The Price Effect of Retail Gentrification*

Louise Guillouët

April 27, 2024

Abstract: Combining barcode-level scanner data with a newly collected dataset on the opening dates of new stores by Whole Foods, a high-end grocery chain, I show that Whole Foods' entry causes prices of products purchased by lower-income households at incumbent neighborhood stores to rise by three percent. By contrast, prices of products purchased by richer households do not change. This differential effect can be explained by the fact that these two groups purchase different products at different stores. While the latter benefit from the pro-competitive effect of entry on prices, the former suffer from the readjustment of surviving incumbent stores.

JEL: D12; D63; L13; L81

Keywords: gentrification, imperfect competition, consumer goods

*guillouet.louise@gmail.com. I thank Blaine Helleloid and Helena Bauer-Mitterlahner for excellent research assistance. I wish to thank John Asker, Gaia Dossi, Amit Khandelwal, Tobias Salz, Wolfram Schlenker, David Weinstein and seminar participants at Columbia University, the University of Warwick, Sciences Po, the French Association for Environmental and Resource Economics and the European Central Bank for helpful feedback. Researcher's own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

There is no universally accepted definition of gentrification.[1] The arrival of new retailers, which serve the new residents as reflected in product assortment, hours and prices (Zukin et al. (2009), Glaeser et al. (2018), Behrens et al. (2023)) is sometimes the signal everybody can agree on. A quintessential example may be Whole Foods' arrival on the corner of 125th street and Malcolm X Boulevard in Harlem, New York City, in 2017 (New York Magazine (2017)). This type of "retail" gentrification has been decried by groups advocating for the low-income households living in these areas, who are already being displaced because of rising costs. Traditional economic intuition, however, suggests that the entry of new businesses is good for consumer welfare.

While the literature has studied the distributional effects of housing upgrading (Guerrieri et al. (2013)), little attention has been given to the distributional effects of the increasing supply in quantity and variety of retail stores. Yet, this is important because low-income households spend a much larger share of their income on retail expenditure than other households[2], so they may be more affected by changes in grocery prices than high-income households would be. Understanding how the business side of gentrification affects people across the income distribution is therefore of prime importance in order to inform policies that seek to address gentrification.

This paper documents one example where the entry of new firms into neighborhoods may benefit high-income households, yet hurt low-income households, further deepening economic inequality. Using an event-study design, I find that the entry of a high-end retailer in a neighborhood does not lead to a pro-competitive price effect overall, and causes the price of products purchased by lower-income households to increase.

In particular, I study the entry of a store from a large, high-end grocery stores chain in the United States: Whole Foods. Started in the 1980s in Texas, this chain focuses on self-defined "quality" foods, excluding a list of artificial ingredients and favoring "natural", environmentally friendly choices. Its motto is "America's healthiest grocery store".[3] The products it carries are typically more expensive than in other grocery stores, that may not meet the quality standards decided by Whole Foods. I focus on a period of rapid expansion: Whole Foods expanded from 106 stores in 103 zipcodes in 2004 to 453

[1] According to the Merriam-Webster dictionary, gentrification is "a process in which a poor area (as of a city) experiences an influx of middle-class or wealthy people who renovate and rebuild homes and businesses and which often results in an increase in property values and the displacement of earlier, usually poorer residents".

[2] In my data I only see which income "bin" households belong to. Taking the higher end of the bin as their income (a lower bound to compute expenditure shares for low-income households), it is clear that the higher a households' income, the lower is the share of Nielsen goods expenditure on income. On average households with income below \$25,000 spend north of 20% of their income on Nielsen goods while households with income above \$25,000 spend less than 10% of their income on Nielsen goods.

[3] See Whole Foods' website for details.

stores in 434 zipcodes in 2017.[4] I study the price responses of incumbent stores with an event-study across treated zipcodes.

The central identification challenge is that entry is endogenous since Whole Foods does not open stores at random, but in areas which are wealthier than average. Following Atkin et al. (2018), who study the expansion of Walmart across Mexico between 2002 and 2014, I exploit the rapid expansion of Whole Foods across the United States between 2004 and 2017. I construct a new dataset based on the list of Whole Foods stores ever in operation in the United States, each of which I link to their opening dates based on newspaper data. I then estimate an event-study specification that compares the prices of the same barcodes purchased at the same (incumbent) stores by households living in a zipcode where a Whole Foods opened during the study period, before Whole Foods opened and after it opened. Conditional on barcode-by-store and time fixed effects, there are no pre-trends in prices before entry, which means that the effect of entry on prices is identified.

I find that when Whole Foods opens a new store in a zipcode, prices paid by poorer households for the same barcodes in the same stores as before, increase by three per cent on average three years after entry. By contrast, prices paid by richer households, for the same barcodes in the same stores as before, which may be a different set than the ones purchased by poorer households, do not appear to increase. Overall, I find that the entry of Whole Foods in a new zipcode does not cause prices to decrease, the pro-competitive effect one would expect from entry in general, with contrasted effects on low-income and high-income households.

I argue this rarely documented "price-increasing entry" event can be explained by the price-sensitivities of consumers in a context of product differentiation. The intuition is that Whole Foods sells products that are not so different from the ones that high-income households purchased at incumbent stores before Whole Foods entered. In order to retain their consumers, incumbent stores must therefore make an effort to lower prices on these goods (the "market-share" effect following the vocabulary of Chen and Riordan (2008)). By contrast, Whole Foods may not sell products that look like the products that low-income households are used to buying in incumbent stores. Some of these households might decide to switch to Whole Foods, but those who stay reveal themselves to strongly prefer the incumbent's products, which is then able to raise its prices (the "price-sensitivity" effect in Chen and Riordan (2008)). I document this mechanism by showing that there is a substantial difference in prices between the chains favored by low-income households and the chains favored by high-income households, when controlling for narrowly defined product categories.

Last, this narrative is supported by evidence of exit of at least some stores in the

[4] I stop the study in 2017 because this is the year when Whole Foods was purchased by Amazon, a sign of its success and the start of a new era for the chain. See New York Times (2017)

zipcodes where Whole Foods enters. Using the US Census's Zipcode Business Patterns, I show that the number of grocery store establishments increased by 0.3 three years after the entry of Whole Foods. It suggests that maintaining profits after Whole Foods' entry is difficult, with some stores exiting. The stores that remain may have been forced to raise prices to survive, and while they are unable to do it on products favoured by high-income households for market-share reasons, they do it on products favoured by low-income households. Moreover, if some of the stores that exit used to cater to lower income households, this makes it even easier for firms that stay on to raise their prices, as the market share effect of entry is compensated by the market share effect of exit.

Price-increasing entries have been documented in several markets. Using data on pharmaceutical drugs, Perloff et al. (1995) show that entry leads to higher prices but also higher consumer welfare in a context of horizontal differentiation, because more consumers find a product that perfectly matches their taste. Using recent data on smartphone applications, Ershov (2018) shows that the entry of extremely popular firms ("superstar") in a market leads to higher entry and lower prices, but also lower quality. Closest to this paper, Ward et al. (2002) shows that when supermarkets introduce private-label goods, the price of branded goods tends to rise. This paper exploits a different angle by focusing on store entry instead of product entry in the supermarket sector which is important for consumer welfare; with no evidence of a change in quality but differential price trends that may reinforce inequality.

One classical explanation for price-increasing entry is the existence of search costs, which could be applied to smartphone applications, potentially to within-store grocery shopping, but not stores themselves as is studied in this paper. Instead, this paper relies on the model developed by Chen and Riordan (2008), who show that in a symmetric two-product, two-firm market, the duopoly price is higher than the monopoly price when the joint distribution of buyer values for the two products are negatively correlated. Deck and Gu (2012) provide an experimental test of this condition. This paper provides a real-life example of the phenomenon.

Disruptions in the retail sector have been studied in several countries outside the United States, including their implications on upstream industries in Romania (Javorcik and Li (2013)) and Mexico (Iacovone et al. (2015)) for example. The paper closest to our setting is Atkin et al. (2018) who show that the opening of a new Walmart outlet (a relatively high-end store) in a Mexican municipality led to lower prices and higher welfare overall, though with a larger benefit for higher-income consumers. The present paper suggests that in fact the entry of a high-end store may end up hurting lower-income households.

I structure the remainder of the paper as follows: Section 2 describes the data used and the empirical strategy. Section 3 presents the results. Section 4 discusses supporting evidence for the mechanism proposed as well as some evidence against alternative mechanisms. Section 5 concludes.

2 Methods

2.1 Nielsen Data

The analysis is based on Nielsen Homescan, a panel dataset from 2004 (the year the dataset begins) to 2017 (the year Whole Foods was purchased by Amazon). 40,000 to 60,000 households participate each year in the panel. I observe some demographic information about the households, including the five-digit zipcode where they live and an income bracket. Participants are incentivized to report any consumer packaged good purchased from any outlet.[5] Participants use an in-home scanner to scan the barcodes, allowing me to observe barcode-level information about each product.

