Is the 'Shop Around the Corner' a Luxury or a Nuisance? The relationship between income and neighborhood retail patterns

Jenny Schuetz
University of Southern California
jschuetz@sppd.usc.edu

Jed Kolko
Public Policy Institute of California
kolko@ppic.org

Rachel Meltzer
Milano The New School for Management & Urban Policy
meltzerr@newschool.edu

Last revised:
September 22, 2010

Abstract

Affluent neighborhoods present a potentially attractive location for retail establishments because of their higher purchasing power and demand for a wide range of specialized goods and services. However, if high income households perceive retail in general – or certain types of retail, such as Big Box stores – as an undesirable use, they may be able to block commercial development through zoning and the political process. In this paper we shed light on these issues by examining the relationship between neighborhood income and several different types of retail presence for 58 large U.S metropolitan areas. We combine detailed data from the National Establishment Time-Series database on retail establishments and employment, by industry category and firm type, with Census data on ZCTA income and demographics. Results indicate that retail density varies with income for certain retail types, such as food service and chain supermarkets and drugstores. In addition, average establishment size increases with income for all retail types. Retail density increases with population density, as expected, and decreases with distance to CBD and with share of owner-occupied housing.

Keywords: Retail employment, firm location, commercial land use, neighborhood income, chain stores, neighborhood amenities

The authors thank PPIC for use of the NETS data and NYU's Furman Center for data support. Schuetz acknowledges financial support from USC's initiative for the Advancement of Scholarship in the Humanities and Social Sciences. Kolko thanks the David A. Coulter Family Foundation for support. Many thanks to Stephaney Kipple and Gabriela Gutierrez for excellent research assistance. All remaining errors and omissions are our own.

I) Introduction

The names of some affluent urban locations are virtually synonymous with upscale shopping and entertainment: for instance, Manhattan's Fifth Avenue, Chicago's Magnificent Mile and Beverly Hills' Rodeo Drive. On the opposite end of the income spectrum, poor urban neighborhoods are often referred to as "food deserts" with few grocery stores and only fast food restaurants (see, for instance, Moore 2010, Osen 2010, Powell et al 2007, Shaffer and Gottlieb 2007, Sloane et al 2005). According to popular media accounts and a few academic studies, one of the most visible signs of a shift in a neighborhood's income or demographics is the arrival of upscale eateries and "boutique" shopping venues (Bruni 2010, Zukin 2010). Collectively, these anecdotes suggest that retail establishments flock to affluent neighborhoods and avoid poor ones.[1] However, high income neighborhoods have sometimes had contentious relationships with developers proposing retail uses, particularly large chain stores (see, for instance, Beaumont 1997, Mitchell 2006, Scroop 2008). In short, although high income neighborhoods may represent desirable locations for retailers, such neighborhoods may also provide a less hospitable development environment than low-income neighborhoods, which may welcome retail as a source of jobs as well as goods and services. To date there has been little empirical research on how neighborhood income (and related characteristics) affects the location of retail establishments within urban areas. In this paper, we take a first step beyond anecdotes to look systematically at the relationship between income and local retail markets.

For our analysis, we combine ZCTA (ZIP Code Tabulation Area) level employment data on retail establishments, by industry category, firm structure and size, from the National Establishment Time-Series (NETS) database, with Census data from the Neighborhood Change Database on household incomes and other characteristics for 58 large metropolitan areas across the United States. We estimate both cross-sectional regressions between retail employment density and income and dynamic models of the change in employment against levels and changes in neighborhood income. Results suggest that employment density for retail as a whole and some categories, such as food service (restaurants and other eating or drinking places), do increase with neighborhood income. Total retail employment in other categories, namely supermarkets, drugstores and clothing, do not vary systematically with income. However, the

[1] According to industry classifications, food service is a separate industry from retail (NAICS codes 72 and 44-45, respectively). However, in this paper we include food and beverage services in our general discussion of retail.

aggregate relationships conceal important differences by firm structure and size: for both supermarkets and drugstores, neighborhood income is negatively related to employment in independent establishments and positively associated with employment in chain establishments. Across all categories, higher neighborhood income is associated with larger establishment size. There is some evidence that income levels and changes are positively associated with retail employment growth, although these results are less robust and consistent.

The paper proceeds in the following way. The following section sets up the theoretical framework for the analysis and Section III summarizes the relevant empirical literature. Section IV describes the data and our empirical strategy, Section V discusses the results, and Section VI concludes and offers directions for future research.

II) Theoretical Framework

In this section we develop a framework for considering the location of retail presence within metropolitan areas and the relationship between retail location decisions and underlying neighborhood characteristics. We first consider the determinants of retail market size and density and use these determinants to identify the types of services and goods that serve the local neighborhood. We then consider how the characteristics of the local neighborhood, in particular income, are likely to affect the size, density and composition of local retail markets. Note that most theoretical models – and conventional wisdom in the development industry – assume that retail establishments follow movements of households, rather than households sorting based on existing retail. Our data and empirical strategy do not allow us to determine the direction of causality but we proceed from that general assumption.

What affects density of retail?

Hotelling (1929) first described a simple spatial model of firm location in a linear city (later modified to accommodate a circular city). This model suggests that the density of stores depends on (1) customer density, (2) transportation costs, (3) frequency of purchases, and (4) store fixed costs. One clear implication from the spatial location model is that there will be different market sizes, and thus different densities of store networks for various product types. In the case of local retail services, potential customers can be categorized into two groups: (i) local residents and (ii) employees at local firms.[2] Therefore, retail store networks will be denser in neighborhoods with higher residential and employment densities. Spatially, this suggests that retail networks will be denser closer to the central business district (CBD), where employment density (and often residential density as well) is high, and in neighborhoods with thicker residential development.[3]

Retail stores will locate more densely and closer to their customers if the retailer has low fixed costs and sells goods that are highly standardized, frequently consumed or involve high transport costs due to perishability or other reasons, so that consumers will not be willing to travel long distances to purchase them (Berry 1967). In the classic example, the market area for ice cream vendors will be very small, due to the highly perishable nature of the good, so in equilibrium there will be a large number of vendors each with a small market area. On the other end of the spectrum, consumers should be willing to travel long distances to purchase goods and services that are expensive and infrequently purchased (such as cars or furniture) or highly differentiated by quality (such as high-end restaurants), implying a smaller number of establishments, each serving quite large geographic markets. Based on this logic, some categories of retail that are most likely to serve the immediate neighborhood include grocery and convenience stores, pharmacies, laundry services, coffee shops and limited service restaurants, gyms, video rental outlets, and beauty salons/barber shops.

In addition, Reilly's "law of retail gravitation" (1931) suggests that consumers will be willing to travel longer distances to reach larger centers of retail, such as malls or central shopping districts.[4] This means that local retail corridors may have a smaller number (density) of establishments and host a more limited range of stores and services, compared to central shopping areas that attract consumers from a larger area. Reilly's model also implies that the presence of other, nearby retail centers (and especially those with a denser collection of services and goods) will also determine the willingness of consumers to travel to a particular retail

---

[2] Customers are also comprised of non-resident and non-employee commuter or tourist populations. In order to keep the framework simple, we assume that these customers are shopping at a select and limited number of retail centers, many of which correspond with the central business district(s).

[3] This formulation assumes a monocentric model of urban development; in the case of a polycentric metropolitan area, the single CBD might be replaced by several employment subcenters. The same relative density predictions hold however.

[4] Reilly's formulation is based on a city-level model, but here we extend it to apply to neighborhood-based retail centers.

corridor. The "pull" of competing retail centers on consumers will be a function of the distance to and size of the retail corridor (see also Huff 1964).

Travel costs depend on the quality and quantity of general transportation infrastructure in the metropolitan area (prevalence and reliability of subway and bus systems, road infrastructure and level of traffic congestion) as well as proximity of residents and stores to transit nodes. This implies that retail markets close to public transportation hubs or freeway entrances will be relatively less costly to access than areas without transit connections. Retailers' preferences over type of transit infrastructure may vary by dominant transportation mode share in the metropolitan area and by type of goods or services provided. For instance, in cities where car ownership is low and a large share of travel is conducted by walking or public transit, proximity to subway stops or heavy pedestrian traffic will be quite valuable. In cities where driving is the main means of transportation, access to freeway entrances and available parking will be essential, although parking fees may be lower. Driving accessibility may also be more important for certain types of retail goods (furniture, home goods, electronic equipment) that are cumbersome to move by public transit. Suppliers may have specific access requirements, such as loading docks for large trucks.

Fixed costs for retailers include a number of factors, some of which vary by neighborhood and others that are specific to the firm and therefore somewhat idiosyncratic.[5] For instance, rents are likely to be higher in high-income (or high wage) neighborhoods, while insurance and security costs increase with neighborhood crime rates. If we assume that labor markets correspond to metropolitan areas (MSAs), then wages for similar positions (sales clerk or shelf stocker) may be relatively similar across neighborhoods. However, there is some anecdotal evidence that employee turnover or training needs are higher in low income neighborhoods (International Council of Shopping Centers 2004), increasing average labor costs in those areas. Marketing and overhead costs may also vary by firm type; specifically multi-establishment firms may choose to pay for advertising campaigns through a central office while single-establishment firms bear all these costs locally. Two other fixed costs that are likely to vary across neighborhoods are local land use regimes (zoning of commercial uses) and

[5] Many retail firm costs are not "fixed" in the traditional sense, but are also not exactly marginal. For instance, building rents are often fixed over lease terms, which may be 5 or ten years long but may offer some flexibility between leases, depending on negotiations between tenant and landlord. Likewise contracts with suppliers, insurance, utilities, etc., may be fixed over a short period of time (1-2 years), and so cannot be directly reduced with marginal productivity.

characteristics of the local building stock. Specifically, restrictions against or incentives for retail occupancy can increase or reduce costs associated with initial set-up. Similarly, the inherent nature of the building stock will determine the feasibility and costs associated with adapting the particular retail business to the existing commercial space. For example, grocery stores often require enough space and a robust enough infrastructure to support freezers, while restaurants require venting from stoves and ovens (International Council of Shopping Centers 2004; Barragan 2010). Availability of suitable land parcels for development may be particularly important for large chains that have a preferred model for their stores (i.e. Big Box), often a model derived in a suburban or low-density context. These retailers often view parking as a key component, and may be more difficult to accommodate in an urban setting.

