Article

Urban Studies

Urban Studies
1–36

DOI: 10.1177/0042098016661268
usj.sagepub.com
SAGE

Neighbourhood differences in retail turnover: Evidence from New York City

Rachel Meltzer
Milano School of International Affairs, Management and Urban Policy, USA

Sean Capperis
NYC Department of Housing Preservation and Development, USA

Abstract
Urban neighbourhoods are defined as much by their commercial character as their residential; retail services not only provide material needs for those living nearby, but less-tangible social and cultural capital as well. It is reasonable to expect, then, that excessive churn in these businesses can threaten the stability of a neighbourhood. Using a longitudinal data set on mixed-use neighbourhoods in New York City, we test whether or not neighbourhoods of varying circumstances and characteristics experience different degrees and types of retail turnover. Results suggest that there are meaningful differences in retail turnover across neighbourhoods. Retail turnover is directly associated with the type of business activity, commercial infrastructure and the neighbourhood's consumer profile. However, when all three sets of factors are considered simultaneously in a regression analysis, consumer-related characteristics explain turnover more than those related to the local commercial environment. Specifically, businesses that provide necessity and more frequently consumed goods/services are more stable and chain establishments are more likely to venture into markets with some housing price discounts, growth potential and possibly less organised opposition. Neighbourhoods with less (and more heterogeneous) general retail (as opposed to food service) concentration, as well as bigger businesses, are more stable. More importantly, bigger households and higher shares of white residents are most strongly associated with less retail churn, and population growth is the strongest predictor of more turnover.

Keywords
commercial markets, neighbourhood change, quality of life, retail, urban services

Received December 2014; accepted June 2016

Introduction

Urban neighbourhoods are defined as much by their commercial character as their residential. Indeed, neighbourhoods are a

Corresponding author:
Rachel Meltzer, Milano School of International Affairs, Management and Urban Policy, The New School, 72 Fifth Avenue, New York, New York, NY 10011, USA.
Email: meltzerr@newschool.edu

function not just of the people who inhabit them, but also the commercial transactions and investments that sustain local economies and activate street life. Retail services, particularly in mixed-use settings, not only provide material needs for those living nearby, but less-tangible social and cultural capital as well (Deener, 2007; Hyra, 2008; Zukin et al., 2009). Jane Jacobs (1961) famously argued that local small businesses are not only good for services and access to jobs, but are critical to the vitality of community life.[1] It is reasonable to expect, then, that excessive churn in these businesses can threaten the stability of a neighbourhood. Furthermore, this interruption could be even more severe under conditions of rapid neighbourhood change or gentrification. While neighbourhood change can bring in new amenities, filling much-needed gaps in local services, it can also introduce unpredictability in what and how these new businesses will serve the community (Ehrenhalt, 1999). Neighbourhood change research has focused extensively on the implications for residential stability (Ellen and O'Regan, 2008; Freeman, 2005; Freeman and Braconi, 2004; McKinnish et al., 2010); much less attention has been paid towards similar implications for local retail services. Patch (2008) suggests that retail change, or 'street gentrification', is a telling manifestation of broader neighbourhood transitions, and yet it has been less thoroughly documented. We aim to fill this gap here. Do neighbourhoods of varying circumstances and characteristics experience different degrees and types of retail turnover? Is this turnover exacerbated under conditions of gentrification?

The literature includes a host of case studies on retail change in particular neighbourhoods, across the globe (see, for example, Zukin et al., 2015), but there are many fewer studies looking systematically at how neighbourhood conditions relate to retail turnover. We present here a longitudinal analysis of microdata on over 1700 neighbourhoods in New York City in order to establish a more comprehensive baseline understanding of which factors explain retail turnover in urban neighbourhoods. New York City is a useful context for studying these issues, as it is comprised of many different types of neighbourhoods and submarkets, all of which have experienced different degrees of socioeconomic change over our study period. Our findings not only identify broad-based correlates of retail vulnerability and opportunity, but also serve as an important foundation for conducting future analyses on the mechanisms and effects of retail turnover. Furthermore, armed with more detailed information on neighbourhood retail markets (and the extent to which they manifest stability), local governments and civic organisations can better target neighbourhood economic development strategies towards both business owners and consumers.

Results suggest that there are meaningful differences in retail turnover across neighbourhoods. Retail turnover is directly associated with the nature of business activity, commercial infrastructure and the neighbourhood's consumer profile. However, when all three sets of factors are considered simultaneously, consumer-related characteristics explain turnover more than those related to the local commercial environment. Food establishments (i.e. restaurants) tend to be a more stabilising presence in neighbourhoods over time – they exhibit less churn compared with general retailers. Businesses that provide frequently consumed and necessity goods and services are more likely to stay in place (compared with those providing less frequently consumed or discretionary goods and services), suggesting that they are less vulnerable to local shocks in business-related costs and/or consumer demand. Chain establishments, compared with independent ones, are less likely to

open up new establishments in New York City, but are more likely to relocate across the city once they have penetrated the market. In addition, chains are more likely to enter neighbourhoods with more commercial space, lower residential vacancy rates, lower housing prices and higher-income households, and less likely to go into neighbourhoods with more owner-occupied homes and more college-educated residents. Neighbourhoods with a greater proportion of food establishments and a wider mix of retail types, as well as bigger businesses, are more stable. These factors, however, are not as important as consumer characteristics, which produce the largest standardised coefficients: bigger households and higher shares of white residents are most strongly associated with less retail churn while population growth is the strongest predictor of more turnover.

Background and literature review

In this paper we specifically focus on the circumstances around retail *turnover*; that is, the frequency and nature of how businesses move in and out of neighbourhood markets, and what those neighbourhoods look like. The literature on neighbourhood change sets up a dichotomy of production- and consumption-based processes, and we build off of this framework. Smith (1979), for example, prioritises the production-based aspect of neighbourhood change, where uneven development and the allocation of capital drive localised economic upgrading. Brueckner and Rosenthal (2009) empirically test for the role of physical reinvestment in neighbourhood change, and find significant (albeit partial) explanatory power. This is in contrast to the consumption-based perspective, where consumer preferences drive neighbourhood redevelopment (Bridge and Dowling, 2001; Ley, 1986; Zukin, 2008); this framework is particularly compelling in the context of retail services and amenities. Retail services, and the nature of their change, very much reflect both the consumption tendencies and cultural identities of local residents. Like Lees (1994) and Hamnett (1991), we hypothesise that retail turnover is both a production- and consumption-based phenomenon. In this section we set up this dichotomy theoretically, and then test it in the remainder of the paper.

We preface the following discussion on two reasonable assumptions. First, we consider neighbourhood retail services as a local amenity for primarily nearby residents; while these services can certainly service local workers as well, we focus on mixed-use communities rather than predominantly commercial ones. Second, we assume that retail generally follows households (or employees) and not vice versa; at the very least the two are correlational.

Why and when does neighbourhood retail change?