Nielsen also runs another program, the Retail Measurement Services (RMS). In this data, stores' locations are only identified up to the county level. For the purpose of the event-study described below, it was not enough to observe the county of stores. By contrast, panelists' locations in the Homescan data are identified at the five-digit zipcode level, which is why it is the data I chose to use for this study.[6] Generally, another advantage of this dataset is that consumers report purchases from any outlet, not just the ones Nielsen has an RMS agreement with. However, stores are anonymized so I do not know whether the purchase I observe are at a Whole Foods or not. In my main specification, I therefore focus on the purchases made at stores for which I have an (anonymized) store identifier, which belong to chains participating in the RMS program, and I know these not to be Whole Foods (Einav et al. (2010)). In robustness checks, I include all purchases and build a store identifier based on the retail chain and the zipcode of the panelist. These may include Whole Foods.

2.2 Whole Foods Stores Locations and Opening Dates

I obtained the list of establishments selling "organic foods and services" (SIC 5499-35) on ReferenceUSA. In 2017 there were 2,145 such businesses in the US, 453 of which were Whole Foods Markets' stores. I cross-checked this list with the current list of stores given on the corporation's website (some of them had closed between 2017 and when writing this article), where I also collected the exact locations of each store.

[5] A subset of households also report nonpackaged groceries such as bulk produce, but I do not use this data. This is mostly produce. Allcott et al. (2019) show that packaged goods represent at least 60% of households' caloric intake from produce. The share is probably even higher once including non-food categories.

[6] The average county has an area of about 1000 square miles while the average zipcode has an area of 90 miles. The median shopper travels 3 miles for groceries.

I then collected the opening dates and announcement dates of new Whole Foods Market stores using ProQuest' news database, and occasionally Factiva. The process was as follows: for each store present in a given city, I entered in the ProQuest search engine the following words "Whole Foods opening + city". I usually found the report in a local newspapers of the event hosted at the store to celebrate the opening day, which gave me the date. I then looked for the earliest piece of news mentioning the upcoming entry of Whole Foods in that city among the first 20 news pieces coming up in the search "Whole Foods opening + city + year" where year is the year of opening. This news piece was either local newspapers again or comments on Whole Foods' quarterly earning reports filed by the SEC. When I went past the first 20 results of ProQuest and Factiva for the first step, I stopped and then collected the opening date by looking at the date of the earliest entry for the store on Yelp, a consumer review website. For these stores I did not complete the second step, as usually these are stores that opened before the start of my data.

2.3 Empirical Strategy

Whole Foods is an interesting chain to study for its positioning in terms of health and environmental quality. It also expanded very fast between 2004 and 2017, as the company was pursuing a growth strategy which relied mostly on sales from new stores.[7] Part of this expansion took place through acquisitions of existing stores operating under different chains, but I do not include them in my analysis.

As a profit-maximizing firm, Whole Foods did not choose the location of its new stores at random. Qualitative evidence from the annual reports shows on the contrary that the exact location of the new stores are carefully chosen, based on density, income, and education levels. Quantitative evidence obtained from the U.S. census' zipcode-level information confirms the intuition that Whole Foods opened new stores in areas that were richer, denser and have a higher level of education on average. Whole Foods therefore opens stores in zipcodes with very different characteristics from the average zipcodes in static sense. There is no reason to expect that these different areas would evolve in the same way whether Whole Foods opens or not.

I therefore focus on the markets where Whole Foods eventually opened a store over the period and perform an event study following Atkin et al. (2018). The obvious identification concern with the event study is that store openings coincided with some kind of pre trends. A common intuition is that Whole Foods opened stores in locations with increasing prices, because they are experiencing income growth for example. This is the classic gentrification argument. This would lead to an upward-biased estimate of

[7] For example, the 2014 annual report states the objective for sales growth in 2015 is 9%, while the objective for comparable stores sales growth in 2015 is less than 5%. The majority of the growth is therefore expected to come from new stores.

the treatment effect of the entry of Whole Foods on other stores' prices. Alternatively, Whole Foods might target markets where there is increasing competition as more and more stores are opening in these areas, so that prices in other stores are decreasing prior to its entry. This would lead to a downward-biased estimate. Last, it could be the case that Whole Foods has expanded as fast as possible in areas where it could potentially make profit, and did not time its entry within these areas. In this scenario, there would not be a substantial bias, as the timing would not be correlated with pre-existing time trends.

There is qualitative evidence supporting the latter hypothesis in the annual reports produced by Whole Foods for its investors. First, the company explains that their high sales growth (approximately 24% compounded annual growth rate between 1990 and 2014) is driven in large part by their opening of new stores, validating the hypothesis that there is a strong drive in the company to expand and to do so rapidly.[8] Second, the reports suggest that the timing of opening is partly exogenous. Indeed, an important part of the work done prior to opening a store is to find the best real estate opportunities, which can take a long time, the exact length of which often *surprises* the company itself. When the company discloses the risks of its activities to investors, it mentions many factors (complexity of development, weather, unions, regulation, etc.) that may delay the opening of the store, even once they have taken possession of the space.[9] Taken together, these last two observations suggest that if the *location* of new stores is not random, the *timing* may be considered quasi-random.

This qualitative evidence does not have to be taken at face value: I can verify the absence of pre-trends quantitatively, in the event-study. I estimate the following regression:

$$\ln p_{bsmt} = \sum_{j=-4}^{12} \beta_j \mathbb{I}(\text{Quarters Since Entry}_{mt} = j) + \delta_{sbm} + \eta_t + \epsilon_{sbmt}$$

where $\ln p_{bsmt}$is the log price of a product (barcode) b in individual store s, in zipcode m and time t. I() is an indicator function, and Quarters Since Entry$_{mt}$ counts the quarters since the first entry of a Whole Foods store in zipcode m at time t (with negative values counting quarters before entry, positive values counting quarters after entry and zero being the quarter the first Whole Foods store enters a zipcode). I chose the event window

[8]From the Annual Report for 2014: target growth for fiscal year is 9%, with comparable store sales growth "in the low to middle single digits" and the remainder from new stores

[9]Here is a quote from the 2008 Annual Report: "The "tender period," which we define as the length of time between a store's tender date and opening date, varies depending on several factors, some of which are outside of our control. These factors include the size of the store and complexity of site development, the impact of weather and unforeseen environmental issues, and issues surrounding construction labor unions and local government authorities, among other things. Furthermore, acquired leases, ground leases and owned properties generally have longer tender periods than standard operating leases because we take possession of these locations earlier in the construction process."

to be within one year before the event and three years after the event to match Atkin et al. (2018)'s specification. I include barcode by store fixed effects and month fixed effects. The coefficients on the "negative" quarters since the event measure the pre-trends, while the coefficients on the "positive" quarters since the event measure the effect of entry.

I define markets by combining administrative data with qualitative evidence about consumers' transportation habits. According to the 2017 National Household Travel Survey[10], the mean shopping trip distance was 7.1 miles, while the median and 75th percentile of shopping travel distance were 3 and 7 miles, respectively. Around each Whole Foods store, I therefore create a circle of radius 10 miles that corresponds to a conservative definition of the area of influence of that store. I intersect it with the zipcode areas map. I define a zipcode as valid for my study if it hosted exactly zero store before 2004 and acquired exactly one such store between 2004 and 2017, including as measured by the areas of influence. I exclude zipcodes that are covered by the 10-mile radius circle but do not host the store itself. Last, to estimate the event study on a balanced sample of areas, I exclude zipcodes where the first Whole Foods opened in the first year of my dataset (2004) or in the last 3 years of my dataset (2015-2017) and zipcodes for which prices are not observed in each month before and after the entry event. I am left with 233 zipcodes.

I cluster standard errors at the zipcode level, as it is the level where errors are most likely to be serially correlated given that the shock is at the zipcode level.

3 Results

3.1 Main results

Figure 1 shows the results of the event study estimation by plotting the coefficients and the 95% confidence intervals obtained on the dummies for the quarters before and after the opening of a new Whole Foods store. The first important conclusion is that I have enough controls to take care of any potential pre-trend: the coefficients on the quarters before the opening are precisely estimated zeros. It confirms the identification assumption that within the areas where Whole Foods chose to enter, the timing was quasi-random, controlling for month and location. Second, it does not seem like Whole Foods had any effect on prices of these goods at incumbent stores overall: the coefficients on the quarter after the opening are precisely estimated zeros. If anything, there might be evidence of a negative trend as the coefficient on twelve quarters after entry is 0.5%, but I cannot reject it is zero.

[figure 1]

I conduct two additional event-studies, keeping the specification the same but di-

[10] https://nhts.ornl.gov/

viding the sample into two groups: households below and above the median income, \$50,000.[11] Figure 2 shows the results for these two separate regressions. Looking at Panel (a), the prices of products paid by households in the lower half of the income distribution at incumbent stores, I first find that the coefficients on the dummies for the quarters before entry are precisely estimated zeros, confirming that there is no pre-trend. Second, I find that there is a positive and statistically significant effect of Whole Foods opening. Prices increase after the opening and continue to increase for three years, stabilizing just below 3% above prices paid in the year before the opening.

