*Journal of Economic Geography*, 2023, 1–26
https://doi.org/10.1093/jeg/lbad034
**Original Article**

# Are local retail services an amenity or a nuisance?

Pooya Ghorbani[1] and Rachel Meltzer[2,*]

[1]New York State Homes and Community Renewal, New York, NY 10022, United States
[2]Department of Urban Planning and Design, Harvard University, Cambridge, MA 02138, United States

*Corresponding author. Department of Urban Planning and Design, Harvard University, 48 Quincy Street, Cambridge, MA 02138, United States. E-mail: rmeltzer@gsd.harvard.edu

## Abstract

We test whether local retail services are considered a nuisance or an amenity and how this distinction is capitalized into residential property values. Using a rich, micro-spatial dataset on property sales transactions and business activity in New York City, we estimate the impact of access to neighborhood retail services on residential sales prices. We construct two instruments to channel supply-side drivers of retail change and to address concerns of endogeneity between changes in retail activity and property values. Results show that retail services that are more frequently consumed and experiential, and are located in relatively more mixed-use neighborhoods are positively capitalized into property values. Residents also pay more to be closer to more diverse retail clusters, and relatively less to be closer to chains. This is true across smaller 1-to-4 family homes as well as larger condos/coops and multi-family rental buildings. The price effects from certain classifications of retail, like restaurants and personal services, are mixed depending on the kind of residential property and the local concentration of the retail. Therefore, the relative strength of the amenity or nuisance effect is very much conditioned on the type of service and the localized neighborhood context.

**Keywords:** retail; amenities; property value; capitalization.

**JEL classifications:** R00, R3

## 1. Introduction

Much of the public finance literature on neighborhood disparities in service provision focuses on the availability and quality of *public* goods and amenities (e.g. Wilson 1987; Massey and Denton 1993; Black 1999; Cutler, Glaeser, and Vigdor, 1999; Jargowsky 2003). However, there is little research on how *private* services affect the quality and desirability of particular neighborhoods. The access to and variety of nearby retail services, like a supermarket, pharmacy, or laundromat, can be part of a household location decision, just as public safety, sanitation, and transportation are. But, how are these local services priced into those housing decisions?

It is not clear, particularly in an urban setting, whether the traditional "separation of uses" is preferred over more mixed-use communities. Theoretical predictions are ambiguous: retail services can be both positively (the "amenity effect") and negatively capitalized (the "nuisance effect"). Which effect is dominant and whether it varies depending on the type of and proximity to services are empirical questions. The findings in this research are particularly timely given the unknown repercussions for urban retail and services in the wake of COVID-19 shutdowns and slowdowns. Quantifying the amenity value of retail services is important for land use and neighborhood planning. It also has fiscal implications, as cities rely heavily on the revenue generated from retail transactions and properties.

**Received:** 20 December 2021. **Editorial decision:** 6 November 2023. **Accepted:** 30 November 2023

In order to estimate price effects from localized retail services, we rely on parcel-level sales data, as well as microdata on retail establishments, all of which span a decade. Therefore, we can isolate effects from retail most likely to serve nearby residents and test for heterogeneous price responses across a range of retail types and micro-distances. We construct two instruments to channel supply-side drivers of retail change and to address concerns of endogeneity between changes in retail activity and property values.

Results from both OLS and 2SLS estimations show that retail services that are more frequently consumed and experiential and are located in denser mixed-use neighborhoods, are positively capitalized into property values. This is true across smaller 1-to-4 family homes as well as larger condos/coops and multi-family rental buildings. There is also a clear price gradient, such that the price premium from neighborhood retail is largest at the closest range. Across the board, residents also pay more for housing that is closer to more diverse retail clusters, and relatively less close to chains. As for specific classes of retail, the patterns are more varied. Restaurants bestow price premia for 1–4 family and larger multi-family homes but generate price discounts for condos/coops, which tend to be closer to higher concentrations of restaurants. The price effects are mixed for personal services (a heterogeneous sub-group of retail) and largely negative for properties near food and beverage establishments. This suggests heterogeneity in the amenity value from retail, and that potential nuisances may counteract any benefit from having a higher concentration of those services nearby.

The article proceeds in the following way. Section 2 provides theoretical framing and Section 3 presents a review of the relevant empirical literature. Section 4 describes the data and empirical strategy, and Section 5 presents the results. Section 6 concludes.

## 2. Are retail services an amenity or a nuisance?

Starting in the 2000s and until the recent pandemic, US cities had experienced a "great inversion," or the return of younger, more affluent, and typically white, households. Empirical evidence has demonstrated that this reversal in trends from the earlier decades of flight is largely due to shifting tastes toward urban amenities (Carlino and Saiz 2008; Couture and Handbury 2020). This aligns with another established pattern, that the magnitude and nature of retail densities are correlated with the demographic and economic compositions of nearby consumers (Waldfogel 2008; Couture and Handbury 2020). Left unanswered is how much households value the proximity to such amenities and how that informs their location choices across neighborhoods within cities. This pattern of preferences is likely different in dense urban areas than more sprawling suburban or rural locations—the access is more proximate by orders of magnitude and the very essence of urban living is the intermingling of residential and commercial uses. We start with the premise that variation in preferences for nearby retail services, if it exists, should be reflected in one of the best metrics for location valuation—real estate prices.

How much will households pay to be close to retail amenities? Is there a price premium for being closer to certain types of services? Theoretically, the direction (and degree) of price capitalization for local retail services is ambiguous. Unlike most (quality-provided) public services, which are overwhelmingly positively capitalized into property values, private services, especially in proximity, can be viewed as both a benefit and a nuisance for local residents. Furthermore, where we assume that differentiation among public services largely relies on their quality, private neighborhood services, especially in denser mixed-use settings, are valued just as much, or perhaps even more, based on their physical proximity (Hotelling 1929). Consumers will pay more to be closer to retail services to forgo the transportation costs of patronizing them from farther away. This premium for homebuyers is likely the largest for retail and services that are a necessity or frequently consumed.

We rely on a simple scenario, a closed city where households choose among a set of neighborhoods with varying densities and bundles of retail services. We abstract away from the quality of those services, but we do allow the type and composition of those services to vary and assume quality will be at least partially captured by these observable features. All else equal, households will choose a location that maximizes their preference for retail *access*, which should be reflected in the price paid for the housing.

## 2.1 Retail services as a local nuisance

Traditional forms of zoning and land use have assumed that any mix of use is undesirable. If these assumptions are accurate, then proximity to non-residential uses should pull down home prices. Specifically, retail can bring with it noise, litter, odors, and increased street traffic. However, the nature and magnitude of any of these factors should vary by the type of retail. For example, restaurants and bars would bring noise, foot traffic, and possibly crime, during late hours, whereas grocery stores provide necessary goods without many negative byproducts. If, on the net, quiet streets and less foot traffic are valued more than convenience or nearby amusement, home prices relatively closer to these kinds of services will be lower than those located farther away.

## 2.2 Retail services as a local amenity

Local retail services can be included among the locational characteristics that are viewed as assets. Having a supermarket, pharmacy, or laundromat, that is, services that are consumed often (Berry 1967), nearby can markedly affect the quality of life and these amenity-based benefits, like other public services, should be reflected positively in the price of the home. Indeed, households might pay more for housing within walking distance of such amenities in lieu of investing in transportation (like a car or public transit). Jane Jacobs (1961) also discusses the importance of an active streetscape to discourage crime and illicit activities and to promote vibrancy. A local retail presence is an important contributor to street life and can even affect mental and social well-being (Cox and Streeter 2019). Safety, and visceral benefits more generally, can bestow meaningful and positive price effects on nearby residential properties.

## 2.3 Qualifications

The strength of either the amenity or nuisance effect is qualified by three important features.[1] First, as discussed above, any positive or negative capitalization will depend on the *type* of retail. Second, the *mix* of retail should matter. As with public services, consumers will not only respond to the intensity of individual services but the composition of the package. We expect that households will prefer diversity and complementarity in nearby services that they frequent more often over any benefits endowed by clusters of similar establishments (which would be more in line with destination comparison-shopping; Nelson 1958).

Third, the degree of capitalization will depend on *proximity* to the retail service(s). In this case, closer may not always be better; indeed, there may be a distance from the retail services that bestows the most value. The service should be close enough to preserve convenience, but with enough of a buffer to distance from the potential noise and activity. That said, cities (and New York in particular) are inherently mixed-use, and therefore the integration of retail and residential may be more appreciated.

## 2.4 Retail is an endogenous amenity

Complicating the above scenario is the fact that retailers and households simultaneously make decisions about where to locate across neighborhoods. For retail establishments, these decisions are based on the characteristics of localized markets and consumers—that is, the very households behind the housing price transactions (Stroebel and Vavra 2019; Borraz et al., 2020). Retailers consider both characteristics of the residents themselves (such as earnings or race) or the wealth endowments, and therefore propensity to spend, indicated by the housing assets (Mian and Sufi 2011). Therefore, the demand for retail amenities is endogenously determined by the factors that drive the localized supply of retail services.

[1] There are several retail features that are hard to measure but may signal things about a neighborhood that affect prices. For example, the physical appearance of the retail storefront can influence the amenity effect. Storefronts that are more contextual and conform to surrounding facades might induce relatively higher price effects than those that are viewed as disruptive or offensive to the local context. Also, smaller and boutique establishments, especially of a certain brand or image, may provide signals about the economic status or even race of who lives in the neighborhood. These signals may be capitalized into prices. Unfortunately, we cannot observe these aspects directly in our data. As a close approximation, we will consider establishment types that are more likely to have appearances tailored to the local context versus those that have more context-neutral designs—for example, standalone compared to chain establishments.

# 3. Empirical literature

## 3.1 Capitalization of private neighborhood services and amenities

Over the past decade, researchers have tried to value amenities in cross-city household location decisions. They find that amenities matter; to what extent depends on the demographics of the household (Chen and Rosenthal 2008; Rappaport 2008; Albouy 2016; Beracha et al., 2018). Unlike these studies, we are concerned with the within-city location decisions of households and how they weigh the benefits and costs of nearby amenities across neighborhoods in the same locality. From this perspective, much of the work has focused on the values of amenities such as open space or parks (Asabere and Huffman 1996; Grudnitski and Do 1997; Benson et al., 1998; Wu, Adams, and Plantinga 2004; Jim and Chen 2006). These studies largely find that proximity to open space or greenery is positively capitalized into property values. We also know that crime, one of the possible nuisance byproducts from retail, tends to devalue residential properties nearby (e.g., Thaler 1978; Taylor 1995; Lynch and Rasmussen 2001).

Perhaps most closely aligned with our approach is Kuang's (2017) study of how the amenity value of restaurants is capitalized into housing prices in Washington, D.C. and a small body of work on how prices respond to the entry of a large retailer. Kuang (2017) relies on Yelp data to measure the quantity and quality of nearby restaurants and finds that both quantity and quality matter, especially those that are well-reviewed (rather than simply more expensive). Also using Yelp data, Davis et al. (2019) find that reviewers in NYC are more likely to patronize restaurants closer to their homes (although they do not capture this proximity premium in housing prices). Unlike these studies, we expand the amenity set beyond restaurants to include other neighborhood services and look at heterogeneity across property types.

Other studies have used the entry of a big box retailer as a supply shock to the area in order to understand various responses in the market, including nearby home prices. Many of these studies find positive price effects but at a distance from the new retailer (Corlija, Siman, and Finke 2006; Johnson and née Lybecker 2018; Daunfeldt et al., 2020). While the identification strategy of these studies is appealing, the price premia are likely explained by mechanisms different than those induced by neighborhood-based retail services (e.g., the labor demand effects). We intend to tease out the value of the services from nearby retailers.