We rely on Hotelling's (1929) model of firm location decisions to motivate the implementation of our analysis; it reflects the production- and consumption-based framework established above. In its simplest form, his model conditions retail density on consumer density, store fixed costs and transportation costs. This suggests that in order to witness any *change* in retail density, one or all of these factors must undergo some shift.[2] In addition to the density of the consumer base, we consider a richer characterisation of the customer market, which accounts for consumption heterogeneity. In this scenario, the demographics of the local market can change; in turn, assuming business location decisions follow local demographic change (and not vice versa), the local business can adapt to these changing conditions and maintain its location, relocate to another

market that better supports its product, or permanently shut down. We expect that the neighbourhood demographics reflect not only consumption preferences, but also less-tangible cultural identities and biases (Ross, 1998).

Fixed costs can also change; here we walk through three ways in which they shift. First, the physical infrastructure or space constraints/opportunities can change over time. New investments in the neighbourhood can make space more appealing, affordable and functional for local retail establishments (in particular those that were not willing to occupy the spaces before), and this can draw particular businesses to a neighbourhood and drive others away. Second, information about the risks associated with operating in a particular neighbourhood can become more accessible over time.[3] For example, increases in activity from other establishments can signal a more hospitable business environment (especially in markets that are otherwise hard to read without very localised knowledge), lowering the entry risk for new businesses. And third, incentives (typically government-induced) can make particular locations more appealing or beneficial. These mediators can influence the behaviour of both producers and consumers. For example, tax incentives or zoning allowances may make it cheaper for the businesses to set up shop and/or compete in a market that they could not have otherwise entered or sustained. Food subsidies (attached to either the establishment or the consumer) can help to make local goods and services more accessible to a broader set of local consumers.

What is the nature of the change?

We also consider the nature of retail change. While overall turnover captures retail stability broadly, it obscures the nature of that turnover. A business can choose to stay, enter or exit a neighbourhood at any point in time. Those that stay can do so for short or extended periods of time. Those that enter can either relocate from another neighbourhood within the same municipality or open up a new establishment entirely. Likewise, those that exit a neighbourhood can relocate to another community within the same municipality or they can close down permanently. The different location decisions not only matter for a particular neighbourhood, but the citywide economy (since it is either a matter of reallocating businesses within the locality or losing/gaining businesses absolutely).

We also recognise that change can manifest itself differently for certain types of retail establishments, which in turn could have different implications for the neighbourhood. First, we distinguish between chain (i.e. multi-store businesses) and independent establishments. Since chain stores are typically more capitalised than independent operators, we anticipate that their turnover will be less pronounced (they can perhaps better withstand local shocks to avoid untimely shutdown or exit).[4] On the other hand, owing to this higher threshold for relocation (or shutdown), we might find their entry into neighbourhoods more selective (and perhaps less frequent) than independent entities. The chains will also be less vulnerable to threats of competition, since they likely capture the market share (because of pricing and/or breadth of service); this will also manifest itself in more stability. Second, we distinguish between necessity and discretionary goods and services. We predict that necessity services (such as groceries, drug stores and banks) will be less vulnerable to shocks to consumer demand than discretionary (or 'luxury') goods and services, which will be less patronised under conditions of economic duress or transition. Not only do discretionary services require more disposable income, but they also may be more particular to the idiosyncrasies of the local consumer base;

both can change in meaningful (and perhaps unpredictable ways) as the neighbourhood undergoes social and economic transitions.

Empirical literature review

Retail differences and location decisions. There is a small body of work that documents the retail differences across neighbourhoods (or markets) of varying consumer characteristics. These studies are presented separately from those that directly address retail change, as they tend to observe retail patterns at one point in time rather than shifts in local retail markets (and their correlates over time). Consistent with expectations, the findings generally demonstrate a correlation between the size and nature of local consumption markets and the size and composition of the local retail market (Berry and Waldfogel, 2003; Davis, 2006; Dinlersoz, 2004). One of these studies, by Waldfogel (2008), 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 behaviour, he finds that there is an association between the mix of locally available chain restaurants and demographic mix by race and education.[5]

To date, much of the research on neighbourhood disparities in commercial services comes from the public health literature. These studies focus on the differences in the locational decisions of establishments across neighbourhoods within a city. Kwate et al. (2013), Powell (2007), Zenk (2005), Bingham and Zhang (1997) and Alwitt and Donley (1997) demonstrate that various retailers, namely banks and supermarkets, opt not to locate in predominantly non-white and poorer ZIP codes even after controlling for purchasing power. Interestingly, Alwitt and Donley find that fast food restaurants are least likely to discriminate across neighbourhoods, whereas Block et al. (2004) and Sloane et al. (2005) find that fast food restaurants are more likely to locate in poorer, predominately black neighbourhoods.

Retail change. Another set of papers focuses specifically on retail *change*. Meltzer and Schuetz (2012) primarily conduct a cross-sectional analysis of retail access across New York City neighbourhoods, using publicly available ZIP code establishment and employment aggregates. They find that although high-income neighbourhoods in New York City have a higher density of retail employment and more chain restaurants, low-income and predominantly black or Latino neighbourhoods have a much higher share of unhealthy fast food restaurants. They also examine the change in retail presence over nearly one decade and correlate it with changes in residential property values. They find that between 1998 and 2008 the rate of retail growth was particularly rapid in neighbourhoods that were initially lower valued and experienced relatively high housing price appreciation compared with the city overall. This is confirmed by another study by Schuetz et al. (2012) that finds lower retail employment density (and smaller establishment size) among higher poverty (and lower income) neighbourhoods, driven largely by reduced employment in chain establishments. While neither income levels nor poverty rates consistently predict retail employment growth, neighbourhoods that experience income

upgrading do see larger gains in retail employment. Immergluck (1999) finds that neighbourhoods that are relatively more minority and less affluent experience declines in commercial investment, as measured by changes in permit activity.

Chapple and Jacobus (2009) use ZIP-code level data on retail businesses (from the National Establishment Time Series dataset) and Census tract-level data (from the Neighborhood Change Database) on neighbourhood economic and demographic characteristics for the San Francisco Bay area to examine the link between retail revitalisation and neighbourhood change. They classify neighbourhoods into five categories of relative income change and show with descriptive crosstabs that retail revitalisation is most strongly associated with gains for middle-income neighbourhoods. They hypothesise that this is, in part, due to their greater ability to attract start-up businesses. Zukin et al. (2009) conduct case studies of two gentrifying neighbourhoods in New York City and find a large increase in the number of independently owned (or local chain) establishments in those neighbourhoods, compared with a small increase in large chain stores. Bates and Robb (2008; 2014) test whether retail establishments are more or less likely to survive in urban, predominantly non-white neighbourhoods and they find that businesses serving predominantly people of colour are less profitable and more likely to close than those serving white customers. Carree and Thurik (1996) find evidence to support both demographic and infrastructure determinants of retail entry/exit. Specifically, retail businesses are more incentivised to enter markets with growing consumer spending and growing unemployment and new firm entry is inhibited by larger floor space requirements.

These studies are largely limited in the scope, detail and variation in retail change; those that do use fine-grained data do so in more limited case settings. Our analysis captures detailed retail turnover activity across a larger geographic area and a longer period of time.

