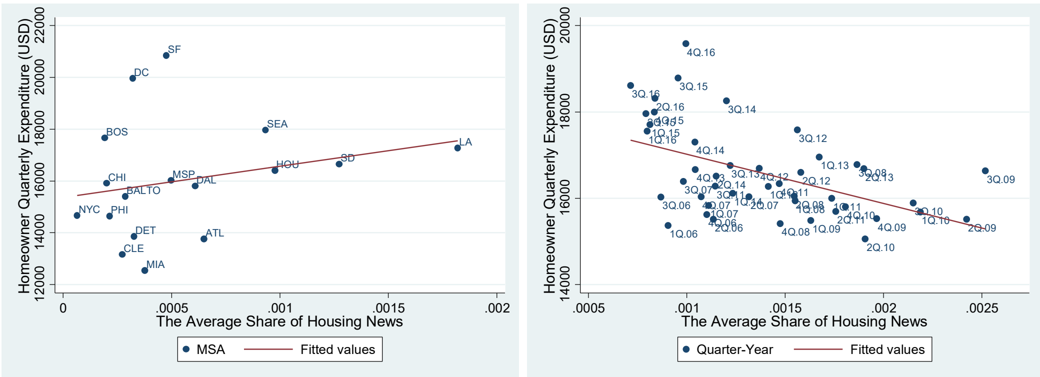
Notes: Panel A displays the relationships between median housing prices and homeowner expenditures by MSA (left) and by quarter-year (right). In the left graph, median housing prices are MSA-level from the National Association of Realtors (www.nar.realtor), and homeowners' quarterly expenditures are MSA-level averages over the period from Q1.2006 to Q4.2016. In the right, the x-axis presents the Federal Housing Finance Agency quarterly national house price index (HPI), and the y-axis presents the average homeowner expeditures across MSAs in the sample for the corresponding quarter-year.

Panel B: Median housing prices (USD, 2006) – Share of housing news by MSA (Left) and by quarter (Right)



Notes: Panel B displays the relationships between the median housing prices and the shares of housing news articles, the ratios of housing news to all news articles. The share of housing news articles are MSA-level averages over the period in the left graph, and averages across MSAs in the sample for the corresponding quarter-year in the right graph.

Panel C: Average quarterly expenditures(USD) – Share of housing news by MSA (Left) and by quarter (Right)



Notes: Panel C displays the relationships between the homeowner expenditures and the shares of housing news articles. See Panels A and B for details on the x- and y-axises.

To do this, I employ **Equation (3.7)** in this section:

$$\textbf{(3.7)}\quad y_{i,c,t} = Owner_i\,[\beta_1\,\rho_{c,t-1}\,\%news_c + \beta_2\,\rho_{c,t-1}] + Renter_i\,[\beta_3\,\rho_{c,t-1}\,\%news_c + \beta_4\,\rho_{c,t-1}] + \boldsymbol{x'}_{i,c,t}\,\boldsymbol{\gamma} + \lambda_t + \phi_{c,j} + \varepsilon_{i,c,t}$$

where $\%news_c$ denotes the city-level percentage ratio of housing news to all news articles published during the period from 2006 to 2016.[71]