Variation in retail density and neighborhood income

For any given type of store or retail product, the Hotelling model implies that the density of stores will be increasing in density of customers (or that distance between stores will be decreasing in customer density). The stylized model assumes that customers are uniformly distributed across space and have homogeneous preferences (i.e. all consumers have the same underlying demand for goods and will purchase given goods at similar frequencies). In reality it is unlikely that all residents of a single neighborhood have the same demand function, either based on income/ability to pay or preferences; so estimating the density of actual rather than potential customers within a given geographic area becomes more complicated. de Palma et. al. (1994) develop a more flexible model that allows for consumer heterogeneity, non-price competition in the form of retail 'variety' and less constrained market boundaries. We apply this intuition to refine our framework.

The primary focus of our current analysis is the relationship between retail markets and local income. Most directly, higher household income implies greater purchasing power among local residents.[6] If we assume that retailers are motivated in their location decisions by profit maximization, retail density should be increasing in the potential for local consumption, or income. Even if higher incomes do not translate into a greater number of purchases, but rather better quality products and services consumed, this still implies rising consumption expenditures

[6] Cash income is not a perfect proxy for purchasing power, especially among lower-income households, who may receive non-cash benefits such as food stamps or housing assistance, and may engage in reciprocal exchange of services in lieu of cash payments. And purchasing power depends not only on current income but also on lifetime income if people smooth consumption over time relative to income fluctuations. Still, income is the most practical empirical indicator of purchasing power.

and thus should induce higher retail density. Besides income, customer preferences are likely driven by characteristics such as race, ethnicity, age and socioeconomic status (Waldfogel 2008). For example, households with children may purchase more groceries and eat in restaurants less frequently than childless households, affecting relative density of establishments in these two categories.

It seems fairly clear that retailers have greater incentive to locate in high income neighborhoods. But as discussed in the introduction, these neighborhoods may also have higher barriers to entry. Particularly if retail involves some disamenities (for example, attracting more automobile and pedestrian traffic, causing noise, excessive litter or safety concerns), high income neighborhoods may use zoning and the political process to block commercial developments. Research on land use in suburban settings shows that communities use zoning to limit commercial activity of most kinds (Hausman and Leibtag 2005). For instance, Boston area suburbs allow business or commercial zoning on approximately five percent of land area, while on average more than 90 percent of land is zoned to allow residential uses (authors' calculations based on Massachusetts Housing Regulation Database (2006)). Even in urban areas, affluent neighborhoods might prefer to confine retail to designated commercial corridors farther from residential areas. Poor neighborhoods, on the other hand, might be more open or less able to manipulate the political system to prevent commercial development.

Variation in size and type of retail and neighborhood income

Variation in retail access may be more nuanced than simply differences in overall density. Specifically, we propose that income and retail density will have a differential relationship depending on the *type* of retail (e.g. grocery versus drug store versus restaurant) and the *size* of the retail establishment. Here, we use size to represent two defining features of local retail stores: (1) the physical space the business occupies and (2) the scope of the business, i.e. the range (and diversity) of goods sold. If retail in general is a normal or luxury good, then retail density overall should be increasing in income, but density may be decreasing in income for specific types of retail that are less desirable. The reverse would hold if retail in general is an inferior good but some products or services are normal or luxury goods. Specifically, establishments such as specialized grocery stores or upscale restaurants are more likely to locate in high income neighborhoods, while establishments selling inferior goods (convenience stores and fast food restaurants) will locate in lower-income areas. This kind of variation within the retail product market is more formally known as horizontal differentiation (de Palma et. al. 1994). Teh (2007) provides an example with liquor stores: she finds that alcohol outlets located in low-SES neighborhoods are seen as disamenities, whereas alcohol outlets located in high-SES neighborhoods – which were more likely to be large grocery stores or upscale wine and liquor stores – were valued by homeowners. In addition, sorting of retail establishments by product quality may be reinforced by zoning, if certain types of food establishments (like bars or fast food places) might attract undesirable crowds or other disamenities.

Income may also be correlated with preferences over the physical size and architectural design of retail establishments, as illustrated in the debate over Big Box Retailers. Anecdotal evidence demonstrates that more affluent communities often protest larger chain retailers, citing loss of neighborhood character (Li 2009). If high income communities have a preference for smaller, locally owned business, retail size should decrease with household income (Zukin 2010). Hausman and Leibtag (2005) show that consumer surplus from increased superstore access is greater for low-income households compared to high-income households. In addition, if higher income households prefer to live in less dense communities (and therefore have more access to car transportation) then retail size will decrease in household income. On the other hand, larger retail establishments may be of value to households, because they can potentially carry a greater variety of goods and offer lower prices (Basker, Klimek, & Hoang 2007 provide evidence for this). Alternatively, the relationship between income and establishment size might be mediated by the population of the local neighborhood. Larger retailers tend to serve larger markets and therefore we might not expect to see significant neighborhood-level variation in size with respect to income (or any other demographics for that matter).

III) Empirical Literature Review

The empirical literature on the relationship between retail presence and local market characteristics is limited. Much of the existing work on retail focuses on a single sector and/or a single geographic area. In addition, the research questions typically center on labor market outcomes rather than linkages between retail presence and consumption markets. Here we summarize the existing research that informs the latter relationship.

A couple of studies link retail market density and size to population size. Berry & Waldfogel's (2003) research on product quality and market fragmentation suggests that as market size increases, the range of product variety and quality widens. They also find that the number of high-quality products grows with market size. Dinlersoz (2004) uses an establishment-level dataset on alcoholic beverage retailers in California to test the difference in the organization of chain versus stand-alone stores. He does find variation across the two types of stores: chain stores expand their scale as market size increases, whereas stand-alone stores tend to grow the number of establishments as market size increases. Davis (2006) looks at the relationship between the distribution of consumers and movie theaters. He finds that demand for the theater (and ticket sales) increases with the number of people living within five miles of the cinema; this increase is less pronounced at further distances.

Waldfogel (2009) exploits the variation in consumer characteristics and empirically tests the relationship between the mix of commercial services and heterogeneity in consumer preferences. He demonstrates that there is considerable heterogeneity across consumer preferences for such services as restaurants and media, and that preferences are strongly correlated with observable population characteristics, such as educational attainment and race/ethnicity. Using 5-digit ZIP-code level data on food and drinking establishments and population characteristics and proprietary data on consumer patronage behavior, he finds that there is an association between the mix of locally available chain restaurants and demographic mix by race and education.

A handful of studies consider the role of consumer income levels in retail location. Glaeser et. al. (2001) explore the role of urban density, and in particular commercial density, in facilitating the growth of consumption centers. Generally they find that high-amenity cities have grown faster than low-amenity cities and that between 1970 and 1990, neighborhoods in Manhattan that are closer to the CBD or a major consumption center have become richer than neighborhoods relatively farther away. Both results suggest that households value access to commercial services and that this preference has strengthened over time.

This conclusion reinforces the findings of a sizable body of literature in public health that explores the differences in the locational decisions of food establishments across neighborhoods. Powell (2007), Zenk (2005) and Alwitt and Donley (1997) demonstrate that various retailers (namely banks and supermarkets) opt not to locate in poorer ZIP codes even after controlling for purchasing power—leading the authors to conclude that retail locational decisions may hinge on a host of factors in addition to an area's market potential. Interestingly, Alwitt and Donley found that fast food restaurants were least likely to discriminate across neighborhoods, whereas Block, Scribner and DeSalvo (2004) and Sloane et al (2005) found that fast food restaurants were more likely to locate in poorer, predominately minority neighborhoods. Meltzer and Schuetz (2010) find that although high-income neighborhoods in New York City have a higher density of retail employment and more chain restaurants, low-income and predominantly black or Latino neighborhoods have a much higher share of unhealthy fast food restaurants.

Finally, Chapple and Jacobus (2009) and Kolko (2009) offer perhaps the most related papers, both using data from the National Establishment Time-Series (NETS) dataset and the Neighborhood Change Database (NCDB). Chapple and Jacobus use ZIP-code level data on retail businesses and Census tract-level data on neighborhood economic and demographic characteristics for the San Francisco Bay area to examine the link between retail revitalization and neighborhood change. They classify neighborhoods into five categories of relative income change and show with descriptive crosstabs that retail revitalization is most strongly associated with gains for middle-income neighborhoods. They hypothesize that this is, in part, due to their greater ability to attract start-up businesses. While they construct a nuanced definition of neighborhood change, their methods are primarily bivariate and leave out controls for neighborhood characteristics that might influence both retail and residential revitalization.

Kolko (2009) looks at the relationship between employment and gentrification at the neighborhood level. He uses the NETS and NCDB data to measure the impact of employment location on neighborhood gentrification during the 1990s for metropolitan areas across the U.S. He finds that, at the tract level, average household income change is positively correlated both with the change in average pay for nearby jobs and with the start-year average pay for nearby jobs. While Kolko focuses on the impact of overall employment on gentrification, we focus on retail presence exclusively and explore the reverse relationship: how well changes in neighborhood income explain changes in local retail presence (as measured by employment as well as establishment density and size). This reverse relationship is particularly appropriate for analyzing retail employment since retail product markets have smaller geographic scope than the markets for many other goods and services. Retail establishments are more likely to follow population than industries that serve other businesses; industries in which output is intangible and can be delivered electronically; industries in which transport costs are low relative to agglomeration economies; and industries that must locate near natural resources.