The body of work on how mixed land use and new urbanist planning affect house prices is also closely related to our analysis. These studies take a hedonic regression approach and incorporate the neighborhood land use characteristics as one of the locational controls in the model. One of the earliest attempts to price mixed land use externalities was by Cao and Cory (1981) in a cross-sectional analysis of single-family homes in Tucson, Arizona in 1970. They find that the amount of industrial and commercial (as well as multi-family and public use) land is positively associated with nearby residential property values. They interpret these findings to indicate that there is some optimal mix of land use in which the land is most productively used, and it does not necessarily conform to the traditional separation of uses under Euclidean Zoning. This is consistent with a study by Sirpal (1994) where he finds that proximity to shopping centers, especially bigger ones, is associated with higher residential housing prices. He interprets this as evidence of positive spillovers growing with the intensity of the nearby retail use (Ellis, Lee, and Kweon (2006) find the opposite in their survey-based analysis).

Since this early work, studies have developed more precise metrics for "mixed-use" and the topography, or walkability, of the local community. Song and Knaap (2003) empirically test the utility of new urbanism by estimating hedonic pricing models for nearly 200 neighborhoods in Washington County, Oregon. They include in their model, among other covariates, an index of land use mix, other measures of accessibility to commercial uses, and a general walkability index. While no single metric gets specifically at the location and nature of retail services, they consider a range of amenities (and disamenities) and how proximity to these various uses can impact property values. The authors find that differences in the composition of urban form are indeed capitalized into housing prices. Specifically, and related to our central question of retail valuation, residents will pay more for better pedestrian access to commercial uses (but they pay less for neighborhoods that themselves contain more commercial uses). This is consistent with the results from their other paper, which tests more broadly for the effect of mixed land uses on single-family housing values (2004) and from another study by Song and Sohn (2007), who use retail land use classification as a proxy for retail services.[2] Koster and Rouwendal

[2] Two studies, however, find null effects from nearby commercial districts and related nuisances (Wu, Adams, and Plantinga 2004; Jim and Chen 2006)

(2012) conduct a similar study in the Netherlands, using semiparametric estimation methods, and find a 2.5 per cent price premium for houses located in mixed-use neighborhoods; this premium is particularly prominent for houses near land occupied by business service and leisure activities. Similarly, Rauterkus and Miller (2011) test for the impact of walkability on land values and find a positive effect. In particular, prices are higher for neighborhoods closer to the central business district, in older historic areas and near university campuses. While he doesn't look at walkability, Couture (2016) also finds gains from amenity density that are driven partly by shorter trip times (by car) to nearby restaurants, but mostly from the variety (and therefore amenity choices) that density supports.

The takeaways are inconclusive and limited in their direct observation of retail activity (versus designated use). We focus on the de facto retail activity and isolate the kind of services more likely to be capitalized as neighborhood amenities. In addition, unlike many of the cross-sectional studies to date (Shimizu et al., 2014; Chiang, Peng, and Chang 2015; Jang and Kang 2015), we address the endogeneity of retail amenities by using longitudinal data, fine-grained geographic controls and an instrumental variable identification strategy in our analysis.

# 4. Data and empirical strategy

## 4.1 Data

We build a longitudinal database of commercial activity and property values in neighborhoods across New York City spanning 2002 to 2010. We rely on two core datasets. First, we use a proprietary dataset, the National Establishment Time Series (NETS) database, with the location, type, and opening/closing dates for nearly the universe of businesses across New York City. The dataset is constructed by Walls and Associates from the Dun and Bradstreet business register. Unlike publicly available government data on employment, the NETS include little or no suppression of employment in small industries or geographic cells. For this analysis, we were provided establishment counts at the city block level, for each industry–year in our study period. In addition, the industry is reported as fine as the six-digit NAICS level, and a headquarters identifier permits the classification of establishments according to firm size and structure.[3]

We recognize the documented limitations of the NETS data (Neumark, Zhang, and Wall 2005; Barnatchez, Crane, and Decker 2017; Crane and Decker 2019). Most concerns relate to the accuracy of the employment and sales fields (neither of which we use in our analysis) and the employment and establishment dynamics over time. While we consider establishments over time, our identification relies on changes in aggregates of establishments rather than the life cycles of individual firms or the tracking of particular establishments over time (Barnatchez, Crane, and Decker (2017) confirm that NETS correlates better with other publicly available data sources at higher geographic aggregates). Our approach is more in line with analyzing repeated cross-sections, rather than following a panel of establishments, the latter of which has been flagged as vulnerable to short-term change calculations. Furthermore, the assessment conducted by Crane and Decker (2019) indicates that any discrepancies between NETS establishment counts and those reported by other sources (such as QCEW and CBP) are smallest for retail-oriented industries (our sector of interest) and consistent over time. Therefore, while the point estimates of establishments at any moment in time may be attenuated, the changes should not be affected by any systematic inconsistencies in the levels over time.

Second, we obtain data for all property sales transactions during the study period from the Furman Center for Real Estate and Urban Policy at New York University and special arrangements with the New York City Department of Finance. Property sales transactions are recorded through a combination of deed transfers and real property transfer tax returns filed with the City Register's office. Each filing contains a parcel identifier for the property sold, the effective date of the transaction, the price, an indicator of property type (to identify 1–4 family, larger multi-family, condominium, and cooperative properties), and additional circumstances of the transfer that allow us to determine whether or not it is an arm's length sale. We drop all non-residential transactions, sales with zero or negative prices, and parcel sales for those parcels with duplicate transactions on the same day. Using a standardized parcel identifier, we are able to add on parcel-level physical characteristics—such as lot size, number

[3] We have access to NETS data aggregated to the city block, based on industry combinations we specifically requested for this analysis. These sectors include: NAICS 44-45, 722, 454, 7223, 52211, 52213, 71394, 812111-2, 81231-2, and 81291. However, it means we cannot go back into the raw data to make new aggregations, including NAICS codes at levels finer than six digits.

**Table 1.** Summary Statistics for prices and hedonics by property type.

| | 1–4 Family | | | Condo/Coop | | | Large multifamily | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **N** | **mean** | **sd** | **N** | **mean** | **sd** | **N** | **mean** | **sd** |
| Price per sq. ft. | 261,944 | **383** | 11,307 | 109,023 | **1,530** | 42,995 | 14,828 | **196** | 204 |
| Gross sq. ft. | 263,304 | **2,149** | 1,340 | 109,314 | **1,733** | 8,743 | 14,923 | **20,072** | 89,753 |
| Number of units | 266,059 | **1.68** | 0.76 | 109,628 | **1.61** | 8.19 | 15,497 | **23.03** | 85.12 |
| Lot front (feet) | 263,314 | **30** | 26 | 109,314 | **131** | 131 | 14,923 | **50** | 55 |
| Stories | 263,304 | **2.14** | 0.65 | 109,314 | **16.48** | 15.02 | 14,923 | **4.2** | 2.31 |
| Age (years in 2010) | 261,886 | **67** | 32 | 108,813 | **38** | 36 | 14,826 | **85** | 19 |
| Years since altered (in 2010) | 263,304 | **54** | 43 | 109,314 | **11** | 30 | 14,923 | **44** | 46 |
| Proportion w/elevator | 263,304 | **0.00** | | 109,628 | **0.76** | | 15,513 | **0.15** | |
| Proportion in historic district | 263,304 | **0.02** | | 109,628 | **0.11** | | 15,513 | **0.09** | |

*Note:* The table displays summary statistics for the sample of property sales transactions, by property type. Mean values are displayed in bold. Dollar values are adjusted to reflect 2013 values.

of stories, number of units, gross square footage, presence of an elevator, age, and year of last alteration—from the Department of Finance's Tax Roll files as well as land use characteristics—specifically whether a property has been landmarked or included in a historic district—from the Department of City Planning's PLUTO database. Table 1 displays summary statistics for sales price and parcel characteristics.

New York City offers a rich testing ground for this kind of analysis: it both provides enough scale to exploit micro-variation and is comprised of a diverse set of neighborhoods in terms of commercial activity and residential real estate. Residents' access to retail is quite localized, and there is variation in the degree of localization. Furthermore, the presence of retail does not always correspond with higher prices—this too varies spatially. Supplementary Appendix A show the variation in both real prices per square foot and the concentration of retail across the entire city as of 2002. This localized variation persists over time, as some neighborhoods experienced strong price growth *and* commercial changes, while other neighborhoods remained stable. In Clinton Hill in Brooklyn, for example, the average price per square foot of a two- to four-family home and the retail density more than doubled from 2002 to 2010. On the other hand, other neighborhoods in Brooklyn, like Canarsie, saw declines in both prices and retail activity. Neighborhoods, like Kew Gardens in Queens, saw prices increase by about 60 per cent and retail decline by 20 per cent; Sugar Hill in Manhattan, for example, experienced the opposite (price declines and retail growth, over the same time period).

## 4.2 Variable construction

The dependent variable of interest is real sales price.[4] We log transform this variable to better fit the non-normal distribution of the data (and to make more intuitive the interpretation of the coefficients). We estimate prices separately for three different property types: single-family, multi-family condominium/co-op, and multi-family rental. This disaggregation is motivated by the fact that hedonic fundamentals look different across these property types and that the occupants of these properties likely differ in characteristics that are correlated with preferences for retail amenities (Stroebel and Vavra 2019). This plays out in our data: we can see in Table 1 and Supplementary Appendix E that the structural features and nearby retail densities vary by property type. We also observe that the relationship between neighborhood residential characteristics and the nature of property transactions and retail densities varies across property types.[5]

The independent variable of interest is the prevalence of nearby retail activity, which will be operationalized in several ways. Again, we focus on retail *access* and do not differentiate by retail quality, as we do not have credible ways to measure this (save for the chain designation, described below). First, to capture the *intensity* of retail access we calculate the number of retail establishments in close

[4] All dollar amounts have been adjusted to 2013 dollars.

[5] The statistics are not shown here, but separating by property type absorbs a lot the demographic variation at the neighborhood level, that could also relate to differences in retail preferences. Estimating property types separately will also largely capture systematic variation across the five boroughs of the city, which tend to be developed in ways that either prioritize one property type (Manhattan, Staten Island, Queens) or diversify across all types (Bronx, Brooklyn). In addition, the analyses will control for unobserved changes in prices at the borough level over time; see Supplementary Appendix B for a distribution of property types by borough.

proximity to the property sale. We first build a comprehensive customer-facing retail classification by including establishments classified as retail trade (NAICS 44-45) or food services (NAICS 722), except retailers without a store-based point of sale (NAICS 454) and food service contractors and caterers (NAICS 7223). We also include various retail services outside of these sectors, including banking (NAICS 52211 and 52213), fitness (NAICS 71394), barber/beauty shops (NAICS 812111-2), laundry and pet care (NAICS 81231-2 and 81291); see Supplementary Appendix C for a full listing of included sectors. This definition is consistent with aspects of earlier ones generated by Meltzer and Capperis (2017), Helling and Sawicki (2003), Bingham and Zhang (1997), DiPasquale and Wheaton (1996), and Stanback et al. (1981).

Using the spatial coordinates of both sales and the city block of the retail establishments, we draw circular areas of three different radii (1/8, 1/4, and 1/2 mile; the corresponding variable names have extensions "_e," "_q," and "_h," respectively) around every sale in the sample and count the number of establishments within each pre-determined distance. For perspective, a person walking in a straight line at about 3 miles per hour could walk 1/8 mile in just over 4 minutes, 1/4 mile in just over 8 minutes, and 1/2 mile in about 17 minutes.