Table 3.9 robustly confirms the effects of the media coverage on housing prices. Household consumption is more responsive to volatile housing prices in cities with a larger share of housing price news. Model 1 is the baseline specification and shows a statistically significant effect only on homeowners. Models 2 and 3 employ the time-invariant percentage share of non-headlined housing news and that of wage news, respectively, as additional control variables as well as placebo news. Yet both have no significant impact. Lastly, Model 5 instruments for the long-term share of housing news with the long-term share of stock news over the same period. The stock news articles are identified through the search queries "S&P 500," "Dow Jones," and "Nasdaq," and a part of the variation in housing news reporting is explained by the variation in stock news reporting (**Figure 3.8**). It is plausible that some newspapers publish more pages in the economy and finance sections with a larger team of staff members, and these sections are likely to cover more housing price news when stock markets are calm than newspapers that are otherwise similar. If this is the case, then the share of housing price news should be strongly correlated with the share of stock news in the long term. More importantly, the stock market event information reported by newspapers rarely reflects local economic conditions, since the search keywords are not location-specific. Therefore, the share of stock news is not likely to have a direct impact on the local housing wealth elasticity. Unobservable local factors may influence a local media outlet's decisions to report stock market events, but the reporting decisions are not likely to vary systematically with households' consumption responses to local housing prices. In other words, the instrument variable is predictive of long-term housing price news volumes but uncorrelated to unobservable factors that affect the local consumption elasticities, suggesting that the exclusion restriction is not violated. With the IV, the key interaction term has a more sizable impact on homeowner spending with an insignificant but negative impact on renters. The results provide some evidence that the IV estimator further alleviates the concern regarding common factors that are expected to affect the consumption elasticities of both homeowners and renters along with housing price news. Overall, a 0.01 percentage point increase in the housing price news share raises the elasticity by 0.0447–0.0684. However, a limitation is that cross-sectional variation in media reporting comes from only

[71] The reason for using the share measure instead of the quantity measure was discussed in section 3.3.2.

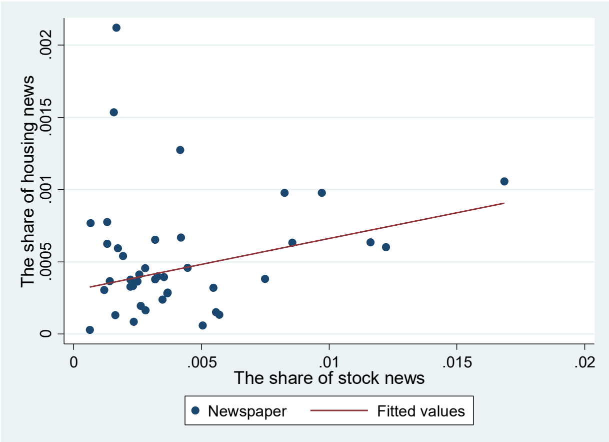
22 cities, so the first-stage F-test does not exceed the Stock and Yogo (2005) thresholds at 20 percent level, suggesting that the instrument is weak.

Table 3.9. Results: alternative specifications

Dep. Var.: ln(Total Expenditure)	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV
Owner × ln(HPI)	-0.480	-1.051	-0.938	0.206	0.425
	(0.986)	(0.996)	(0.895)	(1.119)	(1.270)
Owner × ln(HPI) × % Housing News	1.793**			4.471***	6.842***
	(0.719)			(0.935)	(2.591)
Owner × ln(HPI) × % Non-headlined News		1.485		-7.311**	-12.27**
		(2.362)		(3.430)	(6.104)
Owner × ln(HPI) × % Wage News			-0.224	-0.399	-0.460*
			(0.278)	(0.278)	(0.270)
Renter × ln(HPI)	-0.550	-1.173	-1.131	-0.0350	-0.00782
	(1.017)	(1.016)	(0.952)	(1.172)	(1.297)
Renter × ln(HPI) × % Housing News	0.637			1.266	-1.888
	(0.583)			(1.232)	(1.726)
Renter × ln(HPI) × % Non-headlined News		0.717		-1.260	6.684
		(1.616)		(3.604)	(5.383)
Renter × ln(HPI) × % Wage News			-0.141	-0.289	-0.271
			(0.207)	(0.242)	(0.280)
Observations	92,311	92,311	92,311	92,311	88,819
Adjusted R-squared	0.619	0.619	0.619	0.619	0.093