IV) Data and Methodology

We analyze the relationship between neighborhood income and retail presence with two basic estimation strategies. First we estimate cross-sectional regressions between retail density and median household income in 1990 and 2000, for a variety of retail categories. Second we estimate the impact of both baseline income level and a change in relative neighborhood income (1990 to 2000) on change in retail metrics over a comparable time period (1992-2000).[7] All retail metrics (dependant variables) are constructed from the NETS database, described below, while all right-hand side variables are taken from the Neighborhood Change Database (NCDB), which presents decennial Census data for geographically consistent boundaries. Specific variable definitions and sources are shown in Table 1; summary statistics for all variables are shown in Table 2. Our sample includes the 58 largest MSAs in the U.S. We define "neighborhood" as ZIP Code Tabulation Area (ZCTA), an approximation of U.S. Postal Service ZIP codes created by the Census Bureau. The median population of a ZCTA in our MSAs is 13,700, a reasonable size for locally oriented retail markets (Census tracts are quite small, at least in urban areas, to support much local retail).[8] ZCTA's are assigned entirely to the place and MSA that includes more than half of the ZCTA's population.[9] Therefore our place and MSA boundaries are not exactly consistent with official definitions but are internally consistent, and avoid the quandary of apportioning retail metrics or demographics across ZCTAs that straddle multiple places or MSAs.

The basic form of the regression for the cross-sectional analysis is shown below:

$$Empland_{it} = \alpha Income_{it} + \beta X_{it} + MSA + Yr2000 + \varepsilon_{it}$$

Where $Empland_{it}$ is the employment density of ZCTA i in time t, $Income_{it}$ is the natural log of median household income in ZCTA i in year t, X_{it} is a vector containing population density and a variety of economic, demographic and locational characteristics of ZCTA i in year t (described in more detail below), *MSA* is a set of fixed effects for MSAs, *Yr2000* is a dummy variable for

---

[7] Another typical approach to measure the amenity value of a specific attribute is to include that amenity as a right-hand variable in hedonic regressions of housing prices. We do not use that approach because we lack neighborhood measures of several key variables, namely school quality and crime rates, which would lead to omitted variable bias in such estimations.

[8] ZCTA's were defined by the Census in 2000 only, not 1990. Our 1990 Census data are from GeoLytics, which recalculates 1990 Census long form data to year-2000 Census geographic boundaries. Our ZCTA-level measures therefore reflect consistent boundaries over time, both for Census- and NETS-derived measures.

[9] Some ZCTAs in our sample were split among three places, but in these cases all had greater than 50% in one place and so were assigned to that place. Assignments were based on the MABLE/Geocorr engine, available at http://mcdc2.missouri.edu/websas/geocorr2k.html .

year (1990 is the base year). Because we only have ZCTA-level data on income for two years – 1990 and 2000 – we cannot take full advantage of the annual reporting of retail metrics from the NETS dataset, described below, and simply estimate regressions using retail metrics from the year closest to each Census year (1992 and 2000). Similar regressions using the average of retail metrics over each time period (1992-2000 and 2001-06) produced largely similar results.

All retail metrics are created from the National Establishment Time Series (NETS) database. The NETS is a longitudinal, establishment-level database covering nearly all businesses in the U.S. It is constructed by Walls and Associates from the Dun & Bradstreet business register. Unlike publicly available government data on employment, the NETS includes no suppression of employment in small industry or geographic cells and provides full street address information for each establishment, which we geocoded in order to generate ZCTA-level counts. In addition, industry is reported at the 6-digit NAICS level, and a headquarters identifier permits classification of establishments according to firm size and structure. Finally, because the NETS is longitudinal, we can measure gross employment changes at the establishment level, not just net employment changes: below we present results for retail churn in excess of the level needed to achieve net employment changes.

The primary retail metric is density of employment, calculated by dividing the number of employees in the ZCTA-industry category by total land area of the ZCTA. We also calculate establishment density, using the count of establishments per ZCTA-category and total land area of the ZCTA, and the average establishment size, measured as total employment divided by total establishments. These metrics are calculated separately for all establishments, for those in single-establishment ("independent") firms and those belonging to multi-establishment ("chain") firms.[10] Because our main research focus is on retail that primarily serves the residents of the immediate neighborhood (rather than the type of retail that might attract customers from across the city), and because we are interested in quality of life implications, we have chosen to focus

[10] From NETS data we can identify firm ownership in three ways: single-establishment firms (which we call "independent" for brevity), headquarters of multi-establishment firms (briefly called "chains") and non-headquarters establishments of multi-establishment firms. Ideally we would like to exclude any establishments that do not carry out direct retail (interaction with consumers), or perform little retail relative to other corporate functions, such as personnel or marketing. However, it is likely that some headquarters establishments carry out direct consumer activity while many non-headquarters establishments also carry out general corporate functions, so the headquarters distinction may not be that useful. We estimate our equations both for non-headquarters establishments only and for all establishments belonging to chains. The results are not significantly different, so we present results grouping headquarters and non-headquarters collectively as "chain" establishments.

on several industry categories that meet these criteria: supermarkets (NAICS 6-digit code 445110), pharmacies and personal care stores (NAICS 3-digit code 446), clothing stores (NAICS 3-digit code 448), food service establishments (NAICS 3-digit code 722), and laundry facilities (NAICS code 812). To provide some context we also look at the total number of establishments in retail (NAICS 2-digit 44-45). Note that our "all retail" measure includes many retail industries that we do not look at separately; it also excludes the food service establishments and laundry facilities industries.

The key independent variable is the natural log of median household income. If retail purchases are normal goods, then we would expect to see a positive correlation between income and retail employment density, conditional on population density. As described in Section 2, we would expect retail density to increase with residential density, representing larger potential consumer base, therefore we include a measure of population density. We control for distance from the CBD, as a proxy for employment density.[11] In a sense, the analysis tests hypotheses that retail density is determined by the *quantity* of potential consumers (population and employment density) versus the *quality* or type of potential consumers (income and other characteristics). Distance from CBD should also be correlated with travel costs and accessibility, important cost factors. To control for differences in consumer preferences, we control for a variety of demographic characteristics, namely percent of population with college or graduate degrees, share non-Hispanic black, share Hispanic, share under 18, share over 65, and share foreign born. As noted earlier, it is possible that high-income households are able to manipulate the zoning process to limit undesirable land uses in their neighborhoods. On the assumption that homeowners wield disproportionate power in local land use decisions (see Fischel 2001), we include the share of owner-occupied housing in the ZCTA. Finally, to proxy for the suitability of existing structures to house retail uses, or the availability of land to develop new retail structures, we include the share of housing stock built prior to 1940. Older housing stock is assumed to correlate with smaller existing commercial space.

[11] To identify the CBD, we calculate total employment density in each ZCTA using the NETS data and land area from the Census. The ZCTA within the MSA's primary central city with the highest employment density is designated as the CBD. We then calculate the pairwise distance between each ZCTA and the CBD using latitude and longitude coordinates from the Census for the centroid of each ZCTA. Even if our MSAs are not perfectly monocentric, they all have declining employment density gradients with distance from CBD, so as a first approximation of employment density, this seems reasonable. We run robustness checks stratifying the sample by MSA density gradient, with largely similar results, as shown in Appendix A1.

It is possible that households who choose to live in urban and suburban areas have different preferences over the mix of uses in their environment, with suburban households preferring greater separation of residential and commercial uses. Therefore we also estimate a number of interactions with income, including a dummy variable indicating whether the ZCTA is located in one of the central cities within the MSA, the distance from CBD, the share of the MSA that is mixed residential and commercial use, the overall employment density gradient in the MSA, and the MSA population size.[12] As shown in Appendix Table 1, few of these interactions produced results that are statistically different from zero, and even those with statistical significance were of very small magnitude.

As discussed previously, the income elasticity of demand will differ for various retail goods and services. Similarly, some categories of retail or establishment types may be viewed by neighborhood residents as amenities while others are disamenities. By estimating separate regressions for several different retail categories, we can begin to tease out these distinctions. We might expect that density of establishments selling "necessity" goods and services, such as grocery stores, drugstores, and laundry facilities, will be less sensitive to income: the income elasticity of demand for these goods is presumably less than one, so expenditures would increase with income but at a declining rate. Conversely, goods such as clothing and restaurants may represent "luxury" goods, with high income elasticity of demand. Even within these categories, some establishments such as fast food restaurants or convenience stores may represent "inferior goods" with negative income elasticities. Similarly, the income elasticity for firm type may vary, although it is not immediately obvious in what direction. Some large national (or international) chains are clearly in the luxury market (Barney's clothing or Wolfgang Puck's restaurant chains) while others are more mass-market (Payless Shoes or McDonald's). Therefore we separate establishments based on firm type – chain versus independent – for all retail categories examined.

Although we use employment density as the primary measure of retail density, we also calculate average establishment size to test for differences in market structure by neighborhood

[12] We obtain overall employment density gradient for each MSA by estimating regressions of total employment density (in all industries) against distance from CBD. The coefficient on distance from CBD is then used in the interaction with ZCTA income. For the mixed-use share of the MSA, we calculate the ratio of total employment to population for each ZCTA. ZCTAs with job-population ratios between 0.25 and 0.8 (approximately the 25th and 75th percentiles of the whole sample) as designated as mixed use, and the share of land area within the MSA that is contained by mixed-use ZCTAs is our MSA-level indicator.

income. That is, a network of few large stores or many small stores could yield the same overall employment density, but provide in some senses differential access to consumers. A priori it is unclear whether households with different incomes would prefer different retail networks. Anecdotal evidence suggests that higher-income households may have preferences for small, locally owned stores with more distinctive "character" than large, corporate stores (Zukin et. al. 2009). Smaller stores may also offer a higher level of customer service, which high income households might prefer. And many discount stores targeting lower-income consumers tend to use a large store format, or may prefer lower-rent locations because of their need for large spaces. Alternatively, if low-income households have less access to cars, they may be more dependent on stores within closer proximity, suggesting a higher density of small establishments in low income neighborhoods. The theoretical predictions are ambiguous, but can be tested empirically with our data.