Consistent with retail gravity models (e.g. Reilly 1931), we prioritize counts of establishments, because access to the amenity itself, rather than a density of amenities for efficient comparison-shopping purposes, should be more relevant for neighborhood services. The value of having a service nearby that is patronized individually (e.g., a spontaneous trip to the pharmacy or hardware store) should be better captured by the retail counts; this should be especially true across the smaller, neighborhood-level distances we track and the specific types of retail we document. We expect that retail densities become more relevant when valuing "destination" corridors, which are not the focus of the current analysis. Without clear priors on how big the "local" retail market is, we use the data and the multiple distances to directly test it. Figures in Supplementary Appendix D provide a visual comparison between the three circular areas, census tracts, and ZIP code areas. Supplementary Appendix E also displays typical establishment counts for the three different circular areas.

We also disaggregate the total retail counts to create variables that measure the *types* of retail. This both captures the heterogeneity in services and also isolates the kinds of services most likely to be consumed (and valued) by nearby residents. First, we distinguish across businesses that provide goods or services that are frequently or infrequently consumed (notated as "*_freq*" and "*_infreq*," respectively). For this distinction, we draw heavily from Helling and Sawicki (2003) who consider a subset of "residentiary services" as those businesses that serve local "consumer demand directly" and provide goods or services that are frequently consumed and/or perishable, whereby short travel times are essential to their appeal. Retailers with frequently consumed goods and services include food markets, drug stores, and restaurants; examples of those with infrequently consumed goods and services are home furnishings, automotive goods, and apparel. Frequently patronized establishments are meant to capture the goods and services that are deemed more necessary in a set of neighborhood amenities. Supplementary Appendix F demonstrates how NAICS codes are classified with respect to frequency or infrequency.

Second, we further disaggregate the type of service into more fine-grained categories: restaurants (notated as "*_rest*"), financial services (i.e. banks, check cashing; notated as "*_fin*"), personal care services (ie pharmacies, Laundromats; notated as "*_ps*"); food and beverage stores (i.e. supermarkets, corner stores; notated as "*_fb*"); and other goods (i.e. furniture stores, clothing stores, book stores; notated as "_goods"). Due to the small geography of analysis, we cannot further disaggregate the retail establishments into more singular categories; therefore, these categories are in some cases quite heterogeneous. We also supplement these categorizations with alternative ones based on whether the goods or services are substitutable with online commerce. Those services and goods less substitutable tend to be experiential (e.g. restaurants and theaters) and where face-to-face interactions are necessary (e.g. barber shops and laundromats). See Supplementary Appendix G for the full list of NAICS codes and their substitutability classifications.

Third, we measure the mix of retail using a Herfindahl index of the above retail categories ($k$) concentration: $\sum_{i=1}^{k} s_i^2$. We expect that lower values, indicating a lower concentration in a single sub-sector and more diversity in goods/services provided by those establishments, will be associated with higher prices.

Finally, we classify businesses based on their ownership structure: specifically, whether they are a chain or independently operated establishment (notated as "*_chain*" and "*_indep*," respectively). This distinction will help us identify, albeit crudely, differences in capitalization related to the storefront aesthetic, and possibly the cost and quality of the goods and services provided.

## 4.3 Empirical strategy

### 4.3.1 Hedonic regression

The regression analysis follows a standard hedonic framework, where the nearby retail activity variables discussed above are treated as one of several locational characteristics ($\boldsymbol{Retail}_{it}$) that contribute to the valuation of a particular residential property. Note that we adjust the retail counts down by ten to avoid unreasonably small coefficients (therefore, the unit change in retail is across every ten establishments, rather than a single establishment). Also included in this vector are counts of total commercial establishments for the relevant circular area, to control for any distinct influence from nearby commercial activity, as well as institutional factors that constrain land use and might otherwise affect house prices. We also include property-specific characteristics ($\mathbf{X}_i$), including size, number of units, frontage, number of stories, age, time since a physical alteration, whether or not the building has an elevator and whether or not it lies in a historic district. We also include zip code dummies ($z$) and broader borough–year ($b$, $t$) fixed effects. All standard errors are clustered by zip code. These controls absorb any heterogeneity that is not captured by the specified covariates and ensure that any property sale is compared to an otherwise similar sale within the same micro-neighborhood.

The general regression equation takes the following form:

$$\ln P_{ibzt} = \beta_0 + \beta_1(\mathbf{X}_i) + \beta_2(\mathbf{Retail}_{it}) + d_z + d_{b,t} + \varepsilon_{it}$$

## 4.4 Mitigating against endogeneity

Thus far, we have treated retail as an exogenous amenity (Brueckner, Thisse, and Zenou 1999). However, if it is the case that retail follows residential investment and locates based on the socioeconomic characteristics of those residents—for example, Borraz et al. (2020), Stroebel and Vavra (2019)—then it needs to be treated as endogenous in order mitigate against any bias in estimating the direct effect of retail activity on prices. To improve upon the above "naïve" model, we use an instrumental variables strategy to better isolate the retail supply channel from the demand one in estimating the effect of nearby retail amenities on prices.

First, we instrument for change in retail activity in the surrounding neighborhood using a Bartik-like shift share variable (Bartik 1991), which essentially imposes an exogenous shock on the supply of local retail. We interact macro-level changes in retail- and service-sector establishments with localized exposures to retail activity. We rely on the fact that this Bartik-type shock will be correlated with changes in neighborhood retail activity but not with the outcome of interest, within-neighborhood changes in house prices (controlling for property-level features):

$$\text{Retail\_shock}^k_{j,t} = \frac{\text{commercial\_area}_{j,t}}{\text{total\_area}_{j,t}} \times \sum \left( \frac{E_{k,t}}{E_{k,t-3}} - 1 \right)$$

where $k$ indexes our curated retail industries (i.e. the set of subsectors that fall under service and retail two- and three-digit NAICS industry classifications identified above) and $j$ indexes the neighborhood (e.g. circular area). We compute the predicted change in retail establishments for the neighborhood between $t$ and $t-3$ by interacting national-level changes in retail establishments (again, only those classified as the service and retail two- and three-digit NAICS industry classifications identified above) over that time period with the neighborhood's concentration of commercial square footage at time $t$.[6]

The consistency of our instrument relies on the "shares assumption" (Goldsmith-Pinkham, Sorkin, and Swift 2020)—the fact that we use differential exogenous exposures (commercial square footage) to

[6] We replicate the growth component using state- and MSA-level values and the results are generally the same. We also use employment instead of establishment counts, but the latter performs slightly better. Finally, we experiment with different change period, such as 5 years. We lose observations with the longer lag and the results are not substantially different, and so we opt for the shorter window.

a common shock (national changes in retail establishments).[7] Furthermore, the shares assumption is motivated by our focus on a limited set of industries (those classified as retail or service). We justify the exposures as reasonably exogenous since allowable retail space is determined through the land use code, whose classifications were determined years before the time period of the current study and tend to be slow changing (in our sample the share of total square footage designated commercial increased by about 5 per cent over the decade of the study period). Therefore, the potential for contemporary changes in housing prices to determine retail land use classifications enshrined years prior is unlikely.[8] In addition, we exclude the neighborhood *j's* own growth in retail from the calculation.

We also note that there is a good deal of variation in establishment growth (and contraction) across retail sub-industries, ranging in magnitude from less than 10 per cent to over 100 per cent, and in industry mix across the neighborhoods. No sub-industry is overly concentrated in any particular circular area, preventing unwanted correlation between national-level growth rates and neighborhood-level retail shocks (see Supplementary Appendix I).

Second, we adapt an amenity density instrument from Couture and Handbury (2020) that exploits the variation in both broad and specific NAICS establishment classifications at fine-grained geographies. Specifically, this instrument isolates the potential for retail competition and cannibalization, as well as positive spillovers, in establishment entry into neighborhood-level markets. For this instrument, we rely on the interaction between the commercial square footage "exposure" and industrial drivers of retail entry and exit as the exogenous shock to prices.

The NAICS industrial classifications in our data range from the broader two-digit to the more fine-grained three- and four-digit levels (unfortunately, not as fine-grained as those available to Couture and Handbury 2020) and therefore we estimate the following reduced form model for each radius of the circular area, $r \in \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}\right\}$:

$$est_{j10}^{NAICS3\_4} - est_{j02}^{NAICS3\_4} = \alpha^{NAICS3\_4} + \beta_{j02}^{NAICS3\_4} + \beta_{j02}^{NAICS3\_4|2} + \varepsilon_j^{NAICS3\_4}$$

The coefficients, $\beta_{j02}^{NAICS3\_4}$ and $\beta_{j02}^{NAICS3\_4|2}$, tell us the net change in the number of establishments between 2002 and 2010 associated with the 2002 presence of establishments with the same three/four-digit or two-digit NAICS classification. Supplementary Appendix J indicates that there is variation in the sign and strength of the estimated agglomeration and cannibalization forces. The patterns are also consistent with what Couture and Handbury (2020) find: establishments with the same fine-grained industrial classification are more likely to deter entry (or induce exit), while those with the same broader two-digit classifications are more likely to induce entry. These forces are also stronger in smaller areas where the establishments operate closer to one another.

We then use the sum of fitted values across all of the three/four-digit NAICS categories as the predicted change in establishments, which is then interacted with the "exposure" variable created for each circular area used above (the share of built square footage classified as commercial).

# 5. Results

## 5.1 Is there a retail price gradient?

We start with the OLS model and display the regression results for the pooled sample and then stratified by type of property.

To validate the data, we first run regressions with the hedonics only, but still include geographic and temporal controls (see the first column of Supplementary Appendix K). The coefficients are generally in the expected directions: prices increase when, controlling for other factors, the property is bigger, is newer (or recently renovated), is in a building with an elevator and is either landmarked or located in a historic district.

The next three columns show the results stratified by property type. While many of the coefficients maintain their sign and significance, there are a few notable exceptions. First, properties with more

[7] This is in contrast to shift-share instruments that achieve consistency through the exogeneity of independent shocks (e.g. the growth variable) to a larger set of industries (as proven by Borusyak, Hull, and Jaravel 2022). Ideally, the instrument achieves exogeneity in both the shares and shocks, but that is unlikely in practice. Goldsmith-Pinkett et al. show that satisfying the shares assumption produces consistent estimators as well.

[8] We identify blocks that underwent rezonings or reclassifications of land use and eliminate these from the analysis in robustness checks. This controls for any influence from massive rezonings that could induce retail or residential shocks and threaten the validity of our assumption of commercial land use persistence, during our study period. The results are consistent with those that include the rezoned areas of the city (see Supplementary Appendix H)

**Table 2.** OLS, the association between proximity to retail establishments and residential prices.

| DV = log(price) | 1–4 Family | | | Condos and coops | | | Large multifamily | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Total Estab, 1/8 M | −0.001 | | | 0.0009* | | | 0.001 | | |
| | (0.001) | | | (0.0004) | | | (0.0005) | | |
| Total Estab, 1/4 M | | −0.0006 | | | 0.0005* | | | 0.0005 | |
| | | (0.0005) | | | (0.0002) | | | (0.0003) | |
| Total Estab, 1/2 M | | | −0.0006* | | | 0.0003** | | | 0.0001 |
| | | | (0.0003) | | | (0.0001) | | | (0.0001) |
| Retail Estab, 1/8 M | 0.004 | | | −0.004 | | | 0.001 | | |
| | (0.003) | | | (0.002) | | | (0.003) | | |
| Retail Estab, 1/4 M | | 0.003* | | | −0.003* | | | 0.0008 | |
| | | (0.002) | | | (0.001) | | | (0.0002) | |
| Retail Estab, 1/2 M | | | 0.004*** | | | −0.002* | | | 0.0009 |
| | | | (0.0009) | | | (0.0007) | | | (0.0009) |
| Hedonics? | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| $R^2$ | 0.09 | 0.09 | 0.09 | 0.64 | 0.64 | 0.64 | 0.61 | 0.61 | 0.61 |
| Standard error | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster |
| N | 261,886 | 261,886 | 261,886 | 108,813 | 108,813 | 108,813 | 14,826 | 14,826 | 14,826 |

*Note*: This table shows the association between the proximity to retail establishments and transacted property prices. Columns (1)–(3) for each property type show the effect of retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial establishments in that same area. Standard errors (in parentheses) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of total establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, and ZIP dummies.