[figure 2]

By contrast, looking at Panel (b), the prices of products paid by households in the upper half of the income distribution, I first find that the coefficients on the dummies for the quarters before entry are precisely estimated zeros, confirming that there is no pre-trend. The coefficients on the dummies for the quarters after entry are negative, although they cannot be statistically distinguishable from zero. If anything, there appears to be a downward trend in the prices paid by these households from before entry, which supports the idea that the finding from Panel (a) is not due to an underlying trend of increasing prices.

My main result is therefore that the entry of Whole Foods into a new neighborhood, defined by a zipcode, between 2004 and 2017 caused prices paid by lower-income residents at incumbent stores to increase by three percent on average three years after entry. Because the specification is greedy in data, it is difficult to test the specification on samples that have less than half of the households. The result is qualitatively robust to considering, for example, all households earning less than \$70,000 a year which is the 75th quantile. However the point estimate is slightly lower, at two percent above baseline three years after entry. This suggests that the negative effect of entry on prices is still present, yet negative, for middle-class households.

3.2 Robustness checks

In this section, I present several alternative specifications that reinforce my main finding.

3.2.1 Store definition

As explained above, I make the conservative choice in the main specification to only keep the stores identified in the RMS data, in order to avoid including Whole Foods in the specification. As a robustness check, I therefore propose to look at the entire population

[11] In 2010 (about half-way through my study period), the median household income in the United States was \$49,445 according to the Census. Income in the Nielsen Homescan data are reported using bins, and one of these is up to \$49,999, so I use all the bins below and up to this one to define the lower-income group, and put all the other households in the higher-income group.

of stores where Homescan panelists reported making purchases. However, my regression calls for store fixed effects, and I only observe precise store identifiers for stores belonging to chains that participate in Nielsen's flagship program, the Retail Measurement Services (RMS). Instead in this robustness check, I define stores by the unique combination of a retailer code and resident zipcode. Because these zipcodes are not the zipcodes of the stores, I potentially fail to separate two stores from the same chain located in the same zipcode. But I also probably create too many identifiers when residents from different zipcodes go to the same store. Since it is rare that a chain owns several outlets in the same zipcode, the latter effect probably dominates which would attenuate the estimation.

The results can be consulted in the online appendix (Figures 4 and 5). I find a small positive effect of entry on prices of 1 to 2%, although it is not statistically significant. The main result, which is that entry leads prices of goods purchased by households in the bottom half of the income distribution to increase, still holds. All the coefficients on the dummies for quarters prior to entry are precisely estimated zeros, and almost all coefficients on the dummies for quarters posterior to entry are statistically different from zero at the 5% level, with a leveling off of the price effect at about 4% three years after entry. Last, I estimate that entry may have had a small, 1% positive effect on prices of purchases made by upper-income households in RMS stores, but this effect is not statistically different from zero.

Overall, I find higher effects using all the stores in the Homescan data than using only the RMS stores. This is consistent with the fact that RMS stores typically belong to large chains, and there is evidence that large chains practice national uniform pricing (DellaVigna and Gentzkow (2019)). Therefore, the smaller chains or independent stores included in the Homescan data increase their prices more by reaction to Whole Foods than the national chains are able to. This is especially consistent with the higher effect found for low-income households when including all stores compared to when only keeping RMS stores: Faber and Fally (2021) document the fact that lower-income households tend to source their purchases from smaller firms.

3.2.2 Controls

One may be worried that the effect I find is due to trends that are more granular than the month fixed effects that I use. Following Atkin et al. (2018), I therefore propose to re-estimate the main specification, replacing the month fixed effects by zipcode size by month fixed effects, store type by product category by month fixed effects, and region by month fixed effects. Here, zipcode sizes are just coded as the quintile in terms of the population registered in the American Community Survey in 2011. Store type is a dummy dividing stores between grocery stores, and everything else. Regions are the 9 United States "divisions" as defined by the U.S. Census and attributed by Nielsen to households based on their place of residence. The results shown in the Online Appendix (Figures 6 and 7) are remarkably similar to the one found with the baseline specification, and the estimates are more precise, which is what one would expect from adding relevant controls.

3.2.3 Falsification test

Despite the alternative specifications proposed, one may worry that the effect I am detecting is spurious. I therefore propose a data test. During the collection of the opening dates of Whole Foods stores, I also collected the earliest date on which a media source announced the upcoming opening. This was not feasible for all the "clean entry" event zipcodes, but allows for a relevant comparison. The median time between announcement date and opening date is 22 months: if the effect I estimate earlier is due to entry and not to any other event, I do not expect incumbent stores to "react" to the announcement in the media of the opening of a new store before almost two years. I therefore run again the specification shown above with the announcement date instead of the opening date. I find no detectable effect of the announcement, as shown in the Online Appendix (Figures 8 and 9). This comforts the idea that the effect I detect above is the entry effect.

4 Discussion

The main empirical result of the paper is that the entry of a Whole Foods store into a new zipcode causes the prices of goods consumed by lower-income households to increase. In this section, I introduce a mechanism that can explain this finding and present some empirical evidence supporting it.

4.1 Mechanism

I argue that this relatively surprising phenomenon can be explained by a model of competition with differentiated products. In general, the entry of a firm encourages the incumbent to lower prices in order to attract consumers. This is the market share effect: if firms have to divide the market more, they are willing to lower prices to increase profit. Chen and Riordan (2008) show that in a world of horizontal differentiation, another effect is at play: the price-sensitivity effect. The idea is that when a firm enters, consumers choose not only based on the prices of the incumbent and the entrant but based on the quality or type offered. Consumers who stay with the incumbent reveal their ideal taste to be closer to the incumbent than to the entrant. This would be the case of lower-income households preferring the standard products of incumbents stores to the organic and "healthy" products proposed by Whole Foods. The price-sensitivity effect may therefore dominate the market share effect.

Allowing for asymmetry, it can be shown that the effect of entry on an incumbent's price depends on the initial location of the incumbent. Applying it to the specific empirical context of my study, incumbents that can compete with Whole Foods will fight for their share by raising quality or decreasing prices, while incumbents that cannot compete with Whole Foods will be left with a smaller market share, but will be able to lower quality or raise prices since the consumers who remain are very unlikely to switch to either the other incumbent or Whole Foods.

The detailed model can be found in the online appendix.

4.2 Differentiation

The mechanism suggested above relies on product differentiation. A large literature has shown that lower-income and higher-income households have quite different tastes, for example Allcott et al. (2019) and Handbury (2021). Because of the anonymity required by the Nielsen data, I cannot show the difference between Whole Foods' assortment and the other stores' assortment. Here, I propose to look at the (anonymous) chains of stores most favored by lower-income households compared to higher-income households, and vice versa, in my dataset. I start by defining share of expenditure by retailer and group (lower half vs upper half of the income distribution) over the entire dataset. I then select the retailers that have reached a share above 5%. I define a retailer as catering to lower-income households if this group favors this retailer much more than others.

The chains favored by low-income households sell products that are 5.5% less expensive on average. This could be due to product assortment or to a competitive approach. Controlling for barcodes, stores from the chain favored by low-income households still charge 4.5% lower prices on average, which means that both mechanisms are at play. This is even more obvious when looking at the chain favored by high-income households: without controlling for barcodes, prices are almost twice higher. However when controlling for barcodes prices are "only" 20% higher, suggesting that a lot of the overall price difference is due to higher quality in that chain. These descriptive results support the idea that the stores that low-income households and high-income households shop at and the products they buy are very different, which can explain why the entry of Whole Foods would have a different impact on prices paid by low-income households and prices paid by high-income households.

Detailed results can be found in the online appendix Table 1.

4.3 Complementary mechanism: a few exits

The mechanism suggested above relies on an increase of the number of firms operating in the markets, or zipcodes. An alternative story for the result I obtain is that high-end stores, such as Whole Foods, simply displace lower-end stores which may used to cater to lower-income population by providing cheap goods. It should be noted that the main specification of the empirical analysis controls for barcode by store fixed effects, which means that for low-income households, prices of the goods bought in the same stores increase after the entry of Whole Foods relative to before. Still, a displacement effect could be happening simultaneously, and is consistent with the narrative related by people advocating for barriers to gentrification. I therefore propose to study the evolution of the number of stores in the zipcodes as an outcome. If I find that the number of stores stays constant, it would mean that there is a replacement story going on. I use the Census's Zipcode Business Patterns to do this. Unfortunately, these data are at the yearly level, so the estimation is much less precise than the ones for prices.

$$N_{mt} = \sum_{j=-1}^{3} \beta_j \mathbb{I}(\text{Years Since Entry}_{mt} = j) + \delta_m + \eta_t + \epsilon_{mt}$$

The results for grocery store and convenience stores are shown in Figure 3. Three years after the entry of Whole Foods, the number of grocery stores (excluding convenience stores) operating in the zipcode is predicted to increase by 0.3, a statistically significant effect. This effect is also strictly smaller than one, which suggests there are zipcodes where stores exit. This suggests that the entry of Whole Foods indeed has a market share effect on stores, leading some to lose profit and exit. This could further help surviving firms raise prices.