Notes: The dependent variable is the log of quarterly non-housing total expenditures of each individual household. *HPI* is the MSA-level quarterly house price index. *Owner* and *Renter* are indicators capturing whether the household is a homeowner or a renter. *%Housing News* is a MSA-level percentage ratio of housing news to all news articles published during the period from 2006 to 2016. *%Non-Headlined News* and *%Wage News* are also percentage ratios of non-housing-headlined news and wage news, respectively, to all news articles, varying only by MSA. Model 5 instruments for the share of housing news articles with the share of stock news articles that include search queries, "S&P 500", "Dow Jones", or "Nasdaq". As the dataset is repeated cross-sections, all models include the cohort fixed effects. All models also control for time-varying household characteristics and year-quarter fixed effects. Robust standard errors in parentheses are clustered by MSA-housing tenure. *** p<0.01, ** p<0.05, * p<0.1

Figure 3.8. Correlation between the shares of housing news and stock news



Notes: The fitted line excludes two outliers of which the share of housing news is greater than 0.15%, *the Daily News of Los Angeles* and *The Orange County Register*. Without the two LA media outlets, the correlation is statistically significant with p<0.01.

3.4 Conclusion

Different economic decisions may come from different information environments. In particular, many homeowners do not actively seek housing price information in their everyday lives and are therefore unaware of the precise values of their homes, given the imperfect information and a high degree of uncertainty in housing markets. To what extent, then, might publicity affect the behavior of consumers who are not actively seeking information? This paper pays particular attention to the volume and frequency of information and tests whether homeowner awareness of housing prices can increase consumption responses to house prices. By using the number of articles conveying house price information in local newspapers as a proxy for homeowner awareness of their housing wealth, I robustly find that information from news media can alter homeowner decisions, thereby increasing the elasticity of consumption with respect to housing wealth. The more frequently households are informed about house prices, the larger consumption growth we can expect in response to similar housing price appreciation.

The core contribution of this paper is twofold. First, I show that the quantity of information is predictive of agents' economic decisions. Second, understanding the relationship between information quantity/frequency and outcomes, and finding a more effective way to inform citizens can be very important to practitioners/policymakers for having more accurate predictions and making right choices. Disparity in the amount of information available to different individuals may function as a friction in the macroeconomic policy process and thereby render interventions less effective or less consistent than anticipated.

BIBLIOGRAPHY

ALADANGADY, A. (2017): "Housing Wealth and Consumption: Evidence from Geographically-Linked Microdata," *American Economic Review*, 107, 3415-46.

ALBOUY, D. (2016): "What Are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities," *The Review of Economics and Statistics*, 98, 477-487.

ALBOUY, D., AND B. LUE (2015): "Driving to Opportunity: Local Rents, Wages, Commuting, and Sub-Metropolitan Quality of Life," *Journal of Urban Economics*, 89, 74-92.

ALLCOTT, H., AND M. GENTZKOW (2017): "Social Media and Fake News in the 2016 Election," *Journal of Economic Perspectives*, 31, 211-36.

ANDERSON, M., AND J. MAGRUDER (2012): "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database," *The Economic Journal*, 122, 957-989.

ANGRIST, J. D., AND J.-S. PISCHKE (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Pres.

BAILEY, M., R. CAO, T. KUCHLER, AND J. STROEBEL (2018): "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 126, 2224-2276.

BARBER, B. M., AND D. LOEFFLER (1993): "The “Dartboard” Column: Second-Hand Information and Price Pressure," *Journal of Financial and Quantitative Analysis*, 28, 273-284.

BAUM-SNOW, N., AND F. FERREIRA (2015): "Causal Inference in Urban Economics," in *Handbook of Urban and Regional Economics Vol. 5*.

BERGER, J., A. T. SORENSEN, AND S. J. RASMUSSEN (2010): "Positive Effects of Negative Publicity: When Negative Reviews Increase Sales," *Marketing Science*, 29, 815-827.

BERTRAND, M., AND A. MORSE (2011): "Information Disclosure, Cognitive Biases, and Payday Borrowing," *Journal of Finance*, 66, 1865-1893.

BESHEARS, J., J. J. CHOI, D. LAIBSON, AND B. C. MADRIAN (2018): "Chapter 3 - Behavioral Household Finance," in *Handbook of Behavioral Economics: Applications and Foundations 1*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson: North-Holland, 177-276.