* $P < .05$, ** $P < .01$, *** $P < .001$.

units are discounted in condo/coop transactions. Finally, an elevator generates a significant price premium only for larger multi-family rental buildings. It is not surprising that the magnitudes of all of the hedonic coefficients vary by property type. There are enough differences across the property types that from this point on we display only the stratified results.

Next, we add in measures of retail services. These results are displayed in Table 2. First, we add in a count of total retail establishments nearby, while also controlling for the total number of commercial establishments within 1/8, 1/4, and 1/2 miles.[9] While not shown, the hedonics remain in the regression and are largely unchanged from those displayed in Supplementary Appendix K.

While the estimates vary depending on property type, there is a clear price gradient across the board. Any price changes associated with nearby retail are largest in magnitude at closer distances to the property. However, those at the closest range (i.e. 1/8 mile) tend to be less precisely estimated. For 1–4 family properties, nearby retail services are positively associated with prices, and the price gradient is the least pronounced across the property types. The positive premium is most significant at a 1/4-mile distance or farther. It is also small in magnitude: there is a 0.4 per cent increase in price (or approximately $2,400 based on the mean price in the sample) for an additional ten retailers within 1/2 mile. Although small at the margin, this translates to over $65,000 in higher transaction prices for a typical 1/2 mile area with 274 retail establishments (assuming a constant marginal valuation of additional retailers nearby). For condos/coops, retail has the opposite effect. Retail is negatively capitalized into prices, and, again, only at distances 1/4 mile or bigger. The magnitude is similarly small (about 0.2 per cent for every ten establishments at 1/2 mile). There is no significant price association for larger

[9] For completeness, we replicate the regressions with density measures of retail per total built square footage. As discussed above, the density metric is inherently picking up a different kind of retail amenity than the neighborhood one that we isolate. The results from regression analyses support this. For example, the 2SLS models produce price gradients that are considerably flatter and reversed when using the density metric: the price premia/discount is slightly bigger at farther distances (most significantly for 1–4 family properties). This is also the case for the "mixed-use" models, which still show concentrated effects in the mixed residential-retail neighborhoods. We interpret these differences as consistent with the expectation that the price gradient of retail density metrics will be less sensitive over smaller increments of space and generally increase in value for longer distances traveled. These results are displayed in Supplementary Appendix L for a subset of specifications.

rental buildings (although the coefficients are positive and similarly declining in magnitude over space).

We validate the relevant retail geography in one more way. The retail counts until now have been spatially cumulative; that is the measures of establishments within 1/2 mile of the property sale also include those establishments within 1/8 and 1/4 miles. We redefine the radii as mutually exclusive "donuts" (or "net" counts), such that the establishment counts within 1/8 mile are not included in the 1/4 mile radius, which is in turn not included in the 1/2 mile radius. These results are displayed in Supplementary Appendix M. We control for total commercial activity within 1/2 mile of the property sale. The results are consistent with those presented above, but the premium of distance becomes even more pronounced. The 1–4 family prices are most positively associated and condo/coop prices are most negatively associated at a 1/2 mile distance. For consistency, we present the remaining results for the three circular areas separately, although they are in line with those estimated using the donut measures. We opt for the former specification as it provides more estimating power when we start to disaggregate the establishment counts into smaller subsets of retail and makes the 2SLS estimation presented below less cumbersome (as we will have to address only one endogenous retail area metric at a time, versus the three donut metrics in a single specification).

## 5.2 Retail metric validation

We now run analyses to validate these findings against alternative metrics of retail. In order to identify any nonlinear patterns in price effects, we re-estimate the stratified models using categorical representations of retail presence (the bottom tenth percentile is omitted as the reference category). In the models thus far, we obtain price effects from the marginal establishment; however, the price response could actually depend on the aggregate number of retail establishments near the property. Figure 1 plots out the coefficients from an estimation of prices on deciles of retail concentration for each circular area radius. Due to fewer sales transactions in the condo/coop and multi-family classifications, we are only able to fully estimate the deciles for the larger 1–4 family subsample and display only this property type (although the partial estimations from the other property types suggest similar patterns). The estimates approximate a u-shaped pattern, such that price changes dip around moderately dense retail and then peak at the highest densities. These results suggest that the value of nearby retail amenities is most pronounced when they are minimal (e.g. the difference between having no services and a few) or when they are plentiful and perhaps more varied (we expand on the latter point below).

Next, we re-estimate the above regression using measures of establishment entry and exit over the five years preceding the sales transaction, instead of (and in addition to) annual establishment counts (shown in Supplementary Appendix N).[10] This allows us to interrogate how much of the above estimates are driven by the gross inflow or exit of retail services (versus the net change). When we control the entry and exit of establishments, we notice a few things. First, the entry of retail has no significant association with prices, except for larger multi-family properties. Second, in all property types, retail exits are associated with declining prices (and by relatively large magnitudes). Third, the coefficients on the original retail variables remain largely unchanged from the baseline model. In this iteration, those coefficients capture the price effect from nearby retail, net of any impact from establishments entering or exiting.

Finally, we also test whether price effects are influenced by the mix of retail rather than simply the presence of any retail. We consider the mix of retail across the sub-sector categories defined above (e.g. restaurants, personal services, etc.). We add an indicator, the Herfindahl index, to the baseline model (interacted with the original retail variables), such that a bigger value indicates *less* diversity (or *more* concentration) in the types of retail. Results are displayed in Table 3. For all property types, this parameter is most precisely estimated for the largest area of 1/2 mile. In addition, for all property types, a higher concentration of retail types is associated with lower prices, suggesting that residents value diversity in the local retail mix. Furthermore, including this control does not generally alter the interpretation of the value of the marginal retail establishment (although its magnitude trumps, by

[10] The entry/exit variables are available to us at the tract level, which does not align perfectly with the circular areas (it is typically bigger than the 1/8 mile, but smaller than 1/2 mile areas). *Entry* is measured as the share of establishments in time $t-5$ that have entered into the tract for property $i$ by time $t$; *Exit* is measured as the share of establishments in time $t-5$ that have exited or shut down from the tract for property $i$ by time $t$. The 5-year interval for changes mitigates concerns discussed above about the NETS' accuracy in capturing short-term changes in establishments.

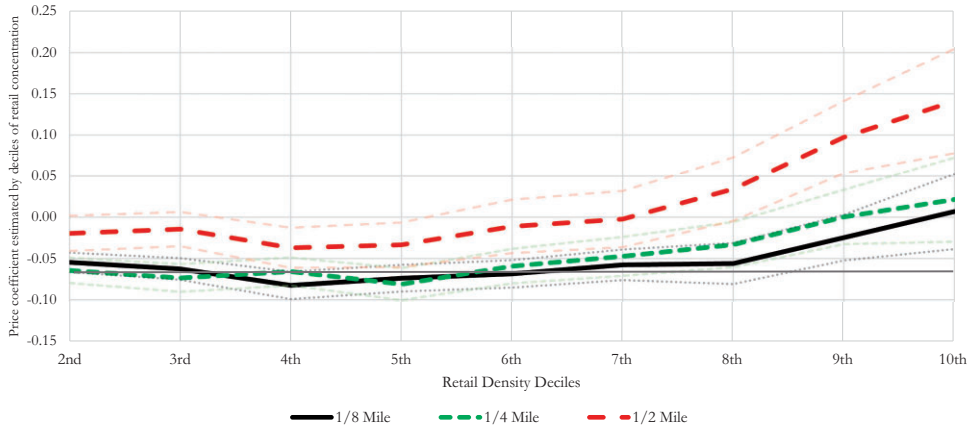


**Figure 1.** OLS, the association between proximity to retail establishments and residential prices, establishment density deciles, 1–4 family. *Note:* This figure shows the association between the proximity to retail establishments and transacted property prices. The x-axis plots retail density (retail per total square footage) deciles and the y-axis plots price effects (the estimated coefficient for each retail density decile). Dotted lines plot price effects for circular areas of varying distances from the block of the property transaction between 1/8 and 1/2 mile. Standard errors (in fine dotted lines) are clustered by ZIP. There is a U-shaped valuation of nearby retail amenities—the price premium is most pronounced when there is any or a lot of retail. Hedonics and control variables included in the regression, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. We show only 1–4 family estimates as the deciles could not be estimated separately for the smaller samples of condo/coops and multifamily sales.

**Table 3.** OLS, association between proximity to retail establishments and residential prices, controlling for retail mix (Herfindahl index).

| | 1–4 Family | | | Condo/Coop | | | Rental | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Retail Estab, 1/8 M | 0.008 | | | −0.007** | | | 0.0006 | | |
| | (0.007) | | | (0.003) | | | (0.005) | | |
| Retail Estab x HI, 1/8 M | −0.012 | | | 0.007 | | | 0.052 | | |
| | (0.019) | | | (0.004) | | | (0.011) | | |
| Retail Estab, 1/4 M | | 0.009* | | | −0.003 | | | 0.003 | |
| | | (0.004) | | | (0.002) | | | (0.003) | |
| Retail Estab x HI, 1/4 M | | −0.019 | | | 0.001 | | | −0.007 | |
| | | (0.012) | | | (0.003) | | | (0.008) | |
| Retail Estab, 1/2 M | | | 0.008*** | | | 0.001 | | | 0.005** |
| | | | (0.002) | | | (0.0008) | | | (0.002) |
| Retail Estab x HI, 1/2 M | | | −0.016* | | | −0.007** | | | −0.015*** |
| | | | (0.007) | | | (0.002) | | | (0.004) |
| Hedonics | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| $R^2$ | 0.09 | 0.09 | 0.09 | 0.64 | 0.64 | 0.65 | 0.61 | 0.61 | 0.61 |
| Standard error | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster |
| *N* | 260,491 | 261,860 | 261,878 | 108,606 | 108,813 | 108,813 | 14,825 | 14,826 | 14,826 |

*Note:* This table shows the association between the proximity to retail establishments and transacted property prices, controlling for the diversity of retail types nearby. Retail diversity is measured by a Herfindahl index of retail types (restaurants, financial services, personal services, food and beverage, and other services) and a higher value indicates less diversity. Columns (1)–(3) for each property type show the effect of retail and retail diversity in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial establishments in that same area. Standard errors (in parentheses) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of total establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, and ZIP dummies.
* $P < .05$, ** $P < .01$, *** $P < .001$.

double or more, the impact of the marginal retail establishment). Together, these findings indicate that the mix of nearby retail plays an important role in the valuation of property, especially for bigger catchment areas.