[figure 3]

5 Conclusion

Policymakers worry about gentrification. An oft-cited positive development is the increased entry of businesses. Since these new businesses *a priori* target the new residents, as opposed to the historical and lower-income residents of these neighborhoods, it is important to study the consequences their entry has on consumer welfare, particularly that of lower-income households.

This paper shows that between 2004 and 2017, the entry of a Whole Foods store, a harbinger of gentrification, into a new zipcode caused prices of consumer packaged goods purchased by lower-income households at incumbent stores of 3% three years after entry. This result is robust to a range of robustness checks, including extending the sample of stores and including a more precise set of controls. I also run a falsification test using announcement dates instead of entry dates and find no effect on prices at incumbent stores.

To explain this finding, I propose a mechanism of Bertrand competition with differentiated products, showing that if the products offered by the competitors are neither too similar nor too dissimilar, the entry of a firm may cause prices to rise. This rare effect is generated by the fact that consumers split across the different firms based on their underlying taste. If the entrant locates close to an incumbent, as Whole Foods locates close to stores serving high-income households, entry has a pro-competitive effect and the incumbent's price decreases in order to extend its customer base (market share effect). But if the entrant locates far from the incumbent, as Whole Foods locates far from stores serving low-income households, entry may have anti-competitive effect and the incumbent's price increases, as the remaining customers reveal themselves to prefer the incumbent's offering (price-sensitivity effect).

While policymakers may hardly discourage entry, policies that encourage a wide assortment of products in stores, including products that cater to lower-income households, may dampen the negative consequences of the business side of gentrification.

References

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell, "Food deserts and the causes of nutritional inequality," *The Quarterly Journal of Economics*, 2019, *134* (4), 1793–1844.

Atkin, David, Benjamin Faber, and Marco Gonzalez-Navarro, "Retail globalization and household welfare: Evidence from mexico," *Journal of Political Economy*, 2018, *126* (1), 1–73.

Behrens, Kristian, Brahim Boualam, Julien Martin, and Florian Mayneris, "Gentrification and Pioneer Businesses," *The Review of Economics and Statistics*, 11 2023, pp. 1–14.

Chen, Yongmin and Michael H Riordan, "Price-increasing competition," *The RAND Journal of Economics*, 2008, *39* (4), 1042–1058.

Deck, Cary and Jingping Gu, "Price increasing competition? Experimental evidence," *Journal of Economic Behavior & Organization*, 2012, *84* (3), 730–740.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform pricing in us retail chains," *The Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo, "Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?," *QME*, 2010, *8* (2), 207–239.

Ershov, Daniel, "Competing with superstars in the mobile app market," *Available at SSRN 3265662*, 2018.

Faber, Benjamin and Thibault Fally, "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *The Review of Economic Studies*, 09 2021, *89* (3), 1420–1459.

Glaeser, Edward L, Hyunjin Kim, and Michael Luca, "Measuring gentrification: Using yelp data to quantify neighborhood change," Technical Report, National Bureau of Economic Research 2018.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, "Endogenous gentrification and housing price dynamics," *Journal of Public Economics*, 2013, *100*, 45–60.

Handbury, Jessie, "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across US Cities," *Econometrica*, 2021, *89* (6), 2679–2715.

Iacovone, Leonardo, Beata Javorcik, Wolfgang Keller, and James Tybout, "Supplier responses to Walmart's invasion in Mexico," *Journal of International Economics*, 2015, *95* (1), 1–15.

Javorcik, Beata S and Yue Li, "Do the biggest aisles serve a brighter future? Global retail chains and their implications for Romania," *Journal of International Economics*, 2013, *90* (2), 348–363.

Perloff, Jeffrey M, Valerie Y Suslow, and Paul J Seguin, "Higher prices from entry: Pricing of brand-name drugs," *U of California, Berkeley, Competition Policy Working Paper No. CPC99-03*, 1995.

Ward, Michael B, Jay P Shimshack, Jeffrey M Perloff, and J Michael Harris, "Effects of the private-label invasion in food industries," *American Journal of Agricultural Economics*, 2002, *84* (4), 961–973.

Zukin, Sharon, Valerie Trujillo, Peter Frase, Danielle Jackson, Tim Recuber, and Abraham Walker, "New retail capital and neighborhood change: Boutiques and gentrification in New York City," *City & Community*, 2009, *8* (1), 47–64.

Main Figures

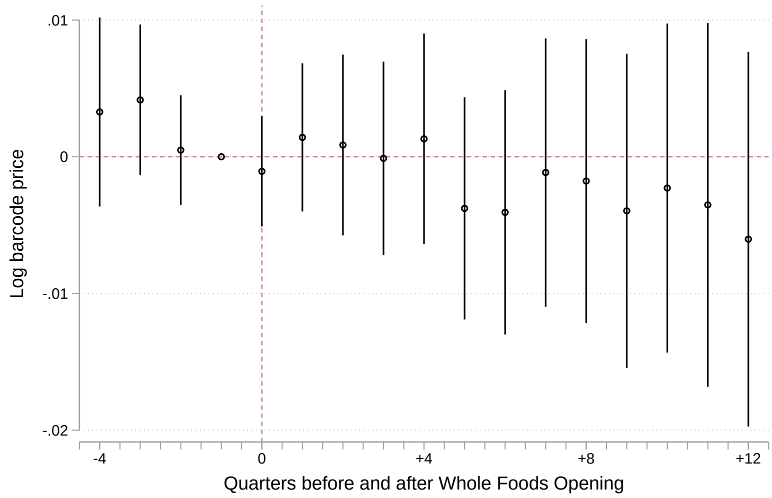
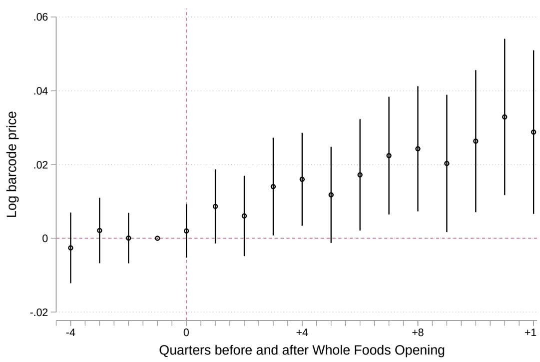
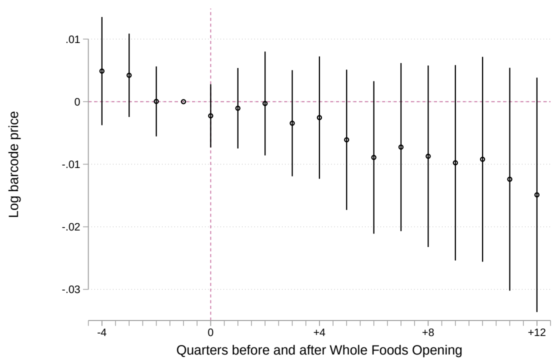


Figure 1: Effect of Entry on Prices at Incumbent Stores

Notes: Figure shows the coefficients and 95% confidence intervals obtained from a regression of log prices on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects and month fixed effects. Here, I only keep stores that participate or have participated in the Nielsen RMS program, and are therefore identified individually. The reference category for each graph is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See Online Appendix (Table 2) for details.





(a) Below median income, only RMS stores (b) Above median income, only RMS stores

Figure 2: Effect of Entry on Prices at Incumbent Stores, by Income Group

Notes: Figure shows the coefficients and 95% confidence intervals obtained from two regressions of log prices on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects and month fixed effects, run separately for households at the bottom of the income distribution (left-hand side Panel) and at the top of the income distribution (right-hand side Panel). Here, I only keep stores that participate or have participated in the Nielsen RMS program, and are therefore identified individually. The reference category for each graph is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See Online Appendix (Table 2) for details.

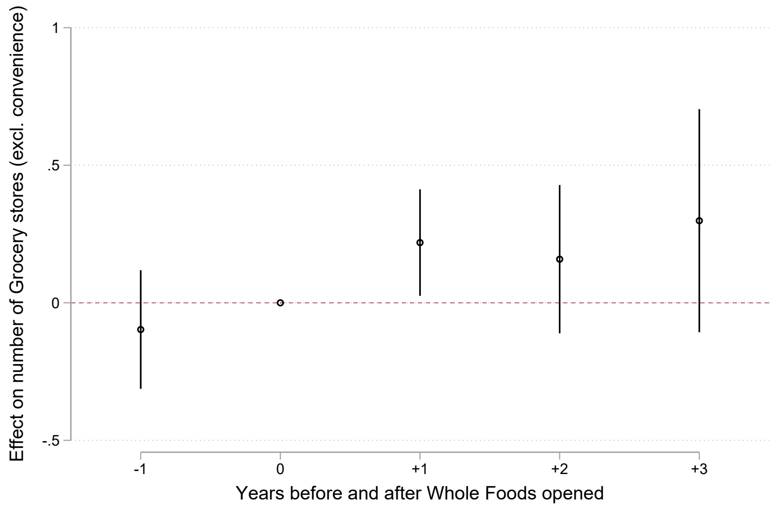


Figure 3: Effect of Entry on Number of Grocery stores in Zipcode

Notes: Figure shows the coefficients and 95% confidence intervals obtained from a regression of number of stores on 5 annual treatment effects, in addition to zipcode fixed effects and month fixed effects. The reference category is the year the Whole Foods store opened. Standard errors are clustered at the zipcode level.