There are several determinants of retail density we are not able to directly measure with the available data. First, operating costs for retailers are likely to be higher in neighborhoods with high crime rates, due to the need for greater security or higher insurance costs. We have no information on crime rates at the ZCTA level, so cannot control for this. Based on existing literature, crime rates are likely to be negatively correlated with income and negatively correlated with retail density, leading to a potential positive bias on the income coefficient by omitting crime rates. Nor can we directly measure neighborhood differences in commercial rents or labor costs (wages, training costs or turnover). Rents are likely positively correlated with income and negatively correlated with retail density, introducing a downward bias into the income coefficient. The direction of correlation between labor costs and income is ambiguous; in a competitive market wages should be roughly similar across neighborhoods within an MSA, but if low-income neighborhoods have higher turnover, then true labor costs may actually be higher in those neighborhoods. Finally, it is not feasible with publicly available data to directly measure transportation access and costs or the availability of building space and land to develop commercial uses. The correlation between transit access and income is ambiguous. If access to transit is perceived as an amenity and positively correlated with income, then the coefficient on income will be biased upward. However, if access to transit is associated with more noise or congestion, then the bias on the income coefficient will be reversed. Likewise, if higher income neighborhoods are more likely to oppose commercial development, then the coefficient on income will be biased downward. In order to address the variation in ease of commercial development, we run a version of the regression where income is interacted with the share of the MSA that is classified as mixed use. These results are displayed in column 4 of Appendix Table A, and show that retail density increases with income in more mixed-use MSAs.

Since neighborhoods are not economically or demographically static, we also model a dynamic relationship between income and retail. We estimate models of the change in retail density as a function of both initial income level and change in income. The basic form of the regression for the dynamic analysis is shown below:

$$\Delta Employment_{ijt,t-1} = \alpha \Delta RelativeInc_{it,t-1} + \beta Inc_{it-1} + \gamma X'_{it,t-1} + MSA + \varepsilon_{it}$$

Where $\Delta Employment_{ijt,t-1}$ is the average annual employment growth rate in ZCTA i within industry category j between time t and t-1, $\Delta RelativeInc_{it,t-1}$ is the change in median household income in ZCTA i relative to the MSA between time t and t-1, Inc_{it-1} is the log of median household income in ZCTA i in year t-1, $X_{it,t-1}$ is a vector of demographic and economic characteristics of ZCTA i (including baseline in year t and change between t and t-1), MSA is a set of fixed effects for MSAs. In all our dynamic specifications, the baseline year is 1990 and the change is between 1990 and 2000. Change in retail metrics occurs between 1992 (the first year data are available) and 2000.

Employment growth rate is calculated using a standard measure:

$$g_{it} = \frac{(Emp_{it} - Emp_{it-1})}{0.5 * (Emp_{it} + Emp_{it-1})}$$

in which Emp_{it} is the number of employees in industry i in time t. As discussed in several previous papers that have used this measure, this growth rate provides a symmetric growth rate that is useful for estimation and, by using a two-year average employment level rather than a single year of employment in the denominator, reduces potential measurement error associated with large single-year deviations from average employment (see Davis et al 1996, Haltiwanger et al 2010 for more discussion). Because we are effectively condensing eight annual estimations of employment change into a single period, we calculate each year-on-year change and average across the period. Calculations were also made of the compound annual growth rate using beginning and ending year employment; regressions using both growth rates are highly similar, so we followed standard practice by using the measure described above.

Note that these growth rates use net change in employment and establishment. We are also interested in the volume of gross job creation and destruction, as an indicator of level of economic activity. So our final metric of retail change is an indicator of churn:

$$Churn_{it} = \frac{Job\ creation_{it,t-1} + (-Job\ destruction_{it,t-1}) - |Net\ emp\ change_{it,t-1}|}{Emp_{it-1}}$$

In which $Job\ creation_{it,t-1}$ is the gross increase in employment from all sources (establishment birth, expansion and in-moves) between periods t and t-1, $Job\ destruction_{it,t-1}$ is the gross decrease in employment (establishment death, contraction and out-moves) between periods t and t-1, $Net\ emp\ change_{it,t-1}$ is the net change in employment in industry i between time t and t-1 (gross job creation less gross job destruction), and Emp_{it-1} is the total employment in industry i in year t-1. This indicates the level of excess employment change, that is job creation and destruction above that which would be needed to create the net gain (or loss) in employment observed (see Davis et al 1996; Davis and Haltiwanger 1992).

The right-hand side of the equation includes both a baseline measure of income (natural log of income in 1990) and a measure of relative change in income between 1990 and 2000. The relative change measure is calculated as follows:

$$\Delta ZCTAInc = \frac{Inc_{ijt}}{Inc_{jt}} - \frac{Inc_{ijt-1}}{Inc_{jt-1}}$$

Where Inc_{ijt} is the median household income in ZCTA i in MSA j in year t, Inc_{jt} is median household income in MSA j in year t. Essentially this measure indicates the change in ratio of ZCTA household income to MSA household income between 1990 and 2000. We use a relative income change measure to indicate upgrading of the neighborhood, relative to the surrounding MSA; this should capture whether a ZCTA is becoming more affluent (thus a more desirable location for retailers), compared to other ZCTAs within the MSA. Intuitively, if a neighborhood's absolute income rises but at a similar or slower pace than surrounding neighborhoods, it is less likely to attract additional retailers than if a neighborhood which experiences smaller absolute gains (or even losses) but whose income growth outpaces other neighborhoods within the MSA. Several recent papers on gentrification or neighborhood change have used relative income gain (or loss) measures (see Ellen and O'Regan 2008, Bostic and Martin 2003, McKinnish Walsh and White 2009). The base regression models were also estimated using several other income change measures, including a simple log of change in median household income, percentage change in income, and difference between percentage change of ZCTA income and MSA income. All regressions also include MSA fixed effects, so results are nearly identical in sign and significance regardless of income measure. .

Besides the change in income and baseline income, right hand side variables include the same baseline metrics of demographics described in the cross sectional model, and changes in these variables (except distance to CBD and central city status, which do not change over time).

All regressions include metropolitan area fixed effects and robust standard errors clustered by Census place, to account for any spatial autocorrelation by political jurisdiction (such as city-wide zoning rules or business start-up fees). All regressions are weighted by ZCTA population, due to large variation in ZCTA size (this reduces distortion of results by sparsely populated ZCTAs on the urban fringe).

V) Results

Descriptive Statistics

Consistent with the general monocentric city model, Figure 1 shows that all retail employment categories have negative density gradients moving away from the CBD, but the rates of decline differ by category. Food service and clothing are the most centralized categories, with the steepest declines in the first few miles from the CBD. Those results are consistent with food service being oriented towards high levels of general employment in the CBD or possibly "destination" restaurants that draw consumers from across the metro area. It is also plausible that residents near the CBD live in smaller housing units and therefore are more likely to eat out. Similarly, there are likely to be large, well-known clothing retailers in most downtown shopping districts. In contrast, supermarkets have the flattest density gradient, consistent with a need in all residential neighborhoods for basic groceries.

Figure 2 suggests that there are large differences in sensitivity of retail density to income across retail categories. Employment densities for overall retail and food service increase steeply with rising household income up to about $50,000 (approximately the sample mean income), then flatten or slightly decline. Given that eating at restaurants is a more expensive substitute for eating at home, and thus less common for low-income households, it is not surprising that food service employment would be quite sensitive to neighborhood income. Other retail categories show employment densities that are nearly flat with respect to income. This makes sense for supermarkets and drugstores, which sell primarily necessities for which income elasticity will presumably be small. The apparent lack of correlation between income and clothing employment is more surprising, but likely masks large variation in the type, quality, and price of goods sold by neighborhood. Figure 3 shows the average establishment size increasing with income for all categories, although at different rates. Clothing and drugstores have the least variation in size by income, while supermarkets have the strongest correlation.

Table 3 shows the mean for all retail metrics across categories and by firm type (chain vs. independent). For comparison purposes, we also include the mean for all industries. Of our retail categories, food service has by far the highest density of employment and establishments, with two thirds the employment density and more than double the establishment density for all retail sectors. Looking at the categories broken out by firm type, independents dominate in establishment density for all categories, but chains dominate in employment density for most (laundry and food service are the only categories with higher employment density in independent establishments) – because chain establishments have more employees, on average, than independent establishments. The ratio of independent to chain employment varies considerably across categories, however. Looking at the average size of establishments, we find that overall retail establishments are quite small, around 8.25 employees, smaller than average size for all industries, and size varies widely by category. Notably, the average size for supermarkets is just under 25 employees, although independents are much smaller (7.6) while chains are much larger (45.8). This suggests that the NAICS category for supermarkets captures many small stores, such as corner bodegas, as well as full service supermarkets. The average size of clothing stores is perhaps surprising; these appear smaller than would be expected of stores in typical suburban malls.[13]

The last three columns of Table 3 compare growth metrics by category. The retail sector overall grew somewhat more slowly than all industries, measured either by employment or establishment growth. Supermarkets and clothing had slower employment growth than retail sector overall, drugstores and laundry approximately the same as the sector and restaurants faster growth. Among supermarkets, drugstores, and food service, chains had faster employment growth than independents, consistent with anecdotal evidence that chains made inroads into

[13] The data on establishment size does not distinguish between full-time and part-time employees, so should be read as total employees on the payroll, not FTEs.

urban markets during this time (Center for an Urban Future 2009). Churn rates also differ by category: for retail overall, excess job change over 8-year period is nearly 40% of baseline employment, lower than the churn rate for all industries (0.488). Supermarkets and drugstores had slightly lower rates of churn while laundry had a much lower rate of churn (0.122). For clothing and restaurants, churn represents nearly half of baseline employment, perhaps suggesting frequent turnover of establishments and potentially short average lifespan.