## 5.3 Identifying localized retail services

As a first attempt to ensure we are identifying localized retail services, we distinguish among retail services and goods that should be more or less frequently consumed. We expect that frequently consumed goods and services will be more desirable in proximity than those less frequently consumed. These results are displayed in Fig. 2a and b. Across all property types, frequently consumed retail goods and services are more often than not positively capitalized in prices. Frequently consumed retail is associated with significantly higher prices mostly for larger radii, with the most precisely estimated changes at 1/2 mile for condo/co-op and large multi-family properties. The fact that the sign of the coefficients, regardless of significance, flips for infrequently consumed goods and services (see Fig. 2b) suggests that the price capitalization is in line with expectations—that prices will increase with more frequently consumed services and goods nearby.[11]

Next, we differentiate between chain and independently owned establishments. This classification, while admittedly crude, attempts to capture differences in the size, selection and, perhaps, aesthetic across the retail establishments.[12] We imagine that chains might have more selection and lower prices, but maybe less character in their street presence and more standardization in their products. In addition, chain establishments likely serve broader markets and are therefore not as locally capitalized as other independent or "boutique" establishments. These results are displayed in Fig. 3a and b. Across all property types, proximity to chain establishments almost uniformly is associated with lower prices, with the biggest price change at smaller radii. The magnitudes of the price changes are also larger than what we have seen before, by an order of magnitude (closer to 2 or 3 per cent price decline in the presence of an additional ten chain establishments). The pattern of price changes associated with nearby independent establishments is less clear. This is not unexpected, as that group of establishments is bigger and more heterogeneous.

## 5.4 Type of retail

Finally, we test for differences in price capitalization across more specific categories of retail to understand better any tradeoffs between local amenity values and nuisance discounts. These results are displayed in Fig. 4a–d. First, we isolate restaurants, which may be considered an amenity but also can produce noise and crowding and other nuisances. The results reflect these mixed effects. For 1–4 family and larger rental buildings restaurants are associated with higher prices nearby; prices for condos/coops near restaurants are relatively lower. The negative price capitalization for condos/coops could be driven by the fact that the concentration of restaurants tends to be about 30 per cent higher around condo/coops than either 1–4 family or larger multi-family buildings (see Supplementary Appendix E). Therefore, the nuisances could outweigh any amenity benefits. In addition, we cannot observe the nature or quality of services—therefore, the differences in capitalization could also be picking up differences in the spatial clustering of those features as well.

Proximity to personal services is also valued differently across properties. They are negatively capitalized into 1–4 family prices, and positively capitalized into condo/coop and multi-family properties (albeit all at farther distances). Again, this difference across property types is likely driven by the difference in concentrations of personal service establishments (which are higher near 1–4 family properties) and also the kinds of personal services that tend to operate in the distinct neighborhoods where the property types cluster. Some, depending on their storefront aesthetic or quality of service, may be

[11] As an additional test, we distinguish between retail services that are more substitutable with online e-commerce and those that are more viable in-person. This distinction is important not only for isolating neighborhood-based services from those that can be consumed at a distance (see Relihan (2022) for a related study), but it also addresses any concern of contemporaneous e-commerce proliferation during our study period. When we plot over time the sub-categories of retail based on their online substitutability, we do see a steeper increase in the presence of services more likely replaced by e-commerce compared to others deemed less substitutable (see Supplementary Appendix O). The findings from these additional regressions (see Supplementary Appendix P) confirm that services deemed less substitutable with online commerce are more positively capitalized across all property types.

[12] As an alternative, we run regressions controlling for retail and the interaction between retail and the average establishment size in the circular area. The results do not introduce any new information about the mediating role of size and confirm the main findings related to price capitalization of retail establishments nearby. Results are available from the authors upon request.

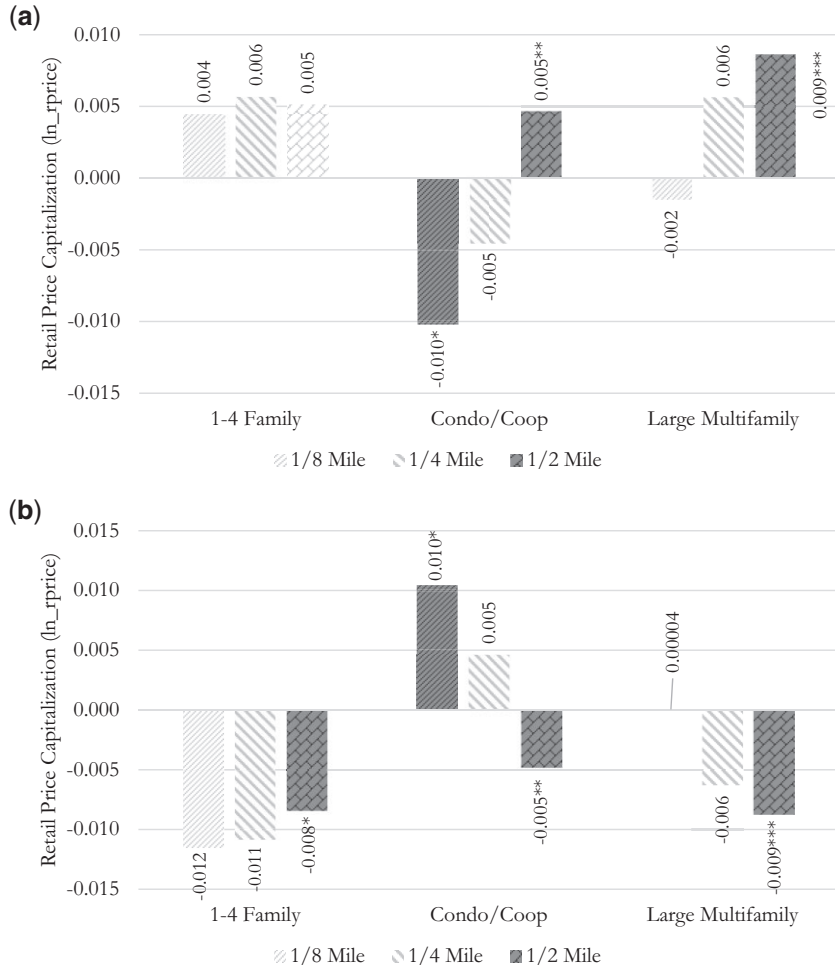


**Figure 2.** (a) OLS, the association between proximity to frequently consumed retail establishments and residential prices. *Note:* This figure shows the association between the proximity to frequently consumed retail establishments and transacted property prices. Bars for each property type show the effect of frequently consumed retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is frequently consumed retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (b) OLS, the association between proximity to infrequently consumed retail establishments and residential prices. *Note:* This figure shows the association between the proximity to infrequently consumed retail establishments and transacted property prices. Bars for each property type show the effect of infrequently consumed retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is infrequently consumed retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$.

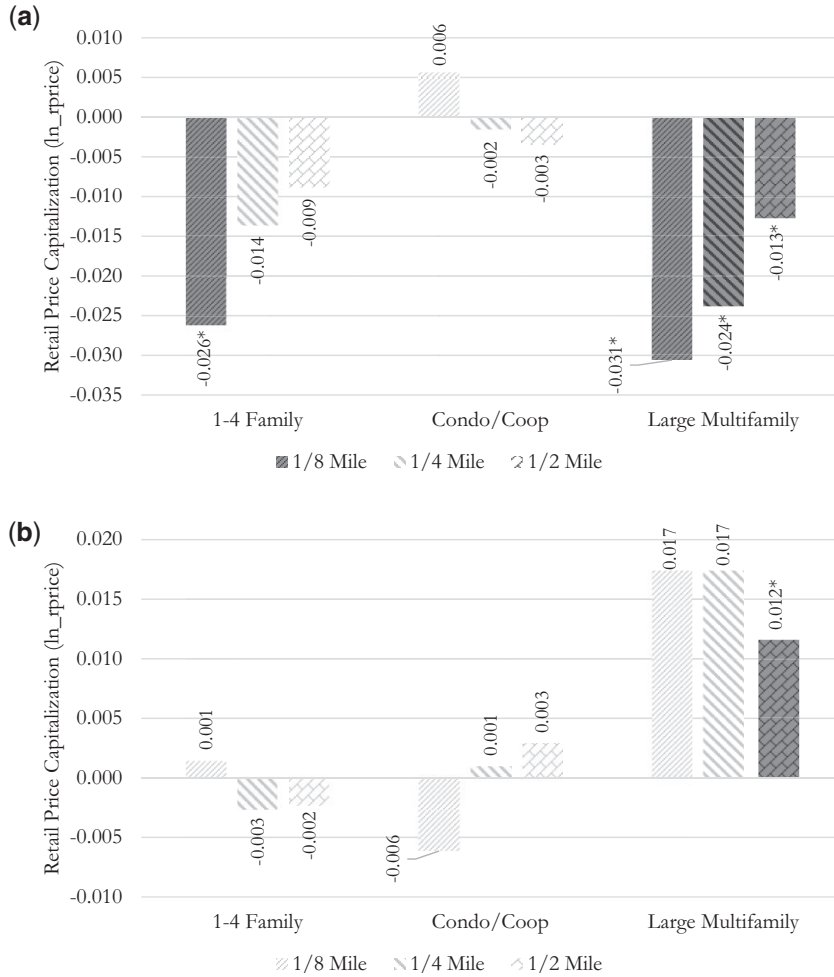


**Figure 3.** (a) OLS, the association between proximity to chain establishments and residential prices. *Note:* This figure shows the association between the proximity to chain retail establishments and transacted property prices. Bars for each property type show the effect of chain retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is chain retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (b) OLS, the association between proximity to independent establishments and residential prices. *Note:* This figure shows the association between the proximity to independent retail establishments and transacted property prices. Bars for each property type show the effect of independent retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is independent retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$.

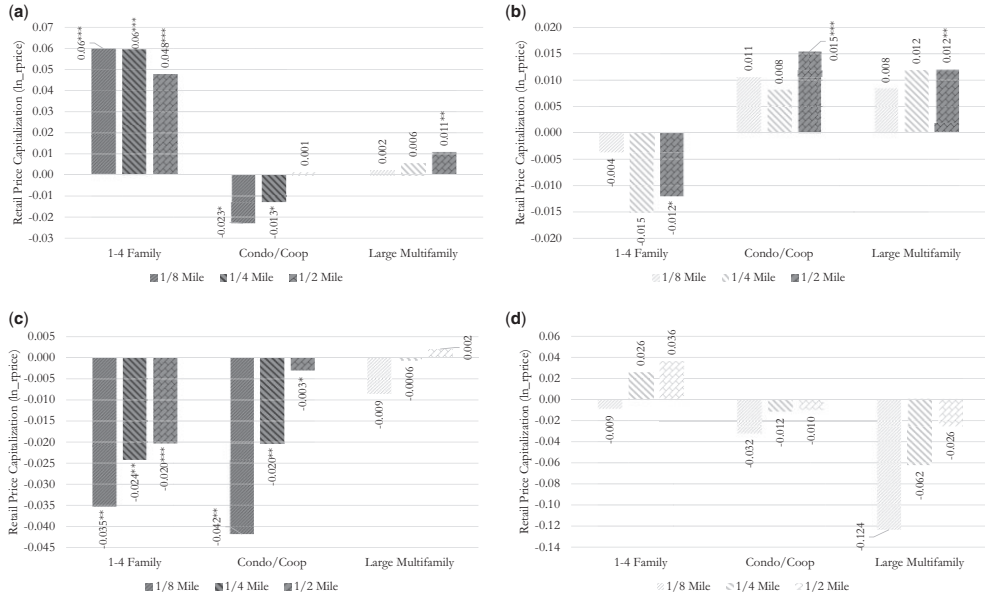


**Figure 4.** (a) OLS, the association between proximity to restaurants and residential prices. *Note:* This figure shows the association between the proximity to restaurants and transacted property prices. Bars for each property type show the effect of restaurants in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is restaurant establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (b) OLS, the association between proximity to personal services and residential prices. *Note*: This figure shows the association between the proximity to personal services and transacted property prices. Bars for each property type show the effect of personal services in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is personal service establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (c) OLS, the association between proximity to food and beverage and residential prices. *Note:* This figure shows the association between the proximity to food and beverage establishments and transacted property prices. Bars for each property type show the effect of food and beverage establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is food and beverage establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (d) OLS, association between proximity to financial services and residential prices. *Note:* This figure shows the association between the proximity to financial services and transacted property prices. Bars for each property type show the effect of financial services in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is financial service establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$.

considered an amenity while others may induce more nuisance effects. Price responses to nearby establishments selling other goods, like clothing and home furnishings, are similarly varied across property types, reflecting the heterogeneity in the retail category and the idiosyncrasies of where and how they cluster.