Online Appendix
The Price Effect of Retail Gentrification*

Louise Guillouët

Columbia University

April 27, 2024

*Columbia University Department of Economics (email: louise.guillouet@columbia.edu). Researcher's own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Model

I follow Perloff et al. (1995) in presenting a simple model of Bertrand competition with differentiated products to explain my results. Perloff et al. (1995) discuss the entry of one firm compared to a monopoly. I study the entrance of an asymmetric competitor in two-firm market.

Suppose there is a mass L of consumers. They are described by an ideal quality $\hat{t}$, distributed uniformly along the Salop circle.

Their utility from consuming good i is described by

$$u_i = v - p_i - c|\hat{t} - t_i|$$

where v is the utility from consuming any good in that category, p_i is the price of the good, t_i is the quality of the good, and c is the marginal utility cost of consuming a good that is distant from one's ideal quality.

The consumer chooses the good that maximizes their utility, and buys it if it generates more utility than the outside option, u_0.

1.1 Local Monopolists

To fix ideas, suppose there is only one good. The consumer will choose it if

$$v - p_i - c|\hat{t} - t_i| \geq u_0$$

w.l.o.g., I normalize $u_0 = 0$. The monopolist can sell to consumers who are within the "monopoly region", that is within the distance x_m of its location:

$$|\hat{t} - t_i| \leq x_m = \frac{v - p_m}{c}$$

The monopolist faces total demand $2x_mL$. It chooses its price to maximize its profit taking into account its marginal cost m, where $v > m$

$$\max_{p_m} \; 2L\left(\frac{v - p_m}{c}\right)(p_m - m)$$

The F.O.C. gives

$$p_m = \frac{v + m}{2}, \; x_m = \frac{v - m}{2c}$$

If the two firms are sufficiently far apart, they will each charge p_m and attract all the consumers in their local monopoly region. I assume this is the baseline situation before Whole Foods opens.

1.2 A Third Firm Enters

Consumers may be located in three areas: between the entrant and one incumbent, between the entrant and the other incumbent, or between the two incumbents. Suppose store A is located z_A away from the entrant, while store B is located z_B away from the entrant, with $z_A > z_B$.

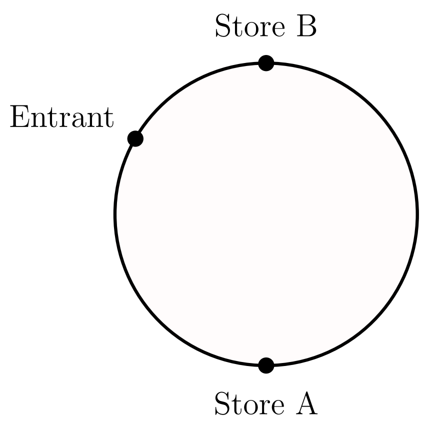


Figure 1: Competitive Landscape after Entry

Since the two incumbents are not truly competing with each other, I will treat each incumbent separately. For the entrant to compete with an incumbent, it must be that the sum of each firm's monopoly region is greater than their distance from each other.

$$\frac{v - p_e}{c} + \frac{v - p_i}{c} > z_i, \;\; \forall i = A, B$$

The marginal consumer is such that the demand faced by the incumbent from the consumers located between the two firms depends on x_i, defined as:

$$v - p_i - cx_i = v - c(z_i - x_i) - p_e \;\forall i = A, B$$

$$x_i = \frac{z_i}{2} + \frac{p_e - p_i}{2c} \quad \forall i = A, B$$

Therefore the demand is kinked at $\bar{p_i} = v - \frac{cz_i}{2}$. I will focus on the case above the kink.[1]

[1]

$$D = \begin{cases} L(x_m + x_i) & \text{if } \frac{v-p_e}{2} + \frac{v-p_i}{2} > z_i \\ 2Lx_m & \text{if } \frac{v-p_e}{2} + \frac{v-p_i}{2} < z_i \;\forall i = A, B \end{cases}$$

Above the kink, the incumbent faces the lower part of the demand curve, such that (assuming symmetry) $v - p < z$.

$$\max_{p_i} 2L\frac{v - p_i}{c}(p - m)$$

At the kink, the derivative of profit is negative

$$2L\frac{v - 2p_i + m}{c}|_{p=v-\frac{cz_i}{2}} = \frac{2L}{c}(v - 2v + cz_i + m) < 0 \text{ since } v > m$$

The incumbent faces demand $L(x_m+x_i)$ (x_m from the consumers between itself and the other incumbent and x_i from the consumers between itself and the entrant). Since it cannot price discriminate, it sets its price to maximize

$$\max_{p_i} L\left(\frac{z}{2}+\frac{p_e-p_i}{2c}+\frac{v-p_i}{c}\right)(p_i-m)$$

$$p_A=\frac{3m+2v+cz_A+p_e}{6},\ p_B=\frac{3m+2v+cz_B+p_e}{6}$$

While the entrant faces demand from consumers located between itself and firm A and between itself and firm B, respectively:

$$\max_{p_e} L\left(\frac{z_A}{2}+\frac{p_A-p_e}{2c}+\frac{z_B}{2}+\frac{p_B-p_e}{2c}\right)(p_e-m)$$

$$p_A=\frac{2m+cz_A+cz_B+p_A+p_B}{5}$$

The solution is

$$p_i=\frac{84m+48v+29cz_i+7cz_{-i}}{132}\ \forall i=A,B\quad p_e=\frac{7cz_A+7cz_B+18m+4v}{22}$$

For any two z_A, z_B such that

$$29cz_B+7cz_A<18(v-m)<29cz_A+7cz_B,$$

$p_A>p_m$ and $p_B<p_m$, so store A raises its price compared to before entry and store B decreases its price. The intuition behind this model is that when the entrant locates somewhat close to the store A ($z_A<z$), but not too close ($z_B>z$), some consumers will greatly prefer the entrant and switch. Meanwhile, the incumbent will keep the consumers who greatly prefer its own product. So the average distance the updated pool of customers has to travel to be at firm A is less than it used to be, which means firm A can raise its price.

1.3 Price Effect

Figure 2 summarizes the effect entry has on prices in the market, depending on how differentiated the entrant is from the incumbent. When the firms are not very differentiated (small z), firms are fierce competitors so entry leads prices to decrease. Intuitively, the market share effect dominates: the incumbent now has to share the market with the entrant, and in order to maintain profits will decrease prices to expand the market base.

so the optimal price is the kink price $p_i=v-\frac{cz_i}{2}>p_m$. It decreases in z until it is back to the (local) monopoly price.

As the distance between the two firms increases, some customers will strongly prefer one product or the other, allowing the incumbent firm to leverage the price-sensitivity effect. The intuition is that customers who stick to the incumbent, while fewer in number than before entry, are on average closer in taste to the incumbent's offering. Therefore, they are on average willing to pay a higher price to buy from this firm. When the level of differentiation is higher than the monopoly region, $z > x_m$, the price-sensitivity effect dominates and entry causes prices to rise. However, the firms are still competing. When products are so differentiated that the firms are local monopolies, $z > \bar{z}$, the market share effect starts to kick in more, incentivizing firms to lower their price to capture more customers, until the firms are optimally behaving like monopolies, so that the price effect of entry is zero.

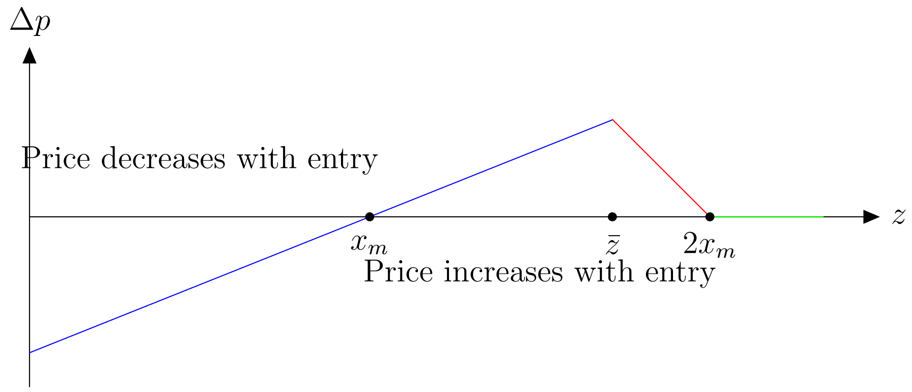


Figure 2: Effect of Distance on Price Effect of Entry

The intuition here is that in general different stores serve low-income and high-income consumers, perhaps meaning that although they are technically competitors, stores can charge high, monopoly level prices because they are very differentiated. When Whole Foods comes in, it competes directly with the stores serving high-income consumers ($z < x_m$), leading to a decrease in price, and less intensely with the stores serving low-income consumers ($z_m < z < 2x_m$), leading to an increase in price in these stores. The mechanism that remains to be tested is the fact that Whole Foods is closer to the taste of rich consumers than poor consumers.