BLOMQUIST, G. C., M. C. BERGER, AND J. P. HOEHN (1988): "New Estimates of Quality of Life in Urban Areas," *The American Economic Review*, 78, 89-107.

BOSTIC, R., S. GABRIEL, AND G. PAINTER (2009): "Housing Wealth, Financial Wealth, and Consumption: New Evidence from Micro Data," *Regional Science and Urban Economics*, 39, 79-89.

BUITER, W. H. (2008): "Housing Wealth Isn't Wealth," *National Bureau of Economic Research Working Paper Series*, No. 14204.

BUSSO, M., J. GREGORY, AND P. KLINE (2013): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103, 897-947.

CALOMIRIS, C., S. D. LONGHOFER, AND W. MILES (2012): "The Housing Wealth Effect: The Crucial Roles of Demographics, Wealth Distribution and Wealth Shares," *Critical Finance Review*, 2, 49-99.

CAMPBELL, J. Y., AND J. F. COCCO (2007): "How Do House Prices Affect Consumption? Evidence from Micro Data," *Journal of Monetary Economics*, 54, 591-621.

CARROLL, C. D., M. OTSUKA, AND J. SLACALEK (2011): "How Large Are Housing and Financial Wealth Effects? A New Approach," *Journal of Money, Credit and Banking*, 43, 55-79.

CASE, K., E., J. QUIGLEY, M., AND R. SHILLER, J. (2005): "Comparing Wealth Effects: The Stock Market Versus the Housing Market," *Advances in Macroeconomics*, 5.

CASE, K. E., J. M. QUIGLEY, AND R. J. SHILLER (2012): "Wealth Effects Revisited 1975-2012," *Critical Finance Review*, 2, 101-128.

CASE, K. E., AND R. J. SHILLER (1989a): "The Behavior of Home Buyers in Boom and Post-Boom Markets," *National Bureau of Economic Research Working Paper Series*, No. 2748.

— (1989b): "The Efficiency of the Market for Single-Family Homes," *The American Economic Review*, 79, 125-137.

CHAN, W. S. (2003): "Stock Price Reaction to News and No-News: Drift and Reversal after Headlines," *Journal of Financial Economics*, 70, 223-260.

CHANDRA, A., AND U. KAISER (2015): "Chapter 9 - Newspapers and Magazines," in *Handbook of Media Economics*, ed. by S. P. Anderson, J. Waldfogel, and D. Strömberg: North-Holland, 397-444.

CHANEY, T., D. SRAER, AND D. THESMAR (2012): "The Collateral Channel: How Real Estate Shocks Affect Corporate Investment," *American Economic Review*, 102, 2381-2409.

CHAUVET, M., S. GABRIEL, AND C. LUTZ (2016): "Mortgage Default Risk: New Evidence from Internet Search Queries," *Journal of Urban Economics*, 96, 91-111.

CHEN, Y., Y. LIU, AND J. ZHANG (2012): "When Do Third-Party Product Reviews Affect Firm Value and What Can Firms Do? The Case of Media Critics and Professional Movie Reviews," *Journal of Marketing*, 76, 116-134.

CHETTY, R., AND A. SZEIDL (2017): "The Effect of Housing on Portfolio Choice," *The Journal of Finance*, 72, 1171-1212.

CHOI, J. J., D. LAIBSON, AND B. C. MADRIAN (2010): "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds," *The Review of Financial Studies*, 23, 1405-1432.

CORRADIN, S., AND A. POPOV (2015): "House Prices, Home Equity Borrowing, and Entrepreneurship," *The Review of Financial Studies*, 28, 2399-2428.

DA, Z., U. G. GURUN, AND M. WARACHKA (2014): "Frog in the Pan: Continuous Information and Momentum," *The Review of Financial Studies*, 27, 2171-2218.

DA, Z. H. I., J. ENGELBERG, AND P. GAO (2011): "In Search of Attention," *The Journal of Finance*, 66, 1461-1499.