More consistent are the results for food and beverage establishments. This kind of retail is associated with lower prices across all property types, especially at the closest distances. And this is true despite higher concentrations around 1–4 family and larger multi-family properties. While the negative price capitalization is surprising, given that this category includes supermarkets (an important neighborhood amenity), it also includes other food and beverage purveyors, like convenience stores and liquor stores, which can be more plentiful and perhaps less attractive. Financial services are also uniformly unrelated to nearby prices. However, the number of establishments in this category is also much smaller.

Altogether these regressions indicate that retail capitalization is significant across most property types and that it declines over distance. The nature of the capitalization, however, varies by property type. This variation is particularly evident when we break down the retail services into classifications that not only have different amenity-nuisance tradeoffs but are also clustered differently across space. However, we consistently find that the retail services and goods that are likely to be locally consumed or experiential are positively capitalized into property values and this result persists regardless of property type.

## 5.5 2SLS results

We now turn to the results where we instrument for retail activity to address concerns about endogeneity. First, we acknowledge that the first-stage results (see Supplementary Appendix Q) for both instruments are reassuring. Both produce significant coefficients for the instrument itself and in most cases *F*-statistics above ten (the first instrument performs better on this front, producing *F*-statistics above thirty in most cases).

The two instruments produce results that overall reinforce the patterns observed in the OLS models, but with more precisely estimated price gradients. The first retail shock instrument (displayed in Table 4) produces positive price effects for 1–4 family properties, which are strongest at close distances and decline over space. The price effects are much larger than the OLS ones, such that at closest distances an additional ten retail establishments can increase prices by about 22 per cent (approximately $130,000) and closer to 4 per cent (approximately $24,000) at farther distances. Adjusted for the mean number of retail establishments within a particular radius, this premium ranges from $468,000 to 658,000 for retail access. Proximity to retail (an additional ten establishments), on the other hand, pushes prices down for condo/coops, with the effects at about 3 per cent (or $31,000) at close distances and declining by an order of magnitude at a 1/2-mile distance to about 0.4 per cent (about $3,000). This translates to a premium between $430,000 to $626,000 (for the entire condo/coop building) for retail access nearby. While the coefficients on the retail variables are similarly negative and declining in magnitude over space for large multi-family properties, none are significant. We interpret the bigger coefficients as evidence of mitigated attenuation bias in the OLS estimates. It suggests that the demand channels, such as housing prices themselves or other neighborhood correlates, may dampen retail entry and activity (perhaps through prohibitive rents or other demographic signals that mediate retail entry and activity). By isolating the supply channel with the instruments, we mitigate any downward bias on the *Retail* coefficients.

Also of note, but not displayed here, is the fact that, when standardized, the coefficients on the retail indicators are similar or bigger in magnitude than any other hedonic control in the model (this varies by property type). The two exceptions are building age and years since an alteration, which are consistently much larger (especially for the condo/coop sample).

The second instrument, based on industrial competition and complementarity, produces similar patterns, albeit less precisely and with magnitudes smaller than the first instrument. The results are displayed in Table 5. These magnitudes are more in line with those produced by the OLS regressions, however, the first-stage diagnostics are weaker for this instrument. We interpret the range of estimates as bounds, whereby the first instrument provides an upper price effect and the OLS and second instrument are closer to the lower bound. Again, proximity to retail induces positive price effects for

**Table 4.** 2SLS, association between proximity to retail establishments and residential prices, retail shock (shift-share) IV

| DV = log(price) | 1–4 Family | | | Condos and Coops | | | Large multifamily | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Retail Estab, 1/8 M | 0.223*** | | | −0.029*** | | | −0.005 | | |
| | (0.030) | | | (0.008) | | | (0.017) | | |
| Retail Estab, 1/4 M | | 0.092*** | | | −0.012*** | | | −0.004 | |
| | | (0.013) | | | (0.003) | | | (0.007) | |
| Retail Estab, 1/2 M | | | 0.036*** | | | −0.004*** | | | −0.002 |
| | | | (0.004) | | | (0.001) | | | (0.003) |
| First-stage F-Stat | 110.80 | 106.20 | 112.11 | 66.84 | 72.18 | 84.29 | 56.15 | 52.21 | 52.49 |
| $R^2$ | 0.01 | 0.04 | 0.06 | 0.62 | 0.63 | 0.64 | 0.61 | 0.61 | 0.61 |
| Standard error | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster |
| Hedonics? | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| N | 261,886 | 261,886 | 261,886 | 108,813 | 108,813 | 108,813 | 14,826 | 14,826 | 14,826 |

*Note:* This table shows the relationship between the proximity to retail establishments and transacted property prices, instrumenting for retail activity with a shift-share metric that interacts national growth in retail establishments with the share of commercial square footage (exposure) in the nearby circular area. Columns (1)–(3) for each property type show the effect of retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial establishments in that same area. Standard errors (in parentheses) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of total establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, and ZIP dummies.
* $P < .05$, ** $P < .01$, *** $P < .001$.

**Table 5.** 2SLS, the association between proximity to retail establishments, and residential prices, complementarity/cannibalization IV.

| | 1–4 Family | | | Condos and Coops | | | Large multifamily | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Retail Estab, 1/8 M | 0.008 | | | −0.009* | | | −0.001 | | |
| | (0.005) | | | (0.004) | | | (0.003) | | |
| Retail Estab, 1/4 M | | 0.019*** | | | −0.006** | | | 0.00006 | |
| | | (0.004) | | | (0.002) | | | (0.00003) | |
| Retail Estab, 1/2 M | | | 0.027*** | | | −0.002 | | | 0.003 |
| | | | (0.008) | | | (0.001) | | | (0.002) |
| First-stage F-Stat | 15.85 | 54.30 | 54.87 | 75.47 | 86.02 | 34.81 | 53.54 | 59.58 | 37.58 |
| $R^2$ | 0.09 | 0.09 | 0.08 | 0.64 | 0.64 | 0.64 | 0.61 | 0.61 | 0.61 |
| Standard error | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster | cluster |
| Hedonics? | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| N | 261,886 | 261,886 | 261,886 | 108,813 | 108,813 | 108,813 | 14,826 | 14,826 | 14,826 |

*Note:* This table shows the relationship between the proximity to retail establishments and transacted property prices, instrumenting for retail activity with a complementarity/cannibalization metric that interacts predicted change in retail establishments with the share of commercial square footage (exposure) in the nearby circular area. Columns (1)–(3) for each property type show the effect of retail in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial establishments in that same area. Standard errors (in parentheses) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of total establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, and ZIP dummies.
* $P < .05$, ** $P < .01$, *** $P < .001$.

1–4 family properties and negative ones for condo/coops. Both effects decline in magnitude as distance from the home increases.[13]

## 5.6 Heterogeneity and robustness checks

In this section, we conduct additional tests to confirm that any unobserved variation, specifically in how retail and property types are allocated across the city, is not driving the results produced thus far.[14] For brevity in exposition, we display only the 2SLS results using the first shift-share instrument (*Retail_shock*) and note where the estimates produced by the second instrument are not consistent with those presented.

First, we know that retail location decisions respond to the preferences (and characteristics) of local consumers, which are also highly correlated with the price of real estate. We stratify the sample by several neighborhood (i.e. census tract) demographic characteristics, including race (share of white households), average household income, and length of commute, to allow the price effects to vary along these dimensions. These characteristics were selected based on previous research that shows they are correlated with differences in retail preferences (e.g., Waldfogel 2008; Meltzer and Schuetz, 2012). We also know that access to home purchases and the fundamentals of local real estate markets can vary with respect to race and income, and expect that the valuation of very local amenities will vary with how tied residents are to their neighborhood (Su 2018). We use time commuting to work to proxy for the latter. The results are shown in Fig. 5a–c. We find that most differences are observed for condo/coop transactions with few or no discernible differences across strata for 1–4 family and large multi-family properties. Specifically, neighborhoods with higher shares of white households, higher average incomes, and shorter typical commutes, drive the negative price capitalization results observed for condo/coop properties. The other strata exhibit similar, but insignificant, patterns. Overall, accounting for residents' demographics sheds some light on, but does not eliminate, price response discrepancies across property types.

Second, we want to make sure we are not simply picking up the perceived benefits of living in more walkable parts of the city where there is likely more retail co-located with residential properties. We expect that the composition of property types will also be related to the mixed-use nature of the neighborhood. We create a matrix of residential and commercial co-mingling by identifying circular areas with high/low residential and retail densities. We stratify the sample in terms of pairings of "high" (above the median) and "low" (below the median) residential and commercial densities. We expect that, all else equal, retail services will be more positively capitalized into prices for areas that are more mixed use (or denser in terms of both residential and retail). The results are presented in Fig. 6a–c. Across all property types, retail is positively capitalized into property values in areas with denser residential use. Furthermore, the biggest price effects are in areas that are both residentially and commercially dense (or relatively more mixed use). There are also clear price gradients such that the positive price capitalization is biggest at small radii and attenuates over distance. These findings are consistent with the expectations that proximate retail services are more valuable in areas that are more walkable and where retail is more integrated with housing. Indeed, any negative price capitalization (for condos/coops and large multi-family properties) is concentrated among areas with less residential use and where, perhaps, commercial activity dominates the market rather than complements it.

Next, we want to make sure we are picking up the amenity value to local residents, apart from any other productive use that might be generated by the commercial real estate where the retail operates. For example, the retail activity may be more valued because of the inherent advantages of its infrastructure or location (proximity to transit, new construction, and visibility), rather than the services provided. To test this, we re-estimate the regression across strata of low, moderate, and high (non-retail) commercial property appreciation. The results are displayed in Fig. 7a–c, and they show no discernible difference across strata. Therefore, the amenity effects estimated above hold.

[13] Property fixed effects are an alternative strategy for addressing unobserved heterogeneity across properties that may be correlated with the amount and composition of local retail. We do replicate the regressions using only a sample of repeat-sales and while the magnitudes of the retail coefficients for the OLS regressions on the repeat sales sample are bigger than those based off of the full sample, any differences go away once we instrument for the retail variables. These results are displayed in Supplementary Appendix R. We rely on the full sample for our main analysis as it is more than twice as large as the repeat-sales sample and there are some concerns of selection bias when restricting to only properties that sell more than once over the study period.

[14] Following the guidance in Goldsmith-Pinkham, Sorkin, and Swift (2020) we will assess the mediating role of correlates with the commercial square footage exposure share in our instrument to rule out the possibility of other causal channels (which would undermine our identifying assumption in the 2SLS approach).