Data

Our study takes place in New York City, a very useful context for studying these issues. It is a dense and diverse city comprised of many different types of neighbourhoods and retail markets. Furthermore, these neighbourhoods experienced varying degrees of socioeconomic changes over our study period, and this is variation that we can exploit in the analysis. Pressures from gentrification are not unique to New York City; local commercial districts from Toronto to Shanghai have been dealing with the challenges of rising rents and shifting demographics (Zukin et al., 2015). We also note that many New York City neighbourhoods are comparable to those in other large US cities. For example, while the median resident lives in a much denser neighbourhood than someone in an otherwise comparable city, the range of densities reflects those experienced in other large cities (Capperis et al., 2015). Typical education levels, unemployment rates, workforce participation rates and racial/ethnic make-ups are comparable with those in other large cities; incomes are also generally comparable, with the exception of slightly higher median household incomes and lower poverty rates (Been et al., 2013; Capperis et al., 2014). In addition, while New York City's overall homeownership rate is lower than that in other large US cities, homeownership rates are higher and closer to the norm in neighbourhoods in the boroughs of Queens and Staten Island (Been et al., 2013).

Our sample of businesses covers close to the universe of retail and food service establishments in New York City. The core data set for this analysis is the National

Establishment Time Series (NETS) database, a longitudinal, establishment-level data set that is constructed by Walls and Associates (2012) from the Dun & Bradstreet business register. Unlike publicly available government data on employment, the NETS data set includes no suppression of employment in small industry or geographic cells and provides full street address information for each establishment. In addition, NETS is more likely to capture non-employer businesses than other public records, such as those issued by the Bureau of Labor Statistics (Neumark et al., 2005).[6] We geocode these businesses' addresses to tax parcels so that we can accurately attach census tracts and then aggregate establishment and employment counts to obtain census tract totals. In addition, industry is reported at the 6-digit North American Industry Classification System (NAICS) level to allow for a fine-grained distinction across retail and food services, and information on whether or not the establishment is a headquarter, branch or standalone outlet to permit classifications according to firm structure. Finally, because the NETS data are longitudinal and establishment-specific, we can measure gross changes in the number of establishments and their employment (versus just net employment changes, which is what the publicly available ZIP code aggregates provide). We recognise the limitations with using NETS to identify very short-term changes in firm characteristics, and therefore process any changes over periods of five or more years (Neumark et al., 2005). This will also mitigate against any lags in the NETS data in observing new firm births (Yang and Aldrich, 2012). Furthermore, we note that the NETS data are less adept at capturing within-city moves (Kaufman et al., 2015); for the small part of our analysis that relies on this metric, we recognise that it captures a subset of the relevant neighbourhoods and likely underrepresents the actual within-city activity.

We collapse the business-level records into census tract aggregates and shares in order to capture turnover for the neighbourhood. Since we are concerned with the neighbourhood's exposure to turnover (rather than the business's), this aggregation makes sense. In order to ensure that we are looking at mixed-use neighbourhoods containing both businesses and resident consumers, we restrict the sample to include tracts with non-zero values for population, housing units and establishments.

We supplement the NETS data with a number of data sets, again, collapsed to the tract-level. We use New York City Department of Finance's tax assessment roll files and the New York City Department of City Planning's Primary Land Use Tax Lot Output (PLUTO) data set to observe changes in the amount of built commercial and residential space. We include building permits filed with the New York City Department of Buildings to measure the number of new buildings and residential units that developers intend to build (i.e. the degree of property investment). To measure changes in housing prices, we obtain residential property sales data from the New York City Department of Finance. Prices (as well as incomes and rents, which are obtained from Census and American Community Survey data described below) are adjusted for inflation using the Consumer Price Index for the New York metropolitan area and expressed in 2012 dollars.

In addition, we merge in tract-level economic and demographic variables from various sources. Geolytics' Neighborhood Change Database (Geolytics, 2003) provides data every 10 years from 1980 through 2000, normalised to consistent census tracts as defined in the 2000 Census. We supplement these data with indicators from the 2010 Census and the American Community Survey's five-year estimates from 2007 to 2011 where Census data are unavailable, which we assume represent

2010 conditions. We also normalise data from these sources to tract definitions set in 2000 to be geographically compatible with measures from the Neighborhood Change Database. Ultimately, we have 2137 census tracts in our sample, spanning 20 years.[7]

Throughout the analysis, we are constrained by the fact that Census-based data are available only every decade, and we adjust the data merging accordingly. In analyses where we observe neighbourhoods every ten years, such as in the cross-tab analyses, we rely on the decennial values from the Census. In analyses where we observe neighbourhoods every five (or more) years, such as in the regression analyses, we compute values for intercensal years (i.e. 1995 and 2005) by assigning the average value from the previous and subsequent year. When a year falls in between the five-year increment (i.e. is not a multiple of five), we then assign Census variables based on the increment of five that mostly closely precedes the year in which the neighbourhood is observed (for example, if we observe a neighbourhood in 1999 we assign Census variables as of 1995; that is, the average of values in 1990 and 2000). In all cases, Census variables are assigned contemporaneously or lagged, in order to mitigate against endogeneity.

Analytical strategy

Since our analysis is focused on *neighbourhood retail* services, we rely on two important analytical features. First, the unit of analysis is the neighbourhood-year. We operationalise neighbourhoods as census tracts as defined in the 2000 Census, which is an area optimally populated by 4000 people (US Census Bureau, 2012). Previous studies have used the census tract to capture neighbourhood communities and markets (Ellen and O'Regan, 2008; McKinnish et al., 2010); it also allows for more fine-grained analysis than the ZIP code (the finest level at which business data are made publicly available) and at a level at which socio-demographic information is readily available over time. The census tract also captures a walkable market area, which, on average, can be traversed in five to ten minutes.[8] While residents could certainly walk farther to access local retail, we aim here to primarily capture the commercial environment in their immediate vicinity and therefore return to the usefulness of the tract operationalisation. We limit our sample to include sectors that include businesses that serve neighbourhoods (versus central business district businesses or manufacturing enterprises). Specifically, we include businesses 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, fitness, barber/beauty shops, laundry and pet care; see Appendix A for a full listing of included sectors. Most of the sample (about 84%) is categorised as general retail trade or services.

Variable construction

We construct a number of variables in order to test the hypotheses set forth in the section 'Background and literature review'. First, we create six measures to assess the movement of businesses into and out of neighbourhoods over time: *stay, entry, exit, birth, death* and *churn*. We define *stay* as the number of businesses that stay in neighbourhood j for the entirety of the change interval, $(t-k,t)$. We define *entry* as the number of businesses that (i) enter neighbourhood j from another neighbourhood i, still within the city's border, any time during the change interval, $(t-k,t)$, and (ii) are still in operation in neighbourhood j at time t. *Exit* is the opposite: the number of businesses that (i)

are in operation in neighbourhood j at time $t-k$, and (ii) exit neighbourhood j to another neighbourhood i, still within the city's border, any time during the change interval, $(t-k,t)$. *Birth* is the number of businesses that (i) enter neighbourhood j as brand new entities in the city (i.e. not previously located in another neighbourhood) any time during the change interval, $(t-k,t)$, and (ii) are still in operation in neighbourhood j at time t. *Death* is the opposite: the number of businesses that (i) are in operation in neighbourhood j at time $t-k$, and (ii) exit neighbourhood j to permanently close down any time during the change interval, $(t-k,t)$.