Results of cross-sectional analysis

The results in Table 4 show that there is a positive relationship between neighborhood income and retail employment density for retail in general, controlling for neighborhood characteristics. A simple bivariate regression, including only year and MSA fixed effects, indicates a significant negative relationship between median household income and employment density. This likely reflects the spatial distribution of income across MSAs – higher income households tend to live farther from the CBD in lower-density neighborhoods, both of which should be associated with lower retail density. Once we control for population density and distance to CBD in column 2, the coefficient on income becomes positive and significant. In order to compare magnitudes of coefficients for the three variables, we estimate standardized betas (with variables normalized to mean zero and standard deviation of one), the coefficient on population density has largest magnitude, 0.72 compared to 0.09 for income and 0.19 for distance to CBD (standardized betas not shown but available upon request). The coefficient on population density also has the strongest statistical significance and in bivariate regressions yields the highest R-squared. The coefficient on distance from CBD is negative and significant, as expected. Overall these results are consistent with predictions that retail density is quite sensitive to density of employment and population, as well as to income. In Column 3 we add a set of standard demographic and economic characteristics, which reduces the magnitude of the income coefficient slightly, but still yields a positive and strongly significant result. Most controls perform as expected. The negative coefficient on share of owner-occupied housing is consistent with an interpretation that homeowners tend to resist commercial development. Retail density declines with share of black and Hispanic population, consistent with some prior research that shows that minority populations generally have less access to retail. The negative coefficient on older housing stock may indicate that older buildings are structurally less suitable for retail use. Column 4 suggests that establishment size is increasing in income. This could fit with an explanation of higher income neighborhoods offering greater demand for a wide range of products and services, or with higher fixed costs (such as rent or obstacles to development) causing retailers to operate larger stores that serve a larger market area.

Next we examine the relationship between income and employment density. Table 5 displays the results by type of retail and for the retail sector overall. The results largely agree with the simple graph in Figure 1: employment density increases with income for all retail and for food service (although the latter is only marginally significant), but there is no statistically significant association between income and employment in supermarkets, drugstores, clothing stores or laundry. It is notable that the coefficients on other controls, particularly population density, distance from CBD and share of owner-occupancy, are all consistent in sign and significance, although magnitude varies by category. That suggests that retail employment for all categories reflects size of potential market and possible NIMBYism, but that income elasticity of demand for or amenity value of products and services varies.

Table 6 explores the relationship between income and firm structure and size, for all retail categories. Each row presents the coefficient on income for total employment density, employment density by firm status (independent vs. chain), and for establishment size. For retail as a whole, income is positively associated with total employment and chain employment, but has no relationship with employment in independent establishments. For two categories, supermarkets and drugstores, the null finding on the relationship between income and employment density masks opposing forces: income is negatively associated with employment at independent establishments but positively associated with chain employment. Intuitively, this suggests that there is roughly consistent demand or amenity value for groceries and pharmacy products, but a difference in preferred firm structure/type. Such preferences may reflect a difference in products (either quality or range) or different pricing by chains. And for all categories, income is positively associated with average number of employees per establishment, although of varying magnitude. Supermarket size shows the greatest sensitivity to neighborhood income, clothing shows the least. It is not possible to assess whether the larger size associated with high income neighborhoods indicates a higher prevalence of big-box retailers, but the results are not consistent with a stated preference for smaller, locally owned stores. Note that because total employment density does not vary systematically by income and establishment size increases with income, this implies that density of establishments is decreasing in income (also confirmed by estimating regressions using establishment density as the dependent variable).

Dynamic regressions

Table 7 switches to examine how the *change* in retail employment over time relates to income levels and changes. Column 1 presents results of change in employment against baseline income and other baseline characteristics. Initial income does not appear to predict employment growth: not only is the coefficient on income not significant, but in magnitude it is close to zero. Few of the control variables have significant coefficients either. Population density is negatively associated with growth, as is black share of population and share of elderly population. The coefficient on educational attainment is positive but only weakly significant. In Column 2, we add the change in relative neighborhood income to the regression. The coefficient on income change is positive and statistically significant, and causes the coefficient on baseline income to increase and be significant as well, suggesting that income growth does attract additional retail.[14] However when we add change measures for the other ZCTA characteristics (Column 3), neither of the income variables remains significant. Most of the controls are not significant predictors of retail employment growth either, although faster population growth is associated with faster retail growth. The final two columns examine whether baseline income or income change predict churn (excessive job creation and destruction), but do not yield significant coefficients on income.

Similar regressions were estimated for changes in employment for all retail categories; results are summarized (control variables are included but not shown) in Appendix Table B. In general, the results on employment change are less consistent and robust than the results on employment levels. Several categories suggest that income level and income growth are positively associated with employment growth, when controlling for levels of other characteristics (Columns 1 and 2). And the coefficient on income change remains significant for supermarkets, when controlling for changes in other demographic and economic characteristics, while both income level and change are significant predictors of drugstore employment growth, controlling for levels and changes in other variables (Column 3). It is notable that employment levels for supermarkets and drugstores were also among the most sensitive to income (at least when separating by firm type). Overall these results are consistent with an interpretation that as

[14] Baseline income and income change have a correlation of -0.32.

neighborhood income increases, small independent groceries and pharmacies are replaced with larger branches of supermarket and drugstore chains, a trend that has been noted anecdotally in some gentrifying areas. The results on retail employment change in general suggest that employment growth in relatively small geographic areas may be somewhat idiosyncratic and is not easily predicted by observable characteristics.

VI) Conclusions and Policy Implications

The urban economics literature on neighborhood amenities has focused mainly on public goods, such as schools, parks and safety. Private goods, such as retail and basic household services, can also have important quality of life implications. Except for limited and largely anecdotal evidence on the dearth of some types of retail (grocery stores, banks, non-fast food restaurants) in poor neighborhoods, we have relatively little evidence on whether retail presence varies within metropolitan areas by neighborhood income. In this paper, we have offered a first analysis of the relationship between income and retail employment density for a variety of retail categories, firm types and sizes.

Results suggest that retail density increases with population density and decreases with distance from CBD and share of owner-occupied housing. The latter result may indicate a NIMBY response of homeowners to commercial uses they perceive as undesirable. In addition, employment density increases with income for some types of retail: for retail overall, food service, and in chain supermarkets and drugstores. Inversely, employment density in independent supermarkets and drugstores decreases with income. Average establishment size, measured as number of employees per establishment, increases with income for all categories. There is some evidence to suggest that growth in retail employment is more rapid in neighborhoods that undergo upgrading (gentrification), particularly for supermarkets and drugstores.

Our results raise a number of questions that invite further research. First, why is there such a consistently strong relationship between income and establishment size? Is this due to differences in operations costs of serving higher income neighborhoods, or greater willingness by large firms (especially regional or national chains) to enter markets perceived as less risky or more profitable? Low-income households presumably have the most to gain from lower prices made possible by economies of scale, yet are less likely to benefit from them. Are there differences in household buying patterns that could explain this? For instance, perhaps low income households have less access to cars and are more dependent on smaller local stores, or have less storage space and so make more frequent trips. Our current data do not allow us to tease out alternative explanations, and would likely need to be supplemented by more micro-level data on household buying patterns to answer the question. In general our results do not match the results from prior, mostly case study research, which finds that low-income neighborhoods have a smaller number of supermarkets.

Most of the categories we examined are basic necessities – food, drugstores, and laundry – which might be expected to have a relatively low income elasticity of demand. But it is perhaps somewhat surprising that employment density in two of the categories that might represent more discretionary spending, clothing and restaurants, are also relatively uncorrelated with neighborhood differences in income. One problem with categorizing establishments based solely on NAICS code is that these codes obscure wide variation in the quality and range of goods and services. For instance, we would expect employment in upscale restaurants to be quite sensitive to income, and employment in coffee shops and delis less so.

We also have limited information on some components of store costs that could be correlated with income, and may introduce bias into our results. Examples include crime rates, which affect security and insurance costs, transportation access and costs, and suitability of existing structures (or availability of land for new development) for commercial uses. Local policies such as zoning or tax incentives for businesses may also affect the incentive or ability to operate retail in neighborhoods of differing income. Obtaining accurate data on such costs or policies at the neighborhood level is infeasible for a large national study, but might be possible for a single MSA.

References

Alwitt, L. and T. Donley. 1997. Retail Stores in Poor Neighborhoods. Journal of Consumer Affairs 31: 139-164.

Barragan, Roberto. September 9 2010. Presentation at Federal Reserve Bank of Kansas City, Economic Development in Underserved Communities conference.

Basker, E., Klimek, S., and Hoang Van, P. 2007. Supersize it: the Growth of Retail Chains and the Rise of the Big Box Retail format. Mimeo, University of Missouri.

Beaumont, Constance. 1997. How Superstore Sprawl Can Harm Communities (And What Citizens Can Do About it). Washington DC: National Trust for Historic Preservation.

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

Berry, Steven T. and Joel Waldfogel. 2003. "Product Quality and Market Size." NBER working paper #9675.

Block, J., R. Scribner, and K. DeSalvo. 2004. Fast food, race/ethnicity, and income: A geographic analysis. American Journal of Preventive Medicine 27(3): 211-217.

Bostic, Raphael W. and Martin, Richard W. 2003. Black home-owners as a gentrifying force? Neighborhood dynamics in the context of minority home-ownership. Urban Studies 40(12): 2427-2449.