DAVIS, D. R., J. I. DINGEL, J. MONRAS, AND E. MORALES (Forthcoming): "How Segregated Is Urban Consumption? Evidence from Yelp," *Journal of Political Economy*.

DEATON, A. (1985): "Panel Data from Time Series of Cross-Sections," *Journal of Econometrics*, 30, 109-126.

DELLAVIGNA, S., AND J. M. POLLET (2009): "Investor Inattention and Friday Earnings Announcements," *The Journal of Finance*, 64, 709-749.

DETTLING, L. J., AND M. S. KEARNEY (2014): "House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to Have a Baby," *Journal of Public Economics*, 110, 82-100.

DRANOVE, D., AND G. Z. JIN (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48, 935-63.

ELIASHBERG, J., AND S. M. SHUGAN (1997): "Film Critics: Influencers or Predictors?," *Journal of Marketing*, 61, 68-78.

ENGELBERG, J. E., AND C. A. PARSONS (2011): "The Causal Impact of Media in Financial Markets," *The Journal of Finance*, 66, 67-97.

FANG, L., AND J. PERESS (2009): "Media Coverage and the Cross-Section of Stock Returns," *The Journal of Finance*, 64, 2023-2052.

FIGLIO, D. N., AND M. E. LUCAS (2004): "What's in a Grade? School Report Cards and the Housing Market," *The American Economic Review*, 94, 591-604.

FLAVIN, M., AND T. YAMASHITA (2002): "Owner-Occupied Housing and the Composition of the Household Portfolio," *The American Economic Review*, 92, 345-362.

FLORES, R. J. O., AND A. P. LOBO (2013): "The Reassertion of Black/Non-Black Color Line: The Rise in Integrated Neighborhoods without Blacks in New York City, 1970–2010," *Journal of Urban Affairs*, 35, 255-282.

FRAZZINI, A. (2006): "The Disposition Effect and Underreaction to News," *The Journal of Finance*, 61, 2017-2046.

FRIBERG, R., AND E. GRÖNQVIST (2012): "Do Expert Reviews Affect the Demand for Wine?," *American Economic Journal: Applied Economics*, 4, 193-211.

GABIELKOV, M., A. RAMACHANDRAN, A. CHAINTREAU, AND A. LEGOUT (2016): "Social Clicks: What and Who Gets Read on Twitter?," *ACM SIGMETRICS Performance Evaluation Review*, 44, 179-192.

GABRIEL, S. A., J. P. MATTEY, AND W. L. WASCHER (2003): "Compensating Differentials and Evolution in the Quality-of-Life among U.S. States," *Regional Science and Urban Economics*, 33, 619-649.

GAO, P., C. LEE, AND D. MURPHY (2018): "Financing Dies in Darkness? The Impact of Newspaper Closures on Public Finance."

GENTZKOW, M. (2006): "Television and Voter Turnout," *The Quarterly Journal of Economics*, 121, 931-972.

GENTZKOW, M., AND J. M. SHAPIRO (2010): "What Drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 78, 35-71.

GENTZKOW, M., J. M. SHAPIRO, AND M. SINKINSON (2011): "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 101, 2980-3018.

GERBER, A. S., D. KARLAN, AND D. 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.

GOBILLON, L., T. MAGNAC, AND H. SELOD (2012): "Do Unemployed Workers Benefit from Enterprise Zones? The French Experience," *Journal of Public Economics*, 96, 881-892.

GREENWOOD, M. J., G. L. HUNT, D. S. RICKMAN, AND G. I. TREYZ (1991): "Migration, Regional Equilibrium, and the Estimation of Compensating Differentials," *The American Economic Review*, 81, 1382-1390.

GUTERBOCK, T. M. (1997): "Review: Why Money Magazine's "Best Places" Keep Changing," *The Public Opinion Quarterly*, 61, 339-355.