Since neighbourhoods can contain a wide range of retail densities, and comparing counts can be misleading, we also calculate the share for each of these measures. They are calculated relative to the total number of establishments in the tract at the beginning of the change interval, $(t-k,t)$. Since there are neighbourhoods with small counts at the beginning of the intervals, we calculate the share using the midpoint method (as an example, we display the calculation for the *entry* variable):

$$Entry\_share_{t-k,t} = \frac{Entry_{t-k,t}}{(Estab_{t-k} + Estab_t)/2} \quad (1)$$

This approach mitigates against inflated shares, due simply to low baselines at the start of the interval.[9] We also create a measure of *churn*, which, for neighbourhood j, is the sum of all possible moves (either into or out of neighbourhood j) divided by the average of the total number of businesses at time $t-k$ and the total number of businesses at time t, consistent with the midpoint method:

$$Churn_{t-k,t} = \frac{(Entry_{t-k,t} + Exit_{t-k,t} + Birth_{t-k,t} + Death_{t-k,t})}{((Estab\_total_t + Estab\_total_{t-k})/2)} \quad (2)$$

We prioritise churn as our summary measure of retail change, as it captures the overall volatility of retail for a particular neighbourhood over time. However, the other five measures will help to disentangle the nature of the churn and provide a more nuanced sense of how businesses flow into and out of the neighbourhood.

We also classify businesses in several ways that relate to local consumer density and the business' fixed costs. First, we categorise each business as either *necessity* or *discretionary*. *Necessity* establishments are those that fulfill more 'everyday' needs (and therefore are likely less vulnerable to shocks in consumer willingness, such as income-induced ones) or are providing for the 'immediate needs of people' (Bingham and Zhang, 1997; Stanback et al., 1981). Stanback et al. (1981) described these as 'residentiary services'. Like Bingham and Zang (1997) we also include a few 'producer services', such as banks, in our definition of retail, since they also provide essential services to local residents. Other examples of subsectors in this category are groceries, drug stores and household goods stores. *Discretionary* establishments, on the other hand, provide more luxury or recreational services or goods that are not considered basic, but certainly enhance quality of life. Examples of subsectors in this category are liquor stores, most restaurants and beauty salons. About half of the businesses in our sample are classified as discretionary, one-third as necessity and, the small remainder, as establishments that provide durable goods (such as car dealers or furniture stores).[10]

We further disaggregate by classifying businesses as providing goods/services that are either *frequently* or *infrequently* consumed. 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. This is a meaningful distinction, also, because it further identifies the types of businesses that contribute to the daily quality of life in the neighbourhood. Examples of subsectors offering frequently consumed goods/services are food (both grocery stores and restaurants), pharmacies, service stations, discount and department stores, banks and laundry; some subsectors with less frequently consumed goods/services include stores offering furniture, housewares, clothing, sporting goods and media (these are the kinds of businesses that also might have market share outside the local neighbourhood). Appendix A demonstrates how NAICS codes align within this typology. There is quite a bit of overlap between necessity and frequency, but discretionary businesses exhibit substantial variation. Together, these classifications allow us to set up a hierarchy of local services, such that *frequently consumed necessity* goods/services are perceived as fundamental to neighbourhood wellbeing, and *infrequently consumed discretionary* goods/services would represent local 'luxuries'.[11]

Finally, we distinguish among businesses based on their organisational structure. We classify a business as *chain* if it is linked to a separate headquarters establishment, if it is itself identified as a branch or headquarters, or if at least one other establishment reports to the same headquarters. There is no minimum for the number of establishments that constitute a chain. We classify a business as *independent* if it does not meet any of the criteria for classification as a chain.

Results

In this section we test whether or not retail turnover varies with dimensions of consumer density and business structure and/or costs.[12] We do this through a series of descriptive cross-tabulations and then multivariate regression analyses to assess which factors are most important in explaining retail turnover. We present here results for changes in retail activity over five-year intervals; therefore, for each year t in the sample, turnover variables are calculated for the interval ($t-5$, t). We thought this a reasonable interval, since the median neighbourhood tenure of a business in the city is five years. For businesses that eventually relocate to another neighbourhood, the average length of stay is just over five years; for those that ultimately close their doors permanently, the average length of stay is about 4.5 years.[13]

As a first cut, we look at broad citywide and borough-wide patterns of retail turnover (see Table 1). The upper panel displays counts and the lower panel displays the corresponding shares. Overall, we see that businesses are more likely than not to stay in place in the same neighbourhood over a five-year interval (as indicated by the stay rate of 0.57). Birth rates (and entry rates) are higher than death rates (and exit rates) and, in general, neighbourhoods experience similar retail churn rates, with the exception of those in Manhattan and Staten Island, which have slightly lower rates. The shares of establishments that stay in place are consistent across the boroughs, as are the shares of deaths. The inter-neighbourhood movement, however, varies. For example, the Bronx experiences the lowest entry rate and Staten Island the highest. Manhattan exhibits the lowest churn and the Bronx the highest. Therefore, there is spatial variation to exploit. These results, however, do not tell us where businesses are relocating (or from where they are relocating), and so we cannot assign any correspondence between the borough rates of entry or exit.

We also see that looking at counts alone can be misleading. For example, while Manhattan has the highest *number* of

Table 1. Retail turnover variables by borough.

	NYC	Bronx	Brooklyn	Manhattan	Queens	Staten Island	Sig. Diff.
Mean number of est.							
Stay	23.7	15.95	17.13	65.74	16.44	18.91	***
Birth	16.65	11.72	12.35	43.29	12.42	12.42	***
Enter	0.6	0.19	0.34	2.228	0.349	0.651	***
Death	15.4	11.87	11.28	40.15	11.12	10.91	***
Exit	0.618	0.209	0.328	2.517	0.292	0.543	***
Mean share of est., t−5 (midpoints)							
Stay	0.574	0.563	0.572	0.601	0.567	0.595	***
Birth	0.428	0.427	0.431	0.403	0.440	0.409	***
Enter	0.015	0.009	0.015	0.018	0.015	0.024	***
Death	0.397	0.430	0.399	0.359	0.401	0.357	***
Exit	0.012	0.008	0.011	0.019	0.011	0.020	***
Churn	0.852	0.875	0.856	0.799	0.866	0.810	***
N	3852	627	1378	545	1116	186	***

Notes: As an example, we interpret the results for NYC: in the average tract, 24 (57.4%) establishments stayed in place between time $t-5$ and t; nearly 17 new establishments were born (42.8% as a share of establishments in $t-5$), 15 existing establishments died (39.7%), one moved into the tract from elsewhere in the city (1.5%), and another moved to another tract in the city (1.2%); between time $t-5$ and t, 85.2% of the establishments (open at time $t-5$) turned over in some form (some combination of births, deaths, or moves in or out). All differences in means are statistically significant at $p < 0.01$.

births during the 5-year interval, it has the smallest share. For this reason, we will display and discuss only the share variables from now on.