Bruni, Frank. September 16 2010. Going Upscale Uptown. New York Times.

Chapple, Karen and Rick Jacobus. 2009. Retail Trade as a Route to Neighborhood Revitalization" In H. Wial, N. Pindus, & H. Wolman (Eds.), Urban and Regional Policy and its Effects. Washington D.C.: Brookings Institution-Urban Institute.

Davis, Peter. 2006. Spatial Competition in Retail Markets: Movie Theaters. The RAND Journal of Economics 37(4): 964-982.

Davis, Steven, John Haltiwanger, Scott Schuh. 1996 . Job Creation and Destruction. Cambridge, MA: MIT Press.

Davis, Steven, and John Haltiwanger. 1992. Gross Job Creation, Gross Job Destruction, and Employment Reallocation. Quarterly Journal of Economics 107(3): 819-863.

Dinlersoz, Emin M. 2004. Firm Organization and the Structure of Retail Markets. Journal of Economics and Management Strategy 13(2): 207-240.

de Palma, Andre, Robin Lindsey, Balder von Hohenbalken, and Douglas S. West. 1994. Spatial Price and Variety Competition in an Urban Retail Market: A Nest Logit Analysis. International Journal of Industrial Organization 12(3): 331-357.

Ellen, Ingrid Gould and Katherine O'Regan. 2008. Reversal of Fortunes? Low-Income Urban Neighborhoods in the 1990s. Urban Studies 45(4): 845-869.

Fischel, William. 2001. The Homevoter Hypothesis. Cambridge: Harvard University Press.

Foster, Lucia, John Haltiwanger and C. J. Krizan. 2006. Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s. The Review of Economics and Statistics 88(4): 748–758.

Glaeser, Ed, Jed Kolko, and Albert Saiz. 2001. Consumer City. Journal of Economic Geography 1(1): 27-50.

Haltiwanger, John, Ron Jarmin and C.J. Krizan. 2010. Mom-and-Pop meet Big-Box: Complements or substitutes? Journal of Urban Economics 67(1): 116-134.

Hausman, Jerry and Ephraim Leibtag. 2005. Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart. MIT and USDA working paper.

Hotelling, H., 1929. Stability in competition. Economic Journal 39: 41–57.

Huff, David L.1964. Defining and Estimating a Trading Area. Journal of Marketing, 24(3): 34-38.

International Council of Shopping Centers. 2004. Developing Successful Retail in Underserved Urban Markets.

Kolko, Jed. 2009. Job Location, Neighborhood Change, and Gentrification. PPIC working paper.

Li, Roland. November 23 2009. City Plans To Rezone Far West Village. The New York Observer.

McKinnish, Tara, Randall Walsh and Kirk White. 2009. Who Gentrifies Low Income Neighborhoods? Journal of Urban Economics (forthcoming).

Meltzer, Rachel and Jenny Schuetz. 2010. Bodegas or Bagel Shops? Neighborhood Differences in Retail & Household Services. Working paper.

Mitchell, Stacy. 2006. Big Box Swindle: The True Cost of Mega Retailers and the Fight for America's Independent Businesses. Boston: Beacon Press.

Moore, Natalie. Aug 12 2010. In Chicago, A Plan to Quench 'Food Deserts'. WBEZ Public Radio.

Osen, Lauren. July 26 2010. Lack of grocery options leads to higher obesity rates in South, East LA. KPCC Southern California Public Radio.

Powell L., et. al. 2007. Food Store Availability and Neighborhood Characteristics in the United States. Preventive Medicine 44(3): 189-195.

Reilly, William J. 1931. The Law of Retail Gravitation. New York: Knickerbocker Press.

Scroop, Daniel. 2008. The Anti-Chain Store Movement and the Politics of Consumption. American Quarterly 60(4): 925-949.

Shaffer, Amanda and Robert Gottlieb. Nov 5 2007. Filling in 'food deserts'. Los Angeles Times.

Sloane, D.C., L.B. Lewis, L.M. Nascimento. 2005. Assessing Healthy Food Options in South Los Angeles Restaurants. American Journal of Public Health 95(4): 668-673.

Waldfogel, Joel. 2008. The median voter and the median consumer: Local *private* goods and population composition. Journal of Urban Economics 63: 567-582.

Zenk, S., et. al. 2005. Neighborhood Racial Composition, Neighborhood Poverty, and the Spatial Accessibility of Supermarkets in Metropolitan Detroit. American Journal of Public Health 95: 660-667.

Zukin, Sharon, Valerie Trujillo, Peter Frase, Danielle Jackson, Tim Recuber, and Abraham Walker. 2009. New Retail Capital and Neighborhood Change: Boutiques and Gentrification in New York City. City and Community 8(1): 47-64.