1.4 Welfare Changes

The total consumer surplus from the two local monopolists is

$$CS_{2m} = 2\left(v - p_m - cE[x|x < x_m]\right) 2x_m = \frac{1}{2}\frac{(v-m)^2}{c}$$

The total consumer surplus from the three firms is:

$$CS_{3f} = \sum_i (v - p_i - cE[x|x < x_i])2x_i$$

$$+(v - p_i - cE[x|x < x_m(p_i)])x_m(p_i) + (v - p_W - cE[x|x < x_W(p_i)])x_W(p_i)$$

Overall, consumer surplus is almost always higher when there are more firms, even if the price increases for one incumbent firm. However, the welfare changes are not homogeneous. Figure 3 is an illustration of how customers located at different points on the circle (linearized for simplicity) around one incumbent and one entrant are affected by the entry of a firm.

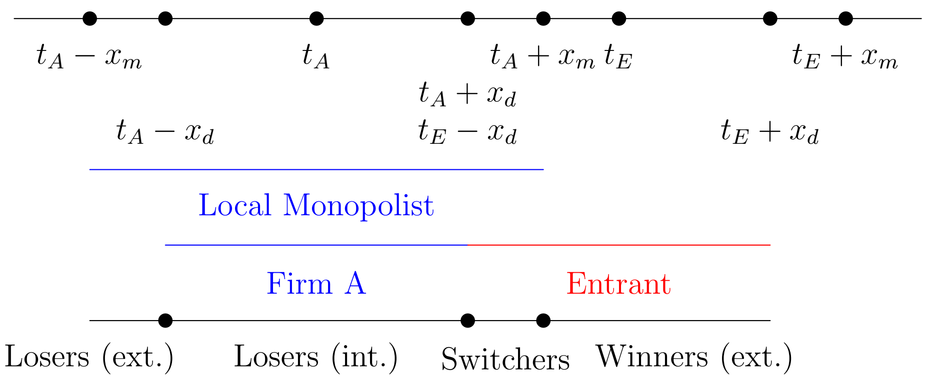


Figure 3: Distribution of Welfare Consequences when the Price Increases after Entry

Note that first, the customers who were not served previously benefit because they are now happy to purchase a product on the market, where previously they might have to drive elsewhere to get their groceries, or eat out. Because more customers are served, I call this a gain at the extensive margin.

Second, the customers who decide to switch to the new product may gain or lose. This may seem surprising given these consumers' ideal tastes are closer to the entrant than to the incumbent. Indeed, when the duopoly price is lower than the monopoly price these customers gain.[2] However, this is not the case when switchers have to pay a higher price under duopoly than under monopoly.

Third, some customers lose welfare as they are buying the same good in both cases but pay a higher price. This is a loss of welfare at the intensive margin.

Last, in this model some customers lose welfare as they no longer find it in their interest to purchase the good now that its price has increased. This is a loss of welfare at the extensive margin.

[2]Suppose an individual located x away from the incumbent switches. It means that $v - p_e - c|z - x| > v - p_A - cx \;\Leftrightarrow\; c|z - x| < cx + p_e - p_A$, they prefer the entrant's good conditional on price. Visually they are located closer to the entrant than to the incumbent. How does his utility after entry compare to his prior utility? $v - p_e - c(z - x) - (v - p_m - cx) = p_m - p_e - cz$ This change is ambiguous, and is negative when the duopoly price is higher than the monopoly price.

2 Additional figures

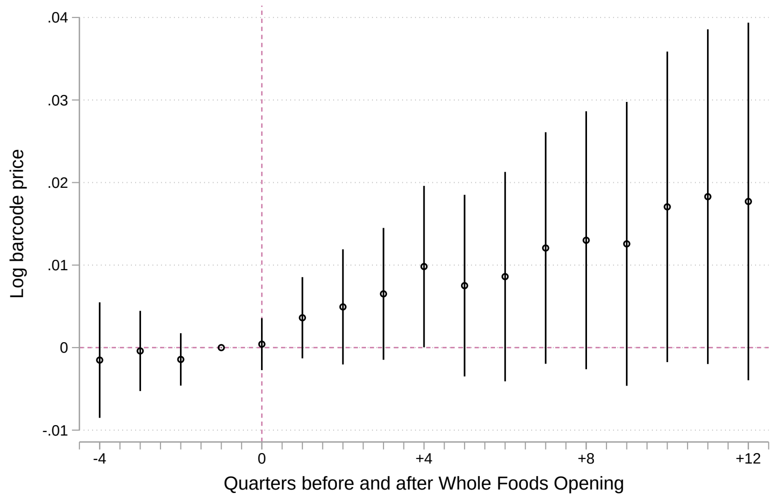
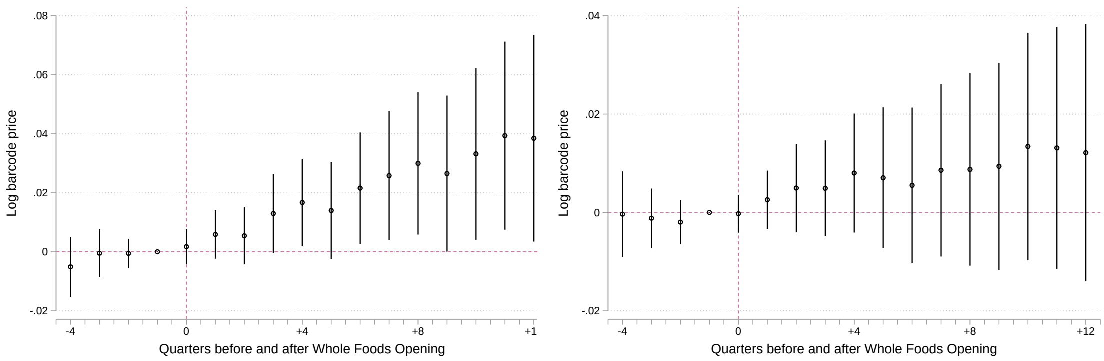


Figure 4: Effect of Entry on Prices at All Stores

Notes: Figure shows the coefficients and 95% confidence intervals obtained from a regression of log prices on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects and month fixed effects. Here, stores are defined by the unique combination of the retailer code attributed by Nielsen, the zipcode of the resident and the store code attributed by Nielsen for those stores that participate in the RMS program. The reference category is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See Table 2 for details.



(a) Below Median Income (b) Above Median Income

Figure 5: Effect of Entry on Prices at All Stores, by Income Group

Notes: Figure shows the coefficients and 95% confidence intervals obtained from two regressions of log prices on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects and month fixed effects, run separately for households at the bottom of the income distribution (left-hand side Panel) and at the top of the income distribution (right-hand side Panel). Here, stores are defined by the unique combination of the retailer code attributed by Nielsen, the zipcode of the resident and the store code attributed by Nielsen for those stores that participate in the RMS program. The reference category for each graph is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See Table 2 for details.

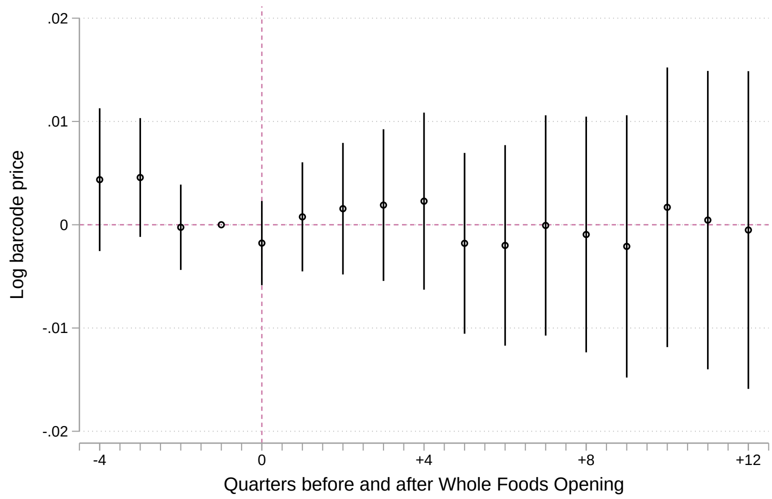
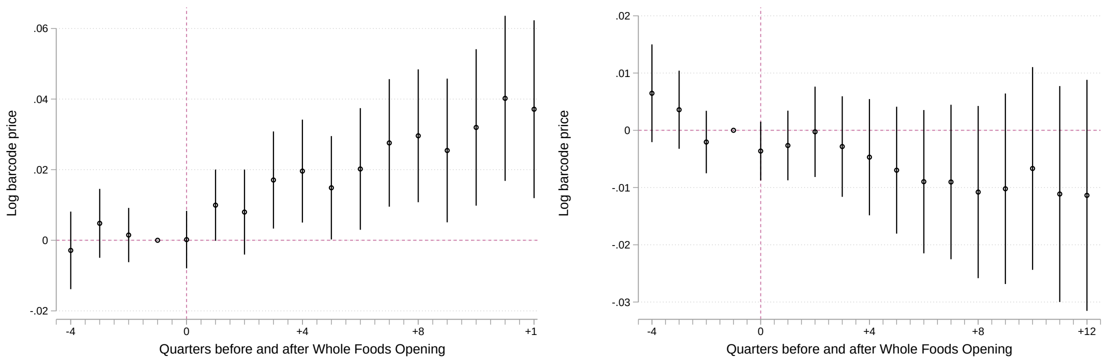


Figure 6: Effect of Entry on Prices at Incumbent Stores, Additional Controls

Notes: Figure shows the coefficients and 95% confidence intervals obtained from a regression of log prices on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects, state-by-month fixed effects, zipcode size-by-month fixed effects and store type-by-product group-by-month fixed effects. Here, I keep only those stores that participate in the RMS program. The reference category is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See the paper and Table 3 for details.