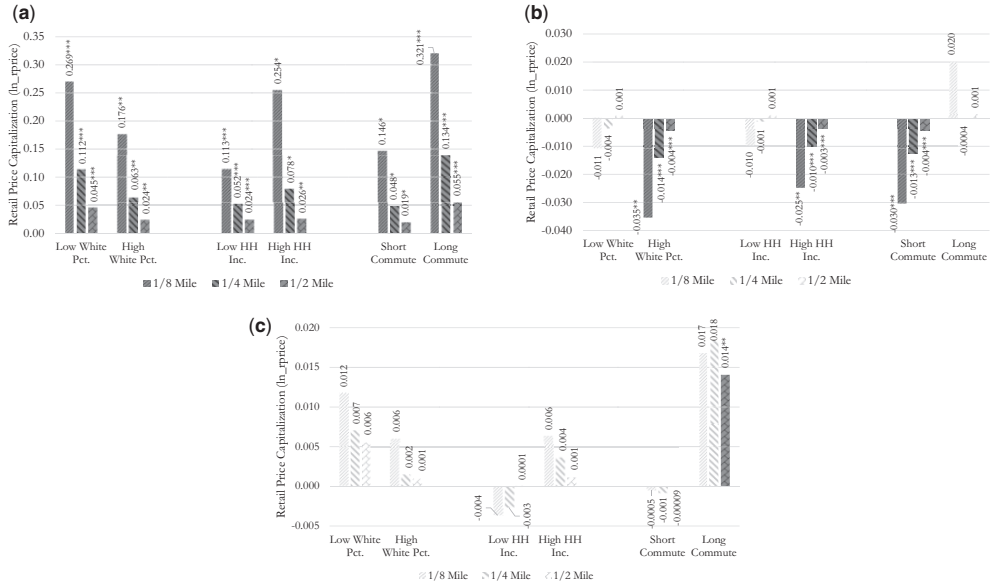


**Figure 5.** (a) 2SLS, the association between proximity to retail establishments and residential prices, stratified by select neighborhood characteristics, 1–4 family. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for 1–4 family homes, instrumenting for retail activity with a shift-share metric and stratified by selected tract characteristics. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified showing those in tracts in the top and bottom quartiles for the share of white households, household income, and commute time. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (b) 2SLS, the association between proximity to retail establishments and residential prices, stratified by select neighborhood characteristics, condo/coop. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for condo/coops, instrumenting for retail activity with a shift-share metric and stratified by selected tract characteristics. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified showing those in tracts in the top and bottom quartiles for the share of white households, household income and commute time. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (c) 2SLS, the association between proximity to retail establishments and residential prices, stratified by select neighborhood characteristics, large multifamily. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for large multifamily properties, instrumenting for retail activity with a shift-share metric and stratified by selected tract characteristics. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified showing those in tracts in the top and bottom quartiles for the share of white households, household income and commute time. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, *** $P < .001$.

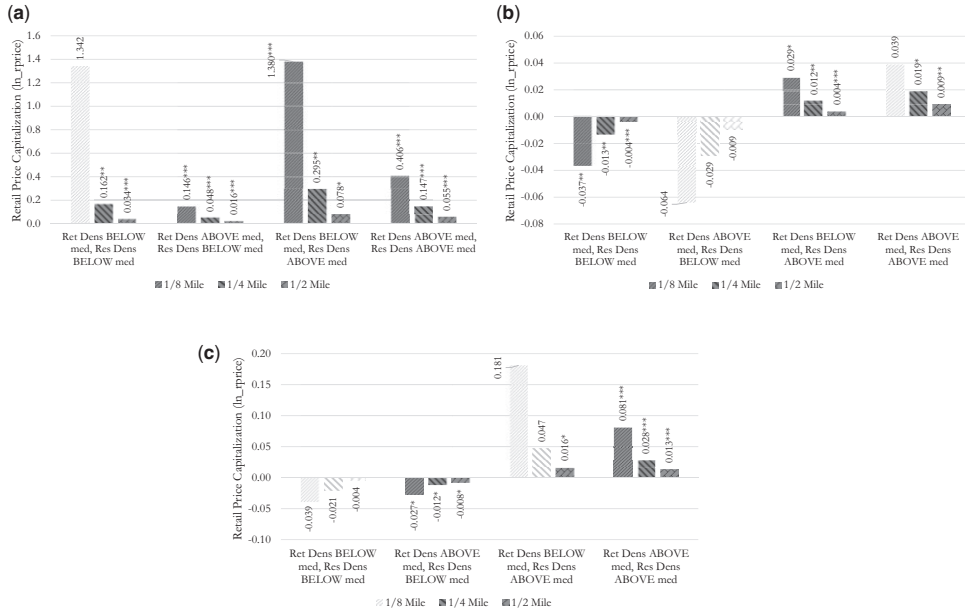


**Figure 6.** (a) 2SLS, the association between proximity to retail establishments and residential prices, stratified by retail/residential density, 1–4 family. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for 1–4 family homes, instrumenting for retail activity with a shift-share metric and stratified by the degree of mixed uses in the neighborhood. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified across the four combinations of median retail and residential densities in the nearby circular areas; areas with higher retail and residential densities are considered the most mixed use. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (b) 2SLS, the association between proximity to retail establishments and residential prices, stratified by retail/residential density, condo/coop. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for condos/coops, instrumenting for retail activity with a shift-share metric and stratified by the degree of mixed uses in the neighborhood. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified across the four combinations of median retail and residential densities in the nearby circular areas; areas with higher retail and residential densities are considered the most mixed use. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < 0.01$, ***$P < .001$. (c) 2SLS, the association between proximity to retail establishments and residential prices, stratified by retail/residential density, large multifamily. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for large multifamily properties, instrumenting for retail activity with a shift-share metric and stratified by the degree of mixed uses in the neighborhood. Bars show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified across the four combinations of median retail and residential densities in the nearby circular areas; areas with higher retail and residential densities are considered the most mixed use. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$.

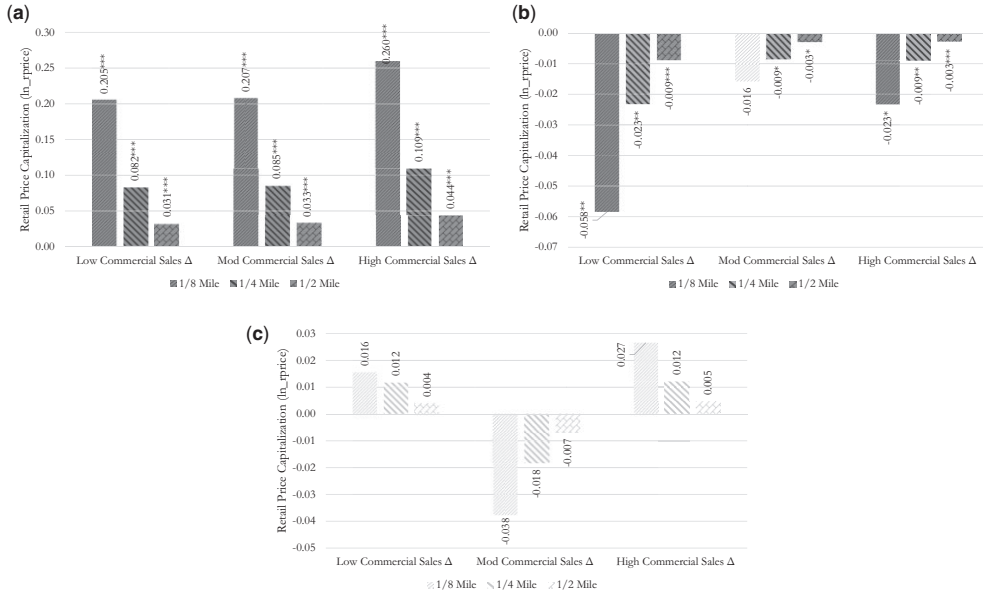


**Figure 7.** (a) 2SLS, the association between proximity to retail establishments and residential prices, stratified by commercial sales growth, 1–4 family. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for 1–4 family homes, instrumenting for retail activity with a shift-share metric and stratified by the magnitude of change in commercial sales prices over the study period. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified across the terciles of changes in commercial prices. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < 0.01$, ***$P < .001$. (b) 2SLS, the association between proximity to retail establishments and residential prices, stratified by commercial sales growth, condo/coop. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for condos/coops, instrumenting for retail activity with a shift-share metric and stratified by the magnitude of change in commercial sales prices over the study period. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified across the terciles of changes in commercial prices. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$. (c) 2SLS, the association between proximity to retail establishments and residential prices, stratified by commercial sales growth, large multifamily. *Note:* This figure shows the relationship between the proximity to retail establishments and transacted property prices for large multifamily properties, instrumenting for retail activity with a shift-share metric and stratified by the magnitude of change in commercial sales prices over the study period. Bars for each property type show the effect of retail establishments in nearby circular areas, of varying distances from the block of the property transaction between 1/8 and 1/2 mile, controlling for the total number of commercial and retail establishments in that same area. The sample of transactions is stratified across the terciles of changes in commercial prices. Shaded bars are estimates significant at least at $P \leq .05$. Standard errors (data not shown) are clustered by ZIP. The dependent variable is the log(real price); the independent variable of interest is retail establishments (counted in 10 s of establishments), which are a subset of all retail establishments (counted in 10 s of establishments). Hedonics and control variables included, but not shown are: property gross square footage (logged), number of units (logged), lot frontage, number of floors, age and age-squared in 2010, years and years-squared in 2010 since last alteration, elevator building, historic designation, borough–year dummies, ZIP dummies. *$P < .05$, **$P < .01$, ***$P < .001$.

Finally, since we identify off of changes in prices (and retail activity), we want to be sure that the estimates are not simply picking up broader macro-economic trends that happen to manifest themselves differently at the local community level during the study period. For example, during the study period some, but not all, neighborhoods experienced gentrification, or extreme price appreciation relative to the city more broadly. To mitigate against any interference from differential price appreciation across neighborhoods—that may or may not be related to retail concentration; see Meltzer (2016) or Glaeser, Luca, and Moszkowski (2020)—we stratify the sample by degrees of housing price appreciation. The results are displayed in figures in Supplementary Appendix S. The patterns of coefficients are not largely different across the price growth strata, while the effects for moderate and higher price growth areas are marginally significant for condos/coops and multifamily properties. Overall, however, any evidence of retail capitalization associated with variable housing price growth is weak.[15]

## 6. Conclusion

Urban neighborhoods are inherently mixed use. This suggests that households who live in these neighborhoods likely value the proximity of private retail amenities. That said, we know very little about how much households are willing to pay to have nearby access to these services. We exploit a detailed and rich dataset on retail activity and property values to estimate the price capitalization of nearby retail services and to understand more broadly how local retail services might contribute to neighborhood quality of life.

Results indicate that retail services are indeed capitalized into housing prices. The direction of the capitalization, however, depends on the nature of the service and the conditions of the local neighborhood. We find that after isolating localized retail services, such as those that are more frequently patronized or experiential, we consistently observe positive price capitalization across all property types. In addition, even at small scales, there are clear price gradients where the positive capitalization attenuates over distance. Most notably, we also find that the price gradients sharpen when we control for the endogeneity of retail amenities in the housing market and when we focus on mixed-use communities, where there is systematic co-mingling of residential and commercial properties. Indeed, in these 2SLS specifications, for the most mixed-use settings, we observe a price premium from an additional ten retail establishments between 6 and 40 per cent for 1–4 family properties, closer to 1 or 2 per cent for coops/condos and between 1 and 8 per cent for larger multifamily properties. For a typical sale in mixed-use settings in the city, this translates into between $1,500 and $9,200 in added value from a retail establishment within 1/4 mile (or about 6 city blocks). When scaled up to reflect multiple establishments and services nearby, this means properties can sell at six- or seven-figure premia in mixed-use settings with more retail services.[16] Furthermore, standardized coefficients indicate that the relative magnitude of the price effect from nearby retail is at least as big as property-level features (with the exception of property age in some cases).

Therefore, even in dense urban settings, the value of nearby retail amenities varies over a small spatial scale. Again, it is most positive when the services are frequently consumed and experiential and when there is an expectation for land use mixing embedded into that local market. On the other hand, some evidence suggests that retailers without obvious neighborhood benefits, like chains or banks, are either not valued at all or push prices down at the margin. Services that can bring nuisances, like noise or traffic, can have mixed price effects, especially if they are plentiful in a small geography.