Figure 1

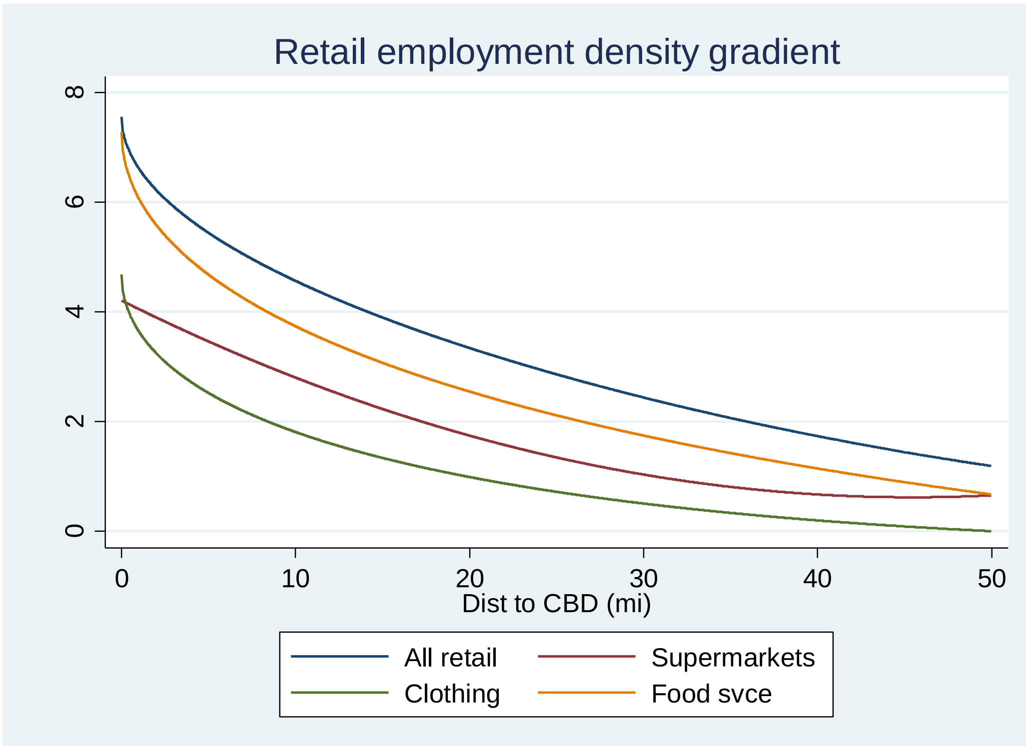

Figure 2

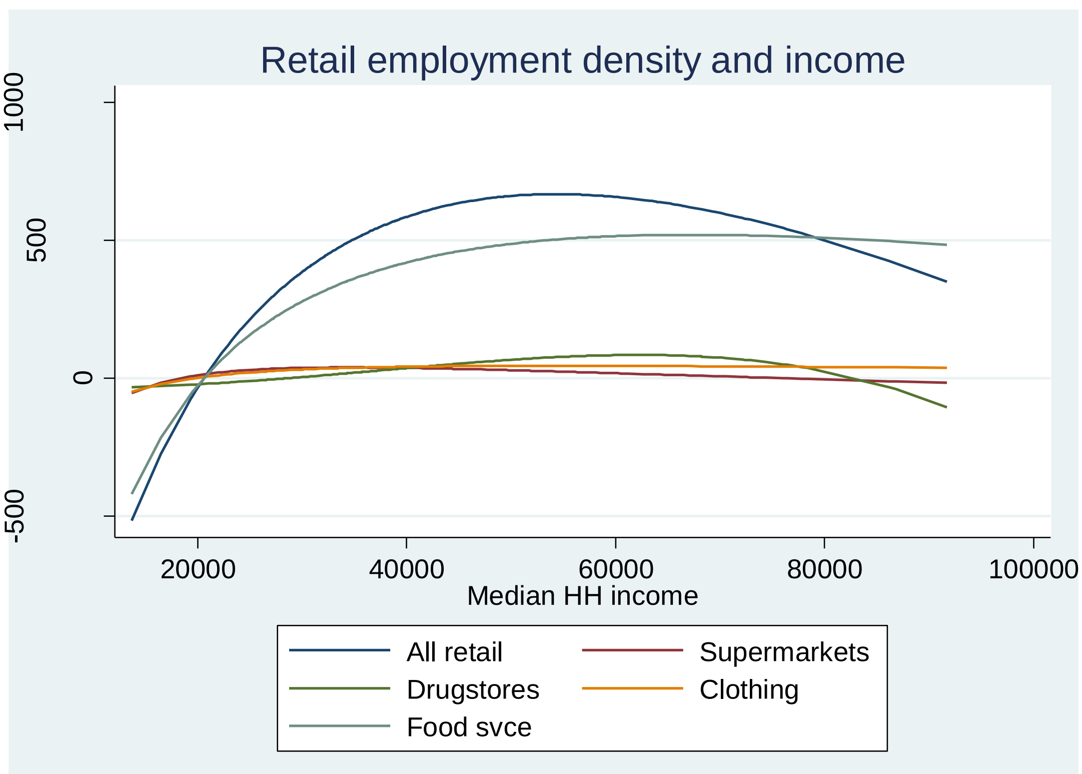

Figure 3

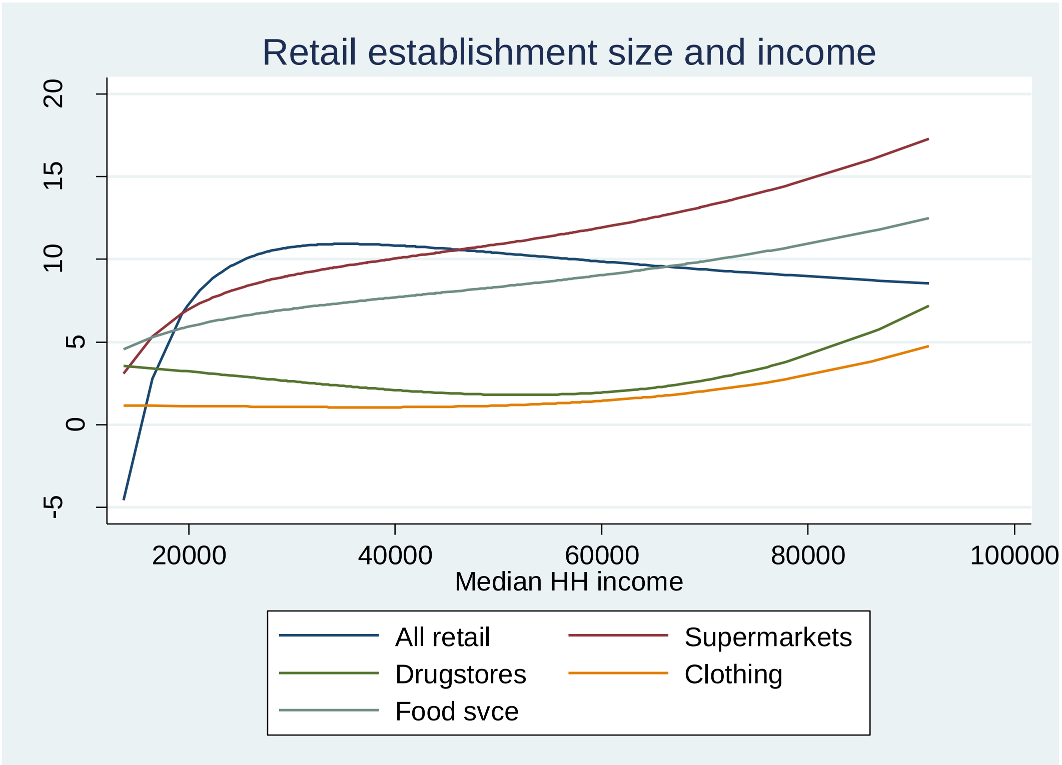
Retail establishment size and income
20
15
10
5
0
-5
20000
40000
60000
80000
100000
Median HH income
All retail
Supermarkets
Drugstores
Clothing
Food svce

Table 1: Variable definitions and sources

Variable	Definition	Source(s)
<u>Retail metrics</u>		
Emp/sq mi	Employees per sq mi (by retail category)	NETS (1992-2006), Census (2000)
Ind emp/sq mi	Employees in independent retail estabs per sq mi	
Chain emp/sq mi	Employees in multi-establishment (chain) retail estabs per sq mi	
Est/sq mi	Establishments per sq mi (by retail category)	
Ind est/sq mi	Independent establishments per sq mi	
Chain est/sq mi	Chain establishments per sq mi	
Emp/estab	Avg employees per establishment (by retail category)	NETS (1992-2006)
Ind emp/estab	Avg employees per independent establishment	NETS (1992-2006)
Chain emp/estab	Avg employees per chain establishment	NETS (1992-2006)
Emp growth	Average annual growth rate in employees (Davis growth formula)	NETS (1992-2000)
Est growth	Average annual growth rate in establishments (Davis growth formula)	NETS (1992-2000)
Emp/est growth	Compound average growth rate in employees/estab	NETS (1992-2000)
Churn	(Gross job creation + gross job destruction - net employment change)/base employment	NETS (1992-2000)
<u>Demographic & economic characteristics</u>		
medhhinc	Median household income	Census (1990, 2000)
Δ ZCTAinc/MSAinc	Change, 1990-2000, (ZCTA median household inc/MSA median HH inc)	Census (1990, 2000)
centcity	= 1 if ZCTA in designated central city, 0 otherwise	OMB (2000)
popland	Population/sq mi	Census (1990, 2000)
distcbd	Distance from ZCTA centroid to Central Business District	Census (1990, 2000)
baplus	% population with BA or graduate degree	Census (1990, 2000)
ownocc	% housing units that are owner-occupied	Census (1990, 2000)
nhblack	% non-Hispanic black population	Census (1990, 2000)
Hisp	% Hispanic population (any race)	Census (1990, 2000)
popkids	% population under 18 years	Census (1990, 2000)
age65pl	% population 65+ years	Census (1990, 2000)
forborn	% population foreign born	Census (1990, 2000)
hsgpre40	% housing built prior to 1940	Census (1990, 2000)

Table 2: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Emp/sq mi	13614	308.46	1632.57	0	56,793
Est/sq mi	13614	35.41	166.18	0	8,400
Emp/estab	13614	8.25	20.39	0	1,759
Emp growth	6807	0.029	0.077	-0.507	0.585
Est growth	6807	0.035	0.060	-0.370	0.497
Emp/est growth	6807	0.015	0.236	-1.000	3.996
Churn	6770	0.391	0.816	0.000	52.224
medhhinc	13614	51,684	21,613	0	206,724
Δ ZCTAinc/MSAinc	6807	-0.006	0.312	-3.847	5.323
centcity	13614	0.253	0.435	0	1
popland	13614	4,037	14,000	0	808,000
distcbd	13614	19.29	15.38	0	239
baplus	13565	24.52	16.72	0	100
ownocc	13542	67.48	20.95	0	100
nhblack	13570	10.71	19.30	0	100
hisp	13570	9.83	16.34	0	100
popkids	13570	25.26	6.61	0	86
age65pl	13570	11.83	6.62	0	100
forborn	13570	9.50	11.87	0	100
hsgpre40	13548	17.11	17.83	0	100

Note: Retail metrics for all retail (44-45). Comparison of retail categories in Table 3.

Table 3: Comparison of retail metrics by category

	Emp/sq mi	Est/sq mi	Emp/estab	Emp growth	Est growth	Churn
All industries	6268.44	322.80	13.17	0.034	0.042	0.488
Independent	2279.53	262.31	6.97	0.025	0.039	
Chain	3988.91	60.48	43.88	0.061	0.072	
All retail	308.46	35.41	8.25	0.029	0.035	0.391
Independent	136.53	28.54	4.79	0.017	0.032	
Chain	171.93	6.87	20.84	0.050	0.051	
Supermarkets	33.71	2.15	24.65	0.021	0.035	0.360
Independent	11.98	1.83	7.64	0.017	0.036	
Chain	21.73	0.32	45.79	0.026	0.028	
Drugstores	14.81	1.47	7.86	0.030	0.031	0.346
Independent	5.92	0.96	4.26	0.007	0.020	
Chain	8.88	0.51	10.97	0.048	0.044	
Food svce	183.64	10.53	13.48	0.040	0.038	0.507
Independent	116.41	8.71	10.41	0.024	0.031	
Chain	67.23	1.82	21.88	0.065	0.064	
Clothing	37.93	4.54	4.57	0.025	0.029	0.480
Independent	14.20	3.26	2.59	0.018	0.028	
Chain	23.73	1.27	6.94	0.017	0.014	
Laundry	7.84	1.67	3.52	0.030	0.035	0.122
Independent	6.39	1.53	3.16	0.024	0.032	
Chain	1.45	0.14	2.96	0.028	0.025	

Notes: All retail includes NAICS 44-45, does not include food service or laundry.

Table 4: How does retail density vary by neighborhood income?

Dep var:	Ln(Emp density)			ln(emp/est)
	(1)	(2)	(3)	(4)
linc	-0.877***	0.441***	0.385***	0.283***
	(0.095)	(0.056)	(0.092)	(0.042)
lpopland		0.829***	0.773***	0.038***
		(0.034)	(0.034)	(0.011)
ldist		-0.389***	-0.418***	-0.087***
		(0.061)	(0.053)	(0.017)
ownocc			-0.014***	-0.006***
			(0.002)	(0.001)
centcity			-0.092**	-0.029
			(0.040)	(0.018)
baplus			-0.0003	-0.0023***
			(0.001)	(0.001)
nhblack			-0.0064***	-0.002***
			(0.001)	(0.000)
hisp			-0.0026**	0.0002
			(0.001)	(0.001)
popkids			-0.022***	-0.006***
			(0.003)	(0.002)
age65pl			0.0156***	0.003
			(0.004)	(0.002)
forborn			0.0007	-0.0008
			(0.001)	(0.001)
hsg pre-1940			-0.011***	-0.0096***
			(0.002)	(0.001)
Fixed effects	Year & MSA	Year & MSA	Year & MSA	Year & MSA
Observations	13,570	13,570	13,542	13,542
R-squared	0.245	0.735	0.773	0.247

Robust standard errors, clustered by place, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Does relationship between income and retail employment vary by retail industry?