Business type and structure

Next, we look at whether retail turnover varies by business type and structure. These are displayed in Tables 2 through 4. We first stratify by super-sector (food versus general retail establishments) and find some consistent differences (see Table 2). Food establishments tend to exhibit less churn, which is driven by smaller rates of birth, death, entry and exit (most dramatically, births). This is also exhibited by the meaningful difference in the share of establishments that stay in place over the 5-year intervals (it is 0.64 for food establishments compared with 0.55 for general retail). The next table (Table 3) displays turnover variables for necessity and discretionary establishments. Necessity retailers exhibit slightly less turnover (i.e. a lower churn rate), and this appears to be driven by fewer births and slightly more businesses that stay in place. While the death rate for necessity services exceeds that for discretionary ones, any losses seem counteracted by the businesses that stay on and by fewer exits due to relocation.

To more precisely capture the type of retail demand, we further disaggregate the necessary and discretionary goods by frequency of use (these results are displayed in Table 3, as well). These analyses show that the necessity services that are more frequented are more likely to stay and contribute to less overall churn (more infrequently consumed necessity services, on the other hand, demonstrate a lot of flux, especially from both birth and exit activities). The same patterns hold for frequently consumed discretionary services compared with those less frequently consumed. Overall, it is consistent with expectations that necessity goods and services would be less vulnerable to local shocks in demand because of their broader

Table 2. Retail turnover variables by business type (food-retail).

	N	Food establishments	Retail establishments	Sig. diff
Share of est., t−5 (midpoints)				
Stay	3847	0.635	0.554	***
Birth	3847	0.318	0.454	***
Enter	3847	0.012	0.015	**
Death	3847	0.391	0.409	**
Exit	3847	0.010	0.013	**
Churn	3852	0.731	0.892	***

Notes: As an example, we interpret some of the cells: in the average tract, 63.5% of food-based establishments and 55.4% of retail-based establishments stayed in place between time $t-5$ and t. The difference is largely due to a higher birth rate among retail establishments (45.5% for retail; 31.8% for food).
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 3. Retail turnover variables by business type (necessity-discretionary; frequent-infrequent).

	Overall			Necessity			Discretionary		
	Necessity	Discre-tionary	Sig. diff.	Frequent	Infrequent	Sig. diff.	Frequent	Infrequent	Sig. diff.
Mean share of est., t−5 (midpoints)									
Stay	0.567	0.564	**	0.590	0.415	***	0.583	0.462	***
Birth	0.427	0.449	***	0.402	0.606	***	0.406	0.598	***
Enter	0.015	0.016		0.011	0.030	***	0.016	0.020	**
Death	0.413	0.394	**	0.397	0.510	***	0.401	0.440	***
Exit	0.011	0.013	*	0.009	0.024	***	0.011	0.018	***
Churn	0.866	0.872	**	0.818	1.170	***	0.834	1.077	***
N	3800	3785		3752	3291		3775	3639	

Notes: As an example, we interpret some of the cells: in the average tract, 86.6% of retail necessities and 87.2% of discretionary establishments operating in time $t-5$ experience some sort of churn between $t-5$ and t. The higher churn among discretionary businesses occurs because of higher average birth, move-induced entry, and move-induced exit rates and appears in spite of a lower average death rate.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

appeal and relatively more frequent consumption; discretionary services should be consumed less under conditions of economic duress or demographic change.

Third, we distinguish independent retailers from chain establishments and replicate the turnover variables (see Table 4). Overall, independent businesses turn over more than the chain establishments. While their probabilities of staying in the same neighbourhood over the 5-year intervals are marginally different, the birth rate of independent businesses is about one-third higher than that for chain establishments. That said, movement across neighbourhoods is more likely for chain establishments, as evinced by the higher entry and exit rates. These findings are consistent with the notion that chains have higher start-up costs (they need bigger spaces) and perhaps a higher threshold for entering a market. However, once they have a presence in the local market, they are more

Table 4. Retail turnover variables by business structure (independent-chain).

	Independent	Chain	Sig. diff.
Mean share of est., t−5 (midpoints)			
Stay	0.568	0.597	***
Birth	0.435	0.359	***
Enter	0.014	0.029	***
Death	0.402	0.389	
Exit	0.012	0.028	***
Churn	0.863	0.805	***
N	3850	2499	

Notes: As an example, we interpret some of the cells: in the average tract, 86.3% of independent establishments and 80.5% of chain establishments operating in time $t-5$ experience some sort of churn between $t-5$ and t. The higher churn among independent businesses occurs because of higher average birth and death rates and appears in spite of lower average move-induced entry and exit rates.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

footloose and move more frequently across neighbourhoods. This suggests then that the bigger fixed costs are attached to the initial market penetration; subsequent moves may be less costly because of newly obtained, on-the-ground information and local political and institutional connections.

To try to understand the types of neighbourhoods that attract chains, we conduct a subanalysis on the neighbourhoods at the point in time at which they receive their first chain. To do this, we conduct a survival analysis of the neighbourhoods that get their first chain during the study period, and compare their characteristics with the neighbourhoods that receive any non-chain new establishments (either from a move or a birth). The neighbourhoods without any chains by the end of the study period, 2010, will be considered 'censored' observations. We use a Cox model with non-proportional hazards to estimate the likelihood of a neighbourhood i getting its first chain between t and Δt, given that it does not yet have a chain establishment by time t (this is also known as the hazard rate, $h_i(t)$).[14] The equation to be estimated is:

$$h_i(t) = \lambda_0(t) exp\left(Commercial_{i\,t-5}, Consumer_{i,\,t-5}\right) \quad (3)$$

In this regression, $\lambda_0(t)$ is the baseline hazard function, i.e. the hazard function for a neighbourhood with all covariates set to 0. $\boldsymbol{Commercial_{i,t-5}}$ is a vector of commercial environment characteristics, including overall establishment density, the share classified as retail and as necessity, the average establishment size, an index of retail diversity and the amount of dedicated commercial space. $\boldsymbol{Consumer_{i,t-5}}$ is a vector of household and housing market characteristics, including race/ethnicity, income, education, poverty status, unemployment, household type and size, homeownership status, foreign born status, vacancy rates, housing prices/rents. Note that we lag these covariates by five years, to mitigate against threats of endogeneity.[15]

The results from this analysis are displayed in Table 5; the first column displays only commercial environment characteristics, the second column only demographics and housing market characteristics, and the final column the full specification with both sets of covariates. We see that there are characteristics of the commercial environment that predict an earlier chain entry. Chain establishments are more likely (i.e. exhibit a hazard rate greater than one) to first enter into neighbourhoods with more businesses

Table 5. First-chain neighbourhoods.