HILGER, J., G. RAFERT, AND S. VILLAS-BOAS (2010): "Expert Opinion and the Demand for Experience Goods: An Experimental Approach in the Retail Wine Market," *Review of Economics and Statistics*, 93, 1289-1296.

HONG, H., AND J. C. STEIN (1999): "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *The Journal of Finance*, 54, 2143-2184.

IHLANFELDT, K. R., AND B. SCAFIDI (2002): "Black Self-Segregation as a Cause of Housing Segregation: Evidence from the Multi-City Study of Urban Inequality," *Journal of Urban Economics*, 51, 366-390.

KLINE, P., AND E. MORETTI (2014): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 129, 275-331.

KUANG, C. (2017): "Does Quality Matter in Local Consumption Amenities? An Empirical Investigation with Yelp," *Journal of Urban Economics*, 100, 1-18.

LOVENHEIM, M. F. (2011): "The Effect of Liquid Housing Wealth on College Enrollment," *Journal of Labor Economics*, 29, 741-771.

LOVENHEIM, M. F., AND K. J. MUMFORD (2013): "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," *Review of Economics and Statistics*, 95, 464-475.

LUCA, M. (2016): "“Reviews, Reputation, and Revenue: The Case of Yelp.Com," *Harvard Business School Working Paper*, No. 12-016.

MIAN, A., K. RAO, AND A. SUFI (2013): "Household Balance Sheets, Consumption, and the Economic Slump," *The Quarterly Journal of Economics*, 128, 1687-1726.

NELSON, P. (1970): "Information and Consumer Behavior," *Journal of Political Economy*, 78, 311-329.

— (1974): "Advertising as Information," *Journal of Political Economy*, 82, 729-754.

OLIVER, J. E., AND T. J. WOOD (2014): "Conspiracy Theories and the Paranoid Style(S) of Mass Opinion," *American Journal of Political Science*, 58, 952-966.

PERESS, J. (2014): "The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes," *The Journal of Finance*, 69, 2007-2043.

PIAZZESI, M., M. SCHNEIDER, AND S. TUZEL (2007): "Housing, Consumption and Asset Pricing," *Journal of Financial Economics*, 83, 531-569.

REINSTEIN, D. A., AND C. M. SNYDER (2005): "The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics," *The Journal of Industrial Economics*, 53, 27-51.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90, 1257-1278.

SAIZ, A. (2010): "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125, 1253-1296.

SILVERMAN, C., AND J. SINGER-VINE (2016): "Most Americans Who See Fake News Believe It, New Survey Says.," *BuzzFeed News*, https://www.buzzfeednews.com/article/craigsilverman/fake-news-survey.

SINAI, T., AND N. S. SOULELES (2005): "Owner-Occupied Housing as a Hedge against Rent Risk," *The Quarterly Journal of Economics*, 120, 763-789.

SOO, C. K. (2018): "Quantifying Sentiment with News Media across Local Housing Markets," *The Review of Financial Studies*, 31, 3689-3719.

STOCK, J. H., AND M. YOGO (2005): "Testing for Weak Instruments in Linear Iv Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. by D. W. K. Andrews, and J. H. Stock. Cambridge: Cambridge University Press, 80-108.

TETLOCK, P. C. (2007): "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, 62, 1139-1168.

— (2010): "Does Public Financial News Resolve Asymmetric Information?," *The Review of Financial Studies*, 23, 3520-3557.

— (2011): "All the News That's Fit to Reprint: Do Investors React to Stale Information?," *The Review of Financial Studies*, 24, 1481-1512.

TETLOCK, P. C., M. SAAR-TSECHANSKY, AND S. MACSKASSY (2008): "More Than Words: Quantifying Language to Measure Firms' Fundamentals," *The Journal of Finance*, 63, 1437-1467.

TIEBOUT, C. M. (1956): "A Pure Theory of Local Expenditures," *Journal of political economy*, 64, 416-424.

YAMASHITA, T. (2007): "House Price Appreciation, Liquidity Constraints, and Second Mortgages," *Journal of Urban Economics*, 62, 424-440.