Dep var:	Ln(Emp density)					
Industry	All retail (1)	Supermarkets (2)	Drugstores (3)	Clothing (4)	Food svce (5)	Laundry (6)
linc	0.385***	0.008	0.050	0.182	0.232*	0.139
	(0.092)	(0.105)	(0.095)	(0.142)	(0.126)	(0.086)
lpopland	0.773***	0.644***	0.494***	0.409***	0.683***	0.383***
	(0.034)	(0.031)	(0.023)	(0.025)	(0.031)	(0.020)
ldist	-0.418***	-0.257***	-0.278***	-0.358***	-0.479***	-0.190***
	(0.053)	(0.044)	(0.037)	(0.053)	(0.054)	(0.037)
ownocc	-0.014***	-0.007***	-0.008***	-0.015***	-0.014***	-0.0095***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)
centcity	-0.092**	0.019	-0.027	-0.068	-0.023	-0.061*
	(0.040)	(0.039)	(0.035)	(0.061)	(0.039)	(0.035)
baplus	0.000	0.002	0.000	0.009***	0.0042**	0.005***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
nhblack	-0.006***	-0.004***	-0.006***	-0.002	-0.008***	0.003***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
hisp	-0.0026**	0.001	-0.005***	-0.001	-0.001	-0.0024**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
popkids	-0.022***	-0.008**	-0.020***	-0.026***	-0.049***	-0.0195***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.003)
age65pl	0.016***	0.017***	0.0195***	0.0199***	0.007	0.007***
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
forborn	0.001	0.003	0.005***	0.009***	-0.0034*	0.005***
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
hsg pre-1940	-0.011***	-0.0032**	0.0017*	-0.0037*	-0.008***	0.007***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Fixed effects	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA
Observations	13542	13542	13542	13542	13542	13542
R-squared	0.773	0.665	0.702	0.613	0.766	0.752

Robust standard errors, clustered by place, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Relationship between income and firm structure, by retail industry

Dependent var:	Ln(Emp density)			Ln(Emp/estab)
Firm type:	All firms	Independents	Chains	All firms
All retail	0.385***	0.0823	0.618***	0.283***
	(0.092)	(0.071)	(0.148)	(0.042)
Supermarkets	0.0083	-0.551***	0.535***	0.655***
	(0.105)	(0.090)	(0.183)	(0.095)
Drugstores	0.0495	-0.300***	0.316**	0.469***
	(0.095)	(0.087)	(0.139)	(0.064)
Clothing	0.182	0.044	0.151	0.240***
	(0.142)	(0.118)	(0.176)	(0.067)
Food svce	0.232*	0.038	0.135	0.326***
	(0.126)	(0.113)	(0.134)	(0.056)
Laundry	0.139	0.106	0.114	0.316***
	(0.086)	(0.082)	(0.088)	(0.052)

Results shown are coefficients on log(income). All regressions include controls for population density, distance to CBD, owner-occupied housing share, central city dummy, share with BA or graduate degree, share non-Hispanic black, share Hispanic, share under 18, share 65 and older, share foreign-born, share housing pre 1940, year and MSA fixed effects.
N = 13,542 for all cells.
Robust standard errors, clustered by place, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Relationship between retail employment growth and income levels and changes

Dep var:	Emp growth			Churn	
	(1)	(2)	(3)	(4)	(5)
Δ ZCTA/MSA		0.0394***	0.019		-0.009
		(0.007)	(0.013)		(0.036)
linc	3.02E-05	0.0154**	0.009	0.008	0.007
	(0.006)	(0.007)	(0.008)	(0.014)	(0.022)
dpop			0.0002**		0.000
			(0.000)		(0.000)
lpopland	-0.0062***	-0.0050***	-0.0032**	-0.001	0.002
	(0.0014)	(0.0014)	(0.0014)	(0.0034)	(0.0031)
ldist	-0.0002	-0.0003	0.0010	0.0129**	0.0067
	(0.0025)	(0.0024)	(0.0022)	(0.0053)	(0.0048)
downocc			0.0002		0.0010
			(0.0002)		(0.0006)
ownocc	0.0001	0.0000	0.0002	0.0003	0.0007**
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0004)
centcity	-0.0012	0.0005	-0.0018	-0.0060	-0.0060
	(0.0026)	(0.0025)	(0.0024)	(0.0048)	(0.0046)
dbaplus			0.0004		-0.0004
			(0.0003)		(0.0006)
baplus	0.00019*	0.0001	0.0001	-0.0001	-0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0003)
dblack			-0.0003**		-0.0001
			(0.0001)		(0.0003)
nhblack	-0.0002***	-0.00013**	-0.0001	0.0005***	0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
dhisp			-0.0001		-0.0012***
			(0.0002)		(0.0003)
hisp	-0.0001	-0.0001	-0.0001	0.0000	-0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0003)
dpopkids			0.0001		0.0021**
			(0.0004)		(0.0010)
popkids	0.0002	0.0003	0.0001	0.0025***	0.0028***
	(0.0002)	(0.0002)	(0.0002)	(0.0005)	(0.0006)
dage65pl			-0.0016***		-0.0039***
			(0.0003)		(0.0007)
age65pl	-0.0011***	-0.0010***	-0.0018***	-0.0014***	-0.0033***
	(0.0002)	(0.0002)	(0.0003)	(0.0005)	(0.0006)
dforborn			0.0002		0.0006
			(0.0002)		(0.0004)
forborn	0.0001	0.0002	0.0002	0.0006**	0.0007**
	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0003)
hsg pre-1940	-0.0001	-0.0001*	-0.00014**	0.00025**	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Fixed effects:	MSA	MSA	MSA	MSA	MSA
Observations	6766	6766	6745	6730	6710
R-squared	0.175	0.199	0.214	0.14	0.173

Robust standard errors, clustered by place, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table A: Interactions between income and ZCTA, MSA characteristics

Dependent var:	Ln(Emp density, all retail)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
linc	0.385***	0.460***	0.404***	0.271***	0.385***	0.359***	0.296***
	(0.092)	(0.099)	(0.137)	(0.087)	(0.095)	(0.095)	(0.098)
linc*central city		-0.110					
		(0.077)					
linc*dist CBD			-0.010				
			(0.044)				
linc*MSA % mixed use				0.00176***			
				(0.0004)			
linc*MSA emp density					9.99E-08		
					(0.00001)		
linc*MSA pop						9.24e-09*	
						-5.1E-09	
lpopland	0.773***	0.774***	0.772***	0.773***	0.773***	0.772***	0.772***
	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.034)	(0.037)
ldist	-0.418***	-0.410***	(0.317)	-0.417***	-0.418***	-0.418***	-0.420***
	(0.053)	(0.054)	(0.458)	(0.052)	(0.054)	(0.053)	(0.056)
ownocc	-0.014***	-0.014***	-0.014***	-0.0138***	-0.014***	-0.0143***	-0.0152***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
centcity	-0.092**	1.087	-0.0922**	-0.0917**	-0.0922**	-0.0913**	(0.063)
	(0.040)	(0.834)	(0.040)	(0.039)	(0.040)	(0.040)	(0.042)
baplus	0.000	-0.001	0.000	0.000	0.000	0.000	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
nhblack	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.00728***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
hisp	-0.003**	-0.003**	-0.003**	-0.003**	-0.003**	-0.003**	-0.00278*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
popkids	-0.022***	-0.024***	-0.022***	-0.0222***	-0.0221***	-0.0216***	-0.0169***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
age65pl	0.016***	0.0155***	0.0155***	0.0151***	0.0156***	0.0158***	0.0189***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
forborn	0.001	0.001	0.001	0.001	0.001	0.000	0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
hsg pre 1940	-0.011***	-0.011***	-0.011***	-0.011***	-0.011***	-0.011***	-0.0132***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Fixed effects	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA	Yr & MSA
Other notes							Excludes NYC & LA
Observations	13542	13542	13542	13542	13542	13542	12441
R-squared	0.773	0.773	0.773	0.775	0.773	0.773	0.747

Robust standard errors, clustered by place, in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table B: Relationship between employment change and income, by industry

Dependent var:	Employment growth rate			Employment churn	
Timing of controls:	Levels only	Levels only	Levels & changes	Levels only	Levels & changes
	(1)	(2)	(3)	(4)	(5)
All retail					
Δ ZCTA/MSA		0.0394***	0.019		-0.009
		(0.007)	(0.013)		(0.036)
linc	0.000	0.0154**	0.009	0.008	0.007
	(0.006)	(0.007)	(0.008)	(0.014)	(0.022)
Groceries					
Δ ZCTA/MSA		0.0503***	0.0277**		0.009
		(0.009)	(0.012)		(0.028)
linc	-0.001	0.019	0.012	-0.010	-0.017
	(0.011)	(0.012)	(0.012)	(0.025)	(0.028)
Drugstores					
Δ ZCTA/MSA		0.0545***	0.0358***		0.000
		(0.007)	(0.011)		(0.027)
linc	0.0197*	0.0410***	0.0340***	-0.030	-0.028
	(0.011)	(0.011)	(0.013)	(0.022)	(0.025)
Clothing					
Δ ZCTA/MSA		0.0365***	0.015		-0.023
		(0.010)	(0.014)		(0.033)
linc	-0.006	0.009	-0.002	-0.0714***	-0.0825***
	(0.010)	(0.012)	(0.012)	(0.025)	(0.029)
Restaurants					
Δ ZCTA/MSA		0.0406***	0.012		0.024
		(0.007)	(0.009)		(0.024)
linc	-0.010	0.006	-0.005	-0.0473**	-0.032
	(0.008)	(0.009)	(0.009)	(0.020)	(0.021)
Laundry					
Δ ZCTA/MSA		0.0396***	0.015		-0.006
		(0.008)	(0.012)		(0.012)
linc	0.004	0.0196**	0.004	-0.005	-0.010
	(0.008)	(0.009)	(0.010)	(0.010)	(0.011)

Results shown are coefficients on log(income) and change in ZCTA income/MSA income. All regressions include controls for levels of: population density, distance to CBD, owner-occupied housing share, central city dummy, share with BA or graduate degree, share non-Hispanic black, share Hispanic, share under 18, share 65 and older, share foreign-born, share housing pre 1940. Columns 3 and 5 also include changes in control variables. All regressions include MSA fixed effects. N = 6,745.
Robust standard errors, clustered by place, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1