	(1) Commercial	(2) Demographic/ economic	(3) All
Lag estab. density	1.002***		1.002***
	(0.000721)		(0.000500)
Lag avg. employees/estab.	1.065***		1.046***
	(0.00888)		(0.0123)
Lag retail (%)	0.669		0.721
	(0.298)		(0.301)
Lag necessity (%)	2.988***		2.247**
	(0.848)		(0.778)
Lag herfindahl index	0.503*		1.058
	(0.198)		(0.335)
Lag estab. density*herfindahl	0.982***		0.987***
	(0.00468)		(0.00355)
Lag log commercial area (2005)	1.388***		1.388***
	(0.0533)		(0.0694)
Lag pop. density		1.000***	1.000
		(2.86e–06)	(2.45e–06)
Lag population change (%)		1.441	1.338
		(0.608)	(0.522)
Lag vacancy (%)		0.000560***	0.00104***
		(0.00121)	(0.00239)
Lag homeownership (%)		0.0360***	0.0712***
		(0.0272)	(0.0532)
Lag non-family hhlds. (%)		1.216	0.347
		(1.790)	(0.487)
Lag average persons/hhld.		1.329	1.140
		(0.331)	(0.268)
Lag. black pop. (%)		0.660	0.779
		(0.439)	(0.475)
Lag. Hispanic pop. (%)		0.171**	0.238**
		(0.136)	(0.174)
Lag white pop. (%)		2.741	3.710*
		(2.259)	(2.897)
Lag foreign born (%)		0.298*	0.350*
		(0.194)	(0.218)
Lag poverty (%)		0.177	0.0912**
		(0.203)	(0.103)
Lag real med. hhld. inc.		1.000***	1.000***
		(8.39e–06)	(7.42e–06)
Lag ratio avg./med. hhld. inc.		2.350***	2.767***
		(0.777)	(0.821)
Lag unemployment (%)		0.825	4.164
		(1.115)	(6.213)
Lag 4-year degree (%)		0.0130***	0.0163***
		(0.0133)	(0.0164)
Lag real gross rent		0.999**	0.999***
		(0.000387)	(0.000352)
Lag log median price/unit		0.252***	0.290***
		(0.0308)	(0.0300)
Observations	1015	1015	1015

Notes: Hazard ratios (robust exponentiated standard errors) shown.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

and commercial space overall and with bigger retail establishments on average. Both of these findings suggest that the nature of the local infrastructure matters: chains often need larger spaces and are more likely to enter into areas that already have some degree of commercial activity (both in terms of the transactions and the permitted use of the land). However, chains are slightly less likely (i.e. exhibit a hazard rate less than one) to first enter into denser neighbourhoods that are also more homogeneous in their retail activity. Chain establishments are also more likely to open up for the first time in neighbourhoods with a higher share of necessity services.[16]

As for residential characteristics, the strongest predictor of first-chain entry is the neighbourhood's ratio of average to median household income (a measure of the neighbourhood's income distribution). Specifically, more positively skewed income distributions are associated with a higher likelihood of chain entry; in other words, the presence of very affluent households invites chain openings. This is consistent with the lower hazard rate associated with higher poverty rates in the fully specified model. Other covariates are also associated with lower probabilities of first-chain entry: higher residential vacancy rates, higher homeownership rates and higher shares of Hispanics, residents with a college degree and, marginally, foreign born. Higher rents and prices are also associated with lower probabilities of chain entry (though the former is very marginal at 0.999). We note that, controlling for other residential and commercial characteristics, neither population density nor population change is associated with first-chain entry. The results imply both structural and socio-political reasons for chain entry. First, neighbourhoods with more commercial space, lower residential vacancy rates and lower prices are more likely to see chains enter – this suggests economic barriers (or incentives) to entry, whereby chains venture into markets with some price discounts and growth potential. Second, the fact that chains are less likely to first enter into neighbourhoods with more owner-occupied and educated households could be a product of organised efforts from those groups against chain establishments (Healy, 2012; Pristin, 2009; Schuetz, 2015).[17]

Commercial environment

In order to understand the influence of the local commercial infrastructure (i.e. supply-side factors) in guiding retailers' location decisions, we stratify the sample by the share of commercial building space dedicated to retail use. We identify commercial and retail space by the actual (and intended) floor area used for these purposes, both vacant and occupied (commercial includes but is not exclusive to retail). In reality, actual use could deviate from the commercial uses permitted by zoning (although the two are very closely related and discrepancies between actual commercial use and zoning classification take place in about only 20% of buildings in our sample). Specifically, the retail share of commercial space is designated as 'high' if the share is more than the 75th percentile, and 'low' otherwise. These results are displayed in Table 6 and they show that neighbourhoods with relatively less retail space experience more churn. These patterns are driven by more entries and births in those areas. While these neighbourhoods might have less infrastructure to support commercial activity, it might be also the case that the lower concentration of retail use means lower rents and untapped markets.

Neighbourhood demographics

Finally, we consider neighbourhood demographics. As discussed earlier, the socioeconomic characteristics of the neighbourhood

Table 6. Retail turnover variables by neighbourhood commercial space.

	Retail share of commercial area		
	⩽ 75th percentile	> 75th percentile	Sig. diff.
Mean share of est., $t-5$ (midpoints)			***
Stay	0.552	0.580	***
Birth	0.474	0.438	***
Enter	0.015	0.010	
Death	0.396	0.382	
Exit	0.011	0.010	***
Churn	0.896	0.839	***
N	1433	505	

Notes: As an example, we interpret some of the cells: establishments in tracts with less retail space as a share of commercial building area (as well as total building area, not shown here) tend to experience more churn. In tracts where the retail share of commercial building area is less than the 75th percentile of all tracts, an average of 89.6% of establishments operating in time $t-5$ experience some sort of churn between times $t-5$ and t. This is in contrast to 83.9% of establishments operating in neighbourhoods with more retail area in time $t-5$.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 7. Retail turnover variables by neighbourhood income/poverty.

	Tract median household income			Poverty rate		
	⩽ City median	> City median	Sig. diff.	⩽ 75th percentile of tracts	> 75th percentile of tracts	Sig. diff.
Mean share of est., t−5 (midpoints)						
Stay	0.556	0.582	***	0.581	0.553	***
Birth	0.444	0.421	***	0.420	0.450	***
Enter	0.011	0.017	***	0.016	0.010	***
Death	0.424	0.385	***	0.388	0.424	***
Exit	0.010	0.013	***	0.013	0.009	***
Churn	0.888	0.835	***	0.838	0.894	***
N	1208	2644		2883	969	

Notes: As an example, we interpret some of the cells: on average, 88.8% of establishments in tracts with a median household income below the city median and 83.5% of establishments in higher-income tracts operating in time $t-5$ experience some sort of churn between $t-5$ and t.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

can reflect actual consumer demand and can also signal to potential retailers about local consumer activity. We recognise that our data cannot entirely capture the consumer activity on the ground, as much of it (especially in lower income communities) is not recorded in the data. However, we rely on data that are publicly and systematically available, and therefore likely available to businesses making their location decisions. The results for this part of the analysis are displayed in Tables 7 through 9. First, we consider income-related measures. We compare neighbourhoods with median household incomes below the citywide median to those with median household incomes above the citywide median (see Table 7). Relatively higher-income neighbourhoods have a

higher share of businesses that stay in place during the five-year interval and lower churn overall. While they lose a higher share of businesses to relocations, they have a lower death rate. This same table also displays retail turnover variables for relatively high and low poverty neighbourhoods.[18] These results are consistent with those for the income-stratified neighbourhoods – movement patterns for lower poverty tracts echo those for higher-income tracts.