The findings from this research shed light on the intra-urban variation in how retail amenities are valued. This information has particular value in light of COVID-19 and its disproportionate impact on storefront retail. Cities will have to grapple with how to rebuild the urban retail landscape and make decisions about the kinds of services that contribute to neighborhood quality of life and the broader

[15] The Great Recession occurred in the middle of our study period. Both housing prices and retail activity declined rapidly (and then gradually recovered) after the Recession, following a markedly different trend than in earlier years. To account for this, we stratify the sample across pre- and post-recession periods. These results are displayed in figures in Supplementary Appendix T. We find that price capitalization is positive and largely driven by post-Recession (2006) transactions.[15] This is consistent with the more recent amenity-driven urban resurgence. Any negative capitalization among condos/coops was a pre-Recession phenomenon, and only significantly present for condo/coops.

[16] These estimates are based on mean prices and establishment counts for areas of the city with higher retail and residential densities. We do not display these statistics here, but prices tend to be slightly higher for 1–4 family properties and lower for the other property types compared to the citywide sample; establishment counts are higher by definition.

municipal fiscal livelihood. Thus far, the results suggest that this is not a one-size-fits-all consideration as the value of retail amenities varies with very localized neighborhood and industry features. However, in areas that are inherently mixed use and for services that are convenient to have nearby and frequently consumed, proximate retail services are on net an amenity.

## Supplementary data

Supplementary data is available at *Journal of Economic Geography* online.

## Conflict of interest statement

Pooya Ghorbani is employed by New York State Homes and Community Renewal (HCR). However, the analyses and opinions in this paper are his own and do not reflect HCR's position on any of the discussed matters.

## References

Albouy, D. (2016) 'What are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities', *Review of Economics and Statistics*, **98**: 477–87.

Asabere, P.K., and Huffman, F.E. (1996) 'Negative and Positive Impacts of Golf Course Proximity on Home Prices', *The Appraisal Journal*, **LXIV**: 351–355.

Barnatchez, K., Crane, L.D., and Decker, R. (2017) *An Assessment of the National Establishment Time Series (NETS) Database*. Finance and Economics Discussion Series 2017-110. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.110.

Bartik, T.J. (1991) *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Berry, B. (1967) *Geography of Market Centers and Retail Distribution*. Englewood Cliffs, NJ: Prentice Hall.

Benson, E.D. et al. (1998) 'Pricing Residential Amenities: The Value of a View', *Journal of Real Estate Finance and Economics*, **16**: 55–73.

Beracha, E. et al. (2018) 'On the Relation between Local Amenities and House Price Dynamics', *Real Estate Economics*, **46**: 612–54.

Bingham, R. D., and Zhang, Z. (1997) 'Poverty and Economic Morphology of Ohio Central-City Neighborhoods', *Urban Affairs Review*, **32**: 766–96.

Black, S.E. (1999) 'Do Better Schools Matter? Parental Valuation of Elementary Education', *The Quarterly Journal of Economics*, **114**: 577–99.

Borraz, F. et al. (2020) *Local Retail Prices, Product Varieties and Neighborhood Change*. CEPR Discussion Paper No. DP16371. https://ssrn.com/abstract=3928649.

Borusyak, K., Hull, P., and Jaravel, X. (2022) 'Quasi-experimental Shift-share Research Designs', *The Review of Economic Studies*, **89**: 181–213.

Brueckner, J.K., Thisse, J., and Zenou, Y. (1999) 'Why is Central Paris Rich and Downtown Detroit Poor: An Amenity-based Theory', *European Economic Review*, **43**: 91–107.

Cao, T.V., and Cory, D.C. (1981) 'Mixed Land Uses, Land-use Externalities, and Residential Property Values: A Reevaluation', *Annals of Regional Science*, **16**: 1–24.

Carlino, G. A., and Saiz, A. (2008) *Beautiful City: Leisure Amenities and Urban Growth*. FRB of Philadelphia Working Paper No. 08-22, https://ssrn.com/abstract=1280157.

Chen, Y., and Rosenthal, S. S. (2008) 'Local Amenities and Life-cycle Migration: Do People Move for Jobs or Fun', *Journal of Urban Economics*, **64**: 519–37.

Chiang, Y., Peng, T., and Chang, C. (2015) 'The Nonlinear Effect of Convenience Stores on Residential Property Prices: A Case Study of Taipei, Taiwan', *Habitat International*, **46**: 82–90.

Corlija, M., Siman, E. and Finke, M.S. (2006) 'Longitudinal Analysis of Big Box Store Construction on Nearby Home Values', *Consumer Interests Annual*, **52**: 187–97.

Couture, V. (2016) *Valuing the Consumption Benefits of Urban Density*. University of Toronto, Working Paper.

Couture, V., and Handbury, J., (2020) 'Urban Revival in America', *Journal of Urban Economics*, **119**: 103267.

Cox, D. A., and Streeter, R. (2019), *The Importance of Place: Neighborhoods Amenities as a Source of Social Connection and Trust*. Washington, DC: American Enterprise Institute.

Crane, L. D., and Decker, R. (2019). *Business Dynamics in the National Establishment Time Series (NETS)*. Federal Reserve Board Working Paper.

Cutler, D., Glaeser, E., and Vigdor, J. (1999) 'The Rise and Decline of the American Ghetto', *Journal of Political Economy*, **107**: 455–506.

Daunfeldt, S.O. et al. (2020) 'Retail and Place Attractiveness: The Effects of Big-Box Entry on Property Values', *Geographical Analysis*, **53**: 467–98.

Davis, D.R. et al. (2019) 'How Segregated is Urban Consumption', *Journal of Political Economy*, **127**: 1684–738.

DiPasquale, D., and Wheaton, W.C. (1996) *Urban Economics and Real Estate Markets*. Englewood Cliffs, NJ: Prentice Hall.

Ellis, C.D., Lee, S.W., and Kweon, B.S. (2006) 'Retail Land Use, Neighborhood Satisfaction and the Urban Forest: An Investigation into the Moderating and Mediating Effects of Trees and Shrubs', *Landscape and Urban Planning*, **74**: 70–78.

Glaeser, E.L., Luca, M. and Moszkowski, E. (2020), *Gentrification and Neighborhood Change: Evidence from Yelp (No. w28271)*. Cambridge, MA: National Bureau of Economic Research.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020) 'Bartik Instruments: What, When, Why, and How', *American Economic Review*, **110**: 2586–624.

Grudnitski, G., and Do, A.Q. (1997) 'Adjusting the Value of Houses Located on a Golf Course', *The Appraisal Journal*, **65**: 261–6.

Helling, A., and Sawicki, D. S. (2003) 'Race and Residential Accessibility to Shopping and Services', *Housing Policy Debate*, **14**: 69–101.

Hotelling, H. (1929) 'Stability in Competition', *Economic Journal*, **39**: 41–57.

Jacobs, J. (1961) *The Death and Life of Great American Cities*. New York, NY: Random House.

Jang, M., and Kang, C. (2015) 'Retail Accessibility and Proximity Effects on Housing Prices in Seoul, Korea: A Retail Type and Housing Submarket Approach', *Habitat International*, **49**: 516–28.

Jargowsky, P. (2003) *Poverty and Place: Ghettoes, Barrios and the American City*. New York, NY: Russell Sage.

Jim, C.Y., and Chen, W.Y. (2006) 'Impacts of Urban Environmental Elements on Residential Housing Prices in Guangzhou (China)', *Landscape and Urban Planning*, **78**: 422–34.

Johnson, D.K., and née Lybecker, K.A. (2018) 'Is Wal-Mart a Bad Neighbor? Repeat Sales Evidence on How Residential Property Values React to a New Big-Box Store', *Open Journal of Social Sciences*, **6**: 15–32.

Koster, H. RA., and Rouwendal, J. (2012) 'The Impact of Mixed Land Use on Residential Property Values', *Journal of Regional Science*, **52**: 733–61.

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

Lynch, A.K., and Rasmussen, D.W. (2001) 'Measuring the Impact of Crime on House Prices', *Applied Economics*, **33**: 1981–9.

Massey, D., and Denton, N. (1993) *American Apartheid*. Cambridge MA: Harvard University Press.

Meltzer, R. (2016) 'Gentrification and Small Business: Threat or Opportunity', *Cityscape*, **18**: 57–86.

Meltzer, R., and Capperis, S. (2017) 'Neighbourhood Differences in Retail Turnover: Evidence from New York City', *Urban Studies*, **54**: 3022–57.

Meltzer, R., and Schuetz, J. (2012) 'Bodegas or Bagel Shops? Neighborhood Differences in Retail & Household Services', *Economic Development Quarterly*, **26**: 73–94.

Mian, A. and Sufi, A. (2011) 'House Prices, Home Equity–Based Borrowing, and the US Household leverage Crisis', *American Economic Review*, **101**: 2132–56.

Nelson, R. L. (1958) *The Selection of Retail Locations*. St. Louis, MO: FW Dodge Corporation.

Neumark, D., Zhang, J., and Wall, B. (2005) *Employment Dynamics and Business Relocation: New Evidence from the National Establishment Time Series*. NBER Working Paper w11647.

Rappaport, J. (2008) 'Consumption Amenities and City Population Density', *Regional Science and Urban Economics*, **38**: 533–52.

Rauterkus, S. Y., and Miller, N. G. (2011) 'Residential Land Values and Walkability', *The Journal of Sustainable Real Estate*, **3**: 23–43.

Reilly, W.J. (1931) *The Law of Retail Gravitation*. New York, NY, USA: WJ Reilly.

Relihan, L. (2022) *Is Online Retail Killing Coffee Shops? Estimating the Winners and Losers of Online Retail Using Customer Transaction Microdata*. London School of Economics CEP Discussion Paper No. 1836.

Stanback, T., Bearse, P., Noyelle, T., and Karasek, R. (1981) *Services: The New Economy*. Totowa, NJ: Allenheld.

Shimizu, C. et al. (2014) *Do Urban Amenities Drive Housing Rent?* Institute of Economic Research, Hitotsubashi University, Working Paper No. 9.

Sirpal, R. (1994) 'Empirical Modeling of the Relative Impacts of Various Sizes of Shopping Centers on the Values of Surrounding Residential Properties', *Journal of Real Estate Research*, **9**: 487–505.

Song, Y., and Knaap, G.J. (2003) 'New Urbanism and Housing Values: A Disaggregate Assessment', *Journal of Urban Economics*, **54**: 218– 38.

Song, Y., and Sohn, J. (2007) 'Valuing Spatial Accessibility to Retailing: A Case Study of the Single Family Housing Market in Hillsboro, Oregon', *Journal of Retailing and Consumer Services*, **14**: 279–88.

Stroebel, J., and Vavra, J. (2019) 'House Prices, Local Demand, and Retail Prices', *Journal of Political Economy*, **127**: 1391–436.

Su, Y. (2018) *The Rising Value of Time and the Origin of Urban Gentrification*. https://ssrn.com/abstract=3216013.

Taylor, R. B. (1995) 'The Impact of Crime on Communities', *The Annals of the American Academy of Political and Social Science*, **539**: 28–45.

Thaler, R. (1978) 'A Note on the Value of Crime Control: Evidence from the Property Market', *Journal of Urban Economics*, **5**: 137–45.

Waldfogel, J. (2008) 'The Median Voter and the Median Consumer: Local Private Goods and Population Composition', *Journal of Urban Economics*, **63**: 567–82.

Wilson, W. J. (1987) *The Truly Disadvantaged*. Chicago: University of Chicago Press.

Wu, J.J., Adams, R.M., and Plantinga, A.J. (2004) 'Amenities in an Urban Equilibrium Model: Residential Development in Portland, Oregon', *Land Economics*, **80**: 19–32.