(a) Below Median Income

(b) Above Median Income

Figure 7: Effect of Entry on Prices at Incumbent Stores, Additional Controls, by Income Group

Notes: Figure shows the coefficients and 95% confidence intervals obtained from a regression of log prices on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects, state-by-month fixed effects, zipcode size-by-month fixed effects and store type-by-product group-by-month fixed effects, run separately for households at the bottom of the income distribution (left-hand side Panel) and at the top of the income distribution (right-hand side Panel). Here, I keep only the stores that participate in the RMS program. The reference category is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See the paper and Table 3 for details.

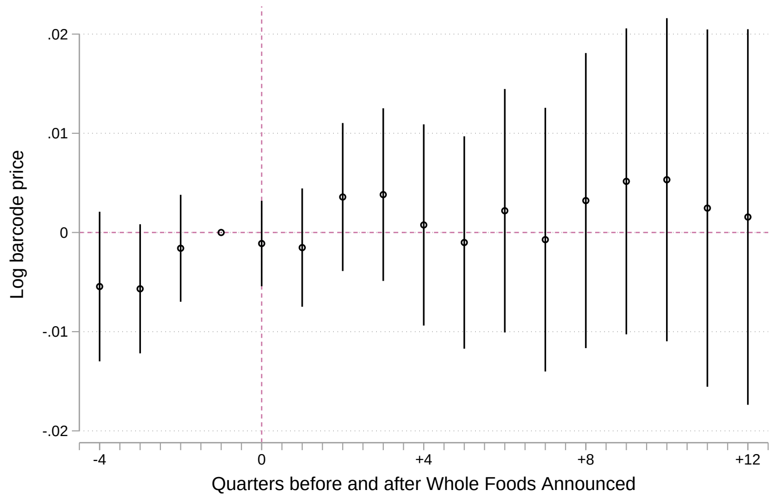
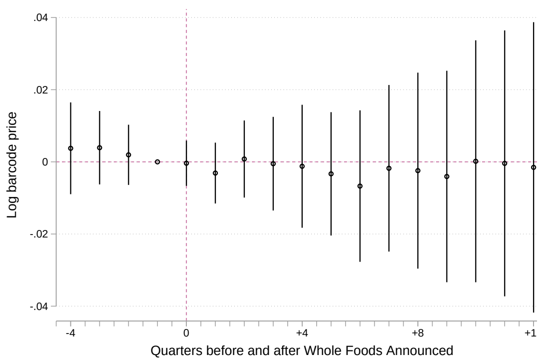
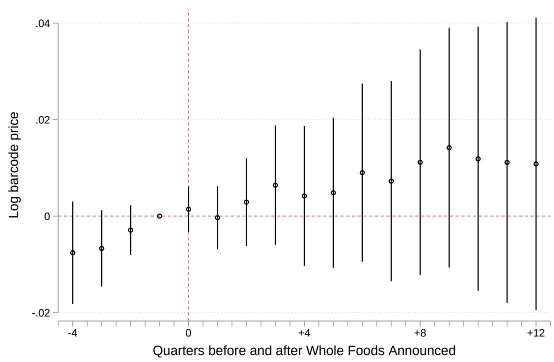


Figure 8: Effect of Announcement on Prices at Incumbent Stores

Notes: Figure shows the coefficients and 95% confidence intervals obtained from a regression of log prices on 16 quarterly "false" treatment effects which are actually announcement effects, in addition to barcode-by-store fixed effects and month fixed effects. Here, I keep only those stores that participate in the RMS program. The reference category is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See the paper and Table 4 for details.



(a) Below Median Income



(b) Above Median Income

Figure 9: Effect of Announcement on Prices at Incumbent Stores, by Income Group

Notes: Figure shows the coefficients and 95% confidence intervals obtained from two regressions of log prices on 16 quarterly "false" treatment effects which are actually announcement effects, in addition to barcode-by-store fixed effects and month fixed effects, run separately for households at the bottom of the income distribution (left-hand side Panel) and at the top of the income distribution (right-hand side Panel). Here, I keep only those stores that participate in the RMS program. The reference category is the price one quarter before the Whole Foods store opening. Standard errors are clustered at the zipcode level. See the paper and Table 4 for details.

3 Additional Tables

	Log price			
	(1)	(2)	(3)	(4)
Low-income Chain	-0.055	-0.045		
	(0.014)	(0.007)		
Hich-income Chain			0.951	0.207
			(0.015)	(0.055)
Zipcode by year FEs	Yes	Yes	Yes	Yes
Zipcode by product by month	Yes	Yes	Yes	Yes
Zipcode by barcode by month	No	Yes	No	Yes
Number of zipcodes	233	233	233	233
R2	0.60	0.90	0.63	0.90
N	4499642	2101400	4499642	2101400

Table 1: Differences Between Chains, by Income Group

Notes: Table reports the coefficients from regressing log price on a dummy identifying a chain that is preferred by low-income households (Columns 1-2) or high-income households (Columns 1-3). Standard errors are clustered at the zipcode level and are reported in parenthesis.

Households	All	< $50k$	≥ $50k$	All	< $50k$	≥ $50k$
	Log price					
	(1)	(2)	(3)	(4)	(5)	(6)
Four Quarters Before	0.003	-0.003	0.005	-0.002	-0.005	-0.000
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)
Three Quarters Before	0.004	0.002	0.004	-0.000	-0.000	-0.001
	(0.003)	(0.005)	(0.003)	(0.002)	(0.004)	(0.003)
Two Quarters Before	0.001	0.000	0.000	-0.001	-0.001	-0.002
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
One Quarter Before	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)
Quarter of Entry	-0.001	0.002	-0.002	0.000	0.002	-0.000
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)
One Quarter After	0.001	0.009	-0.001	0.004	0.006	0.003
	(0.003)	(0.005)	(0.003)	(0.002)	(0.004)	(0.003)
Two Quarters After	0.001	0.006	-0.000	0.005	0.005	0.005
	(0.003)	(0.006)	(0.004)	(0.004)	(0.005)	(0.005)
Three Quarters After	-0.000	0.014	-0.003	0.007	0.013	0.005
	(0.004)	(0.007)	(0.004)	(0.004)	(0.007)	(0.005)
Four Quarters After	0.001	0.016	-0.003	0.010	0.017	0.008
	(0.004)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)
Five Quarters After	-0.004	0.012	-0.006	0.008	0.014	0.007
	(0.004)	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)
Six Quarters After	-0.004	0.018	-0.009	0.009	0.022	0.006
	(0.005)	(0.008)	(0.006)	(0.006)	(0.010)	(0.008)
Seven Quarters After	-0.001	0.023	-0.007	0.012	0.026	0.009
	(0.005)	(0.008)	(0.007)	(0.007)	(0.011)	(0.009)
Eight Quarters After	-0.002	0.025	-0.009	0.013	0.030	0.009
	(0.005)	(0.009)	(0.007)	(0.008)	(0.012)	(0.010)
Nine Quarters After	-0.004	0.022	-0.010	0.013	0.027	0.009
	(0.006)	(0.010)	(0.008)	(0.009)	(0.013)	(0.011)
Ten Quarters After	-0.002	0.028	-0.009	0.017	0.033	0.013
	(0.006)	(0.010)	(0.008)	(0.010)	(0.015)	(0.012)
Eleven Quarters After	-0.003	0.035	-0.012	0.018	0.039	0.013
	(0.007)	(0.011)	(0.009)	(0.010)	(0.016)	(0.012)
Twelve Quarters After	-0.005	0.032	-0.015	0.018	0.038	0.012
	(0.007)	(0.012)	(0.010)	(0.011)	(0.018)	(0.013)
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Store-barcode FEs	Yes	Yes	Yes	No	No	No
Retailer-zipcode-barcode FEs	No	No	No	Yes	Yes	Yes
N	2322396	698893	1550254	3105372	959659	2055517
R2	0.90	0.91	0.90	0.92	0.93	0.92
Store-barcode cells	507622	158702	346414	673997	215494	455809
N zipcodes	233	211	225	233	211	225

Table 2: Effect of Entry on Prices (Main Results)

Notes: Table reports regressions of log price on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects and month fixed effects. Columns 1-3 show the regressions results only keeping stores that have a Nielsen store code (see Figures **??** and **??**). Columns 4-6 expand the sample by defining the stores as the unique combination of the retailer code attributed by Nielsen, the zipcode of the resident and the Nielsen store code when available (see Figures 4 and 5). Columns 1 and 4 show the regression results using all households in the panel. Columns 2 and 5 show the regression results using only households in the bottom half of the income distribution (defined as less than $50,000). Columns 3 and 6 show the regression results using only households in the top half of the income distribution. Standard errors are clustered at the zipcode level and are reported in parenthesis.