We supplement this income analysis by also stratifying the turnover variables by housing prices and rents. These are good measures of what people are willing to pay to reside in a particular neighbourhood (which should be a function of the local retail services). We use the citywide mean as a benchmark for rents and housing prices. First we look at the turnover variables in neighbourhoods with rents that are typically lower than the citywide median and compare them with the outcomes for neighbourhoods with rents that are above the citywide median. The statistics (displayed in Table 8) show that churn is significantly higher in the relatively lower rent areas. This is primarily driven by the higher shares of business deaths and, somewhat less, births. Next, as a way to operationalise gentrification, we classify neighbourhoods by housing price appreciation over the change interval and, again, peg the neighbourhood's prices changes to the average change for the city overall.[19] In neighbourhoods that experienced price changes that exceeded those of the city overall, i.e. gentrification, we see significantly higher churn, driven primarily by births, compared with neighbourhoods with price changes lower than the citywide mean. The higher share of births in the context of appreciating prices could either mean that the commercial space was not fully saturated or that new space has been built in those areas.[20] This is consistent with subanalyses that look at business turnover across neighbourhoods with relatively more or less retail growth: neighbourhoods with faster retail growth experience more churn, because of higher entry/birth rates.[21]

Table 8. Retail turnover variables by neighbourhood housing prices and rents.

	Tract median gross rent relative to city median			Tract change in housing prices relative to city change		
	⩽ City median	> City median	Sig. diff.	Stable or declining	Increasing	Sig. diff.
Mean share of est., t−5 (midpoints)						
Stay	0.558	0.586	***	0.582	0.571	***
Birth	0.442	0.418	***	0.409	0.436	***
Enter	0.012	0.017	***	0.014	0.015	
Death	0.419	0.381	***	0.401	0.395	
Exit	0.011	0.013	***	0.013	0.012	*
Churn	0.883	0.829	***	0.837	0.858	***
N	1615	2232		1234	2274	

Notes: As an example, we interpret some of the cells: on average, 88.3% of establishments in tracts with a median gross rent below the city median and 82.9% of establishments in higher rent tracts operating in time $t-5$ experience some sort of churn between $t-5$ and t. Additionally, establishments in tracts with median housing prices increasing faster than the city's overall rate experience more churn than tracts with stable or declining housing prices.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 9. Retail turnover variables by neighbourhood race/ethnicity.

	Predominantly black	Predominantly Hispanic	Predominantly white
Mean share of est., t−5 (midpoints)			
Stay	0.547	0.546	0.609
Birth	0.442	0.460	0.388
Enter	0.010	0.010	0.021
Death	0.445	0.427	0.355
Exit	0.009	0.010	0.017
Churn	0.905	0.907	0.782
N	711	527	1142

Notes: As an example, we interpret some of the cells: of establishments operating in time $t-5$, on average, 90.5% in predominantly black tracts, 90.7% in predominantly Hispanic tracts, and only 78.2% in predominantly white tracts experience some sort of churn between $t-5$ and t.
All differences in means, relative to predominantly white, are statistically significant at $p < 0.01$.

Finally, we consider different racial/ethnic compositions (see Table 9); here we compare neighbourhoods designated as predominantly non-Hispanic white to those designated as predominantly non-Hispanic black or Hispanic. We define a race/ethnicity category as predominant if it comprises more than 60% of the tract's total population. Predominantly white neighbourhoods experience lower retail churn (more stability) overall compared with predominantly black and predominantly Hispanic neighbourhoods. This is primarily due to lower deaths and births, and more businesses that stay in place. However, predominantly white neighbourhoods receive and lose more businesses because of inter-neighbourhood relocation.

Multivariate analysis

In an attempt to discern which set of factors (commercial environment or consumer-related) is most influential in explaining retail turnover, we run multivariate regressions. Following the assumption made above that residential/consumer shifts will precede, and therefore help to explain, retail turnover, we specify the OLS model in the following way:

$$Estab\_Turnover_{i,p,t} = \beta_0 + \beta_1(\boldsymbol{Commercial}_{i,t-5}) + \beta_2(\boldsymbol{Gentrify}_{,t,t-5}) + \beta_3(\boldsymbol{Residential}_{i,t-5}) + d_t + d_p + \varepsilon_{it} \quad (4)$$

Here, *Estab_Turnover* is a turnover variable, including *Stay_share, Birth_share, Enter_share, Exit_share, Death_share*, and *Churn*. The vectors, ***Commercial*** and ***Residential***, include the variables capturing the commercial environment and household characteristics (those same ones itemised previously for the hazard analysis). And ***Gentrify*** controls for whether or not the neighbourhood had lower housing prices than the citywide median in time $t-5$ and whether or not it experienced price/rent appreciation over the period ($t-5$, t), relative to the city as a whole.[22] We also include year dummies, public use microdata area (PUMA)[23] fixed effects and clustered standard errors. So that we can better compare the relative impacts of each covariate, we display standardised coefficients.[24] These results are displayed in Tables 10–13.

We start with models that include only characteristics of the commercial environment, displayed in column 1 of Table 10. First, we consider the churn outcomes, and

Table 10. OLS regression of turnover variables on commercial characteristics.

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag estab. density	−0.0284	0.0142	−0.0297**	−0.00362*	0.00474	0.000146
	(0.0190)	(0.00952)	(0.0141)	(0.00207)	(0.00766)	(0.00169)
Std. lag rel. estab. % change	0.00408	−0.00204	−0.0201***	0.00172	0.0223***	0.000206
	(0.00657)	(0.00328)	(0.00537)	(0.00153)	(0.00313)	(0.000596)
Std. lag avg. employees/estab.	−0.0386**	0.0193**	−0.0411***	0.00505**	−0.00239	−0.000151
	(0.0160)	(0.00799)	(0.0116)	(0.00232)	(0.00604)	(0.00156)
Std. lag retail (%)	0.0325***	−0.0162***	0.0178***	0.00432***	0.0103**	7.56e−05
	(0.00905)	(0.00452)	(0.00559)	(0.00127)	(0.00402)	(0.00176)
Std. lag necessity (%)	−0.0124	0.00618	−0.00260	−0.00343***	−0.00383	−0.00252**
	(0.00793)	(0.00397)	(0.00597)	(0.00128)	(0.00402)	(0.00111)
Std. lag independent (%)	−0.00233	0.00116	−0.00591	0.00113	0.00380	−0.00135
	(0.00631)	(0.00316)	(0.00524)	(0.00129)	(0.00398)	(0.000936)
Std. lag herfindahl index	0.0514***	−0.0257***	0.0742***	0.00568***	−0.0287***	0.000200
	(0.0123)	(0.00616)	(0.00879)	(0.00185)	(0.00519)	(0.00222)
Std. lag estab. density*herfindahl	−0.0320	0.0160	−0.0407*	−0.00154	0.00812	0.00215
	(0.0205)	(0.0102)	(0.0234)	(0.00206)	(0.0109)	(0.00252)
Std. lag log commercial area (2005)	−0.00303	0.00151	0.0128*	−0.00334**	−0.0128***	0.000221
	(0.00903)	(0.00452)	(0.00645)	(0.00138)	(0.00423)	(0.000654)
Std. 2005	−0.00116	0.000579	0.0568***	0.00412***	−0.0635***	0.00137***
	(0.00283)	(0.00141)	(0.00307)	(0.000652)	(0.00279)	(0.000418)
Std. 2010	0.0216***	−0.0108***	0.0321***	0.000116	−0.00901***	−0.00153***
	(0.00292)	(0.00146)	(0.00342)	(0.000529)	(0.00269)	(0.000430)
Constant	0.867***	0.567***	0.457***	0.0188***	0.376***	0.0144***
	(0.00341)	(0.00171)	(0.00351)	(0.000541)	(0.00290)	(0.000470)
Observations	5596	5596	5596	5596	5596	5596
R-squared	0.061	0.061	0.192	0.047	0.231	0.015
Number of PUMAs	55	55	55	55	55	55