Households	All (1)	< $50k (2)	≥ $50k (3)	All (4)	< $50k (5)	≥ $50k (6)
	Log price					
Four Quarters Before	0.004	-0.003	0.006	-0.001	-0.005	-0.000
	(0.004)	(0.006)	(0.004)	(0.003)	(0.005)	(0.004)
Three Quarters Before	0.005	0.005	0.004	-0.001	0.001	-0.003
	(0.003)	(0.005)	(0.003)	(0.002)	(0.004)	(0.003)
Two Quarters Before	-0.000	0.001	-0.002	-0.002	-0.000	-0.004
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)
One Quarter Before	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)
Quarter of Entry	-0.002	0.000	-0.004	0.000	0.002	-0.001
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)
One Quarter After	0.001	0.010	-0.003	0.003	0.009	0.001
	(0.003)	(0.005)	(0.003)	(0.002)	(0.004)	(0.003)
Two Quarters After	0.002	0.008	-0.000	0.005	0.007	0.005
	(0.003)	(0.006)	(0.004)	(0.003)	(0.005)	(0.004)
Three Quarters After	0.002	0.017	-0.003	0.008	0.015	0.005
	(0.004)	(0.007)	(0.004)	(0.004)	(0.006)	(0.005)
Four Quarters After	0.002	0.020	-0.005	0.010	0.019	0.007
	(0.004)	(0.007)	(0.005)	(0.005)	(0.007)	(0.006)
Five Quarters After	-0.002	0.015	-0.007	0.009	0.017	0.008
	(0.004)	(0.007)	(0.006)	(0.005)	(0.008)	(0.007)
Six Quarters After	-0.002	0.020	-0.009	0.010	0.025	0.007
	(0.005)	(0.009)	(0.006)	(0.006)	(0.009)	(0.007)
Seven Quarters After	-0.000	0.028	-0.009	0.013	0.031	0.009
	(0.005)	(0.009)	(0.007)	(0.007)	(0.010)	(0.008)
Eight Quarters After	-0.001	0.030	-0.011	0.013	0.033	0.008
	(0.006)	(0.010)	(0.008)	(0.007)	(0.011)	(0.009)
Nine Quarters After	-0.002	0.025	-0.010	0.013	0.028	0.009
	(0.006)	(0.010)	(0.008)	(0.008)	(0.013)	(0.010)
Ten Quarters After	0.002	0.032	-0.007	0.019	0.035	0.015
	(0.007)	(0.011)	(0.009)	(0.009)	(0.014)	(0.011)
Eleven Quarters After	0.000	0.040	-0.011	0.020	0.042	0.014
	(0.007)	(0.012)	(0.010)	(0.010)	(0.015)	(0.011)
Twelve Quarters After	-0.001	0.037	-0.011	0.020	0.041	0.015
	(0.008)	(0.013)	(0.010)	(0.010)	(0.016)	(0.012)
Store type-group-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Region-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode size-month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Store-barcode FE	Yes	Yes	Yes	No	No	No
Retailer-zipcode-barcode FEs	No	No	No	Yes	Yes	Yes
N	2318840	693983	1546151	3102732	955704	2052462
R2	0.90	0.92	0.90	0.92	0.93	0.92
Store-barcode cells	506476	157170	345065	673121	214267	454794
N zipcodes	233	211	225	233	211	225

Table 3: Effect of Entry on Prices (Additional Controls)

Notes: Table reports regressions of log price on 16 quarterly treatment effects, in addition to barcode-by-store fixed effects, state-by-month fixed effects, zipcode size-by-month fixed effects and store type-by-product group-by-month fixed effects. Columns 1-3 show the regressions results only keeping stores that have a Nielsen store code (see Figures 6 and 7). Columns 4-6 extend the sample by defining the stores as the unique combination of the retailer code attributed by Nielsen, the zipcode of the resident and the Nielsen store code when available. Columns 1 and 4 show the regression results using all households in the panel. Columns 2 and 5 show the regression results using only households in the bottom half of the income distribution (defined as less than $50,000). Columns 3 and 6 show the regression results using only households in the top half of the income distribution. Standard errors are clustered at the zipcode level and are reported in parenthesis.

Households	All	$< \$50k$	$\geq \$50k$	All	$< \$50k$	$\geq \$50k$
	Log price					
	(1)	(2)	(3)	(4)	(5)	(6)
Four Quarters Before	-0.005	-0.005	-0.006	-0.003	0.004	-0.008
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.005)
Three Quarters Before	-0.006	-0.002	-0.007	-0.003	0.004	-0.007
	(0.003)	(0.005)	(0.004)	(0.003)	(0.005)	(0.004)
Two Quarters Before	-0.002	-0.003	-0.001	-0.002	0.002	-0.003
	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)
One Quarter Before	0.000	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)	(.)
Quarter of Entry	-0.001	0.003	-0.002	-0.000	-0.000	0.001
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)
One Quarter After	-0.002	0.003	-0.002	-0.003	-0.003	-0.000
	(0.003)	(0.005)	(0.004)	(0.003)	(0.004)	(0.003)
Two Quarters After	0.004	0.007	0.002	0.001	0.001	0.003
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Three Quarters After	0.004	0.007	0.004	0.002	-0.001	0.006
	(0.004)	(0.007)	(0.006)	(0.005)	(0.007)	(0.006)
Four Quarters After	0.001	0.008	-0.001	-0.000	-0.001	0.004
	(0.005)	(0.009)	(0.007)	(0.006)	(0.009)	(0.007)
Five Quarters After	-0.001	0.006	-0.002	-0.001	-0.003	0.005
	(0.005)	(0.009)	(0.008)	(0.006)	(0.009)	(0.008)
Six Quarters After	0.002	0.005	0.003	0.000	-0.007	0.009
	(0.006)	(0.011)	(0.009)	(0.008)	(0.011)	(0.009)
Seven Quarters After	-0.001	0.007	-0.003	0.001	-0.002	0.007
	(0.007)	(0.012)	(0.010)	(0.009)	(0.012)	(0.011)
Eight Quarters After	0.003	0.009	0.002	0.004	-0.002	0.011
	(0.008)	(0.014)	(0.011)	(0.010)	(0.014)	(0.012)
Nine Quarters After	0.005	0.011	0.004	0.005	-0.004	0.014
	(0.008)	(0.015)	(0.012)	(0.011)	(0.015)	(0.013)
Ten Quarters After	0.005	0.017	0.002	0.005	0.000	0.012
	(0.008)	(0.018)	(0.012)	(0.012)	(0.017)	(0.014)
Eleven Quarters After	0.002	0.018	-0.003	0.004	-0.000	0.011
	(0.009)	(0.020)	(0.014)	(0.013)	(0.019)	(0.015)
Twelve Quarters After	0.002	0.021	-0.004	0.003	-0.002	0.011
	(0.010)	(0.021)	(0.014)	(0.013)	(0.020)	(0.015)
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Store-barcode FEs	Yes	Yes	Yes	No	No	No
Retailer-zipcode-barcode FEs	No	No	No	Yes	Yes	Yes
N	1910174	591344	1254819	2506189	792956	1635723
R2	0.89	0.90	0.90	0.91	0.92	0.92
Store-barcode cells	430400	137839	289855	553594	180869	369904
N zipcodes	180	162	177	180	162	177

Table 4: Effect of Announcement on Prices (Falsification Test)

Notes: Table reports regressions of log price on 16 quarterly announcement effects (see the paper for details), in addition to barcode-by-store fixed effects and month fixed effects. Columns 1-3 show the regressions results only keeping stores that have a Nielsen store code (see Figures 8 and 9). Columns 4-6 extend the sample by defining stores as the unique combination of the retailer code attributed by Nielsen, the zipcode of the resident and the Nielsen store code when available. Columns 1 and 4 show the regression results using all households in the panel. Columns 2 and 5 show the regression results using only households in the bottom half of the income distribution (defined as less than \$50,000). Columns 3 and 6 show the regression results using only households in the top half of the income distribution. Standard errors are clustered at the zipcode level and are reported in parenthesis.