Notes: All outcome variables measure changes between t and $t-5$ and are expressed as shares of establishments over the average of t and $t-5$. All specifications include PUMA fixed effects. 'Rel. estab. % change' refers to percent change in number of establishments in a tract between time t and $t-5$ relative to the change citywide. Establishment density is calculated per square mile of land. Robust standard errors in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 11a. OLS regression of turnover variables on gentrification variables (prices).

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag low price/unit	0.00565	−0.00283	0.0106**	−0.00308***	−0.00150	−0.000393
	(0.00532)	(0.00266)	(0.00401)	(0.00102)	(0.00313)	(0.000620)
Std. lag low price/unit* rising price/unit	0.00147	−0.000736	−0.00215	−0.000421	0.00413	−9.05e−05
	(0.00413)	(0.00207)	(0.00329)	(0.000517)	(0.00276)	(0.000500)
Std. 2000	−0.0402***	0.0201***	−0.0784***	0.00199***	0.0347***	0.00148***
	(0.00359)	(0.00180)	(0.00477)	(0.000444)	(0.00355)	(0.000407)
Std. 2005	−0.0429***	0.0214***	−0.0235***	0.00583***	−0.0281***	0.00293***
	(0.00399)	(0.00199)	(0.00376)	(0.000759)	(0.00194)	(0.000380)
Std. 2010	−0.0241***	0.0121***	−0.0542***	0.00120**	0.0289***	−3.37e−05
	(0.00394)	(0.00197)	(0.00404)	(0.000508)	(0.00273)	(0.000277)
Constant	0.929***	0.536***	0.580***	0.0148***	0.322***	0.0118***
	(0.00334)	(0.00167)	(0.00361)	(0.000483)	(0.00224)	(0.000263)
Observations	7168	7168	7168	7168	7168	7168
R-squared	0.032	0.032	0.136	0.025	0.158	0.012
Number of PUMAs	55	55	55	55	55	55

Notes: All outcome variables measure changes between t and $t-5$ and are expressed as shares of establishments over the average of t and $t-5$. All specifications include PUMA fixed effects. Robust standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 11b. OLS regression of turnover variables on gentrification variables (rents).

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag low rent	0.0112	−0.00560	0.00739	−0.000852	0.00488	−0.000209
	(0.00720)	(0.00360)	(0.00513)	(0.000673)	(0.00335)	(0.000519)
Std. lag low rent* rising rent	−0.00487	0.00243	−0.00373	0.000145	−0.000794	−0.000483
	(0.00462)	(0.00231)	(0.00343)	(0.000582)	(0.00258)	(0.000481)
Std. 2000	−0.0399***	0.0200***	−0.0781***	0.00179***	0.0349***	0.00147***
	(0.00329)	(0.00165)	(0.00471)	(0.000445)	(0.00371)	(0.000388)
Std. 2005	−0.0427***	0.0213***	−0.0233***	0.00574***	−0.0280***	0.00295***
	(0.00376)	(0.00188)	(0.00367)	(0.000759)	(0.00200)	(0.000363)
Std. 2010	−0.0243***	0.0121***	−0.0546***	0.00123**	0.0291***	3.71e−05
	(0.00379)	(0.00189)	(0.00399)	(0.000511)	(0.00282)	(0.000272)
Constant	0.930***	0.535***	0.582***	0.0147***	0.322***	0.0116***
	(0.00319)	(0.00159)	(0.00344)	(0.000469)	(0.00248)	(0.000235)
Observations	7388	7388	7388	7388	7388	7388
R-squared	0.032	0.032	0.135	0.021	0.157	0.012
Number of PUMAs	55	55	55	55	55	55

Notes: All outcome variables measure changes between t and $t-5$ and are expressed as shares of establishments over the average of t and $t-5$. All specifications include PUMA fixed effects. Robust standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

then use the results for the specific types of turnover to broaden our understanding of overall churn. Neighbourhoods experience more churn when the local commercial activity is more homogeneous, and when there are higher shares of retail. This

Table 12. OLS regression of turnover variables on consumer characteristics and gentrification variables.

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag pop. density	−0.0240***	0.0120***	−0.0146***	−0.00339***	−0.00376	−0.00228***
	(0.00774)	(0.00387)	(0.00532)	(0.00106)	(0.00342)	(0.000520)
Std. lag population change (%)	0.0314	−0.0157	−0.0177	0.00249	0.0468**	−0.000222
	(0.0302)	(0.0151)	(0.0319)	(0.00325)	(0.0194)	(0.00323)
Std. lag vacancy (%)	0.00621	−0.00311	0.000216	−0.000541	0.00658**	−4.22e−05
	(0.00564)	(0.00282)	(0.00453)	(0.000905)	(0.00316)	(0.000575)
Std. lag homeownership (%)	0.00398	−0.00199	−0.0128	0.00383**	0.0145**	−0.00156
	(0.00990)	(0.00495)	(0.00873)	(0.00146)	(0.00573)	(0.00113)
Std. lag non-family hhlds. (%)	−0.0480***	0.0240***	−0.0305***	−0.00191	−0.0151**	−0.000551
	(0.0120)	(0.00598)	(0.0102)	(0.00155)	(0.00612)	(0.00154)
Std. lag average persons/hhld.	−0.384***	0.192***	−0.170*	−0.00779	−0.202***	−0.00330
	(0.125)	(0.0623)	(0.0927)	(0.0148)	(0.0618)	(0.0118)
Std. lag black pop. (%)	0.00931	−0.00465	0.00622	−0.00409	0.00591	0.00127
	(0.0171)	(0.00853)	(0.0142)	(0.00286)	(0.00751)	(0.00129)
Std. lag Hispanic pop. (%)	0.00343	−0.00172	−0.00364	−0.00145	0.00739	0.00113
	(0.0127)	(0.00637)	(0.0106)	(0.00207)	(0.00677)	(0.00105)
Std. lag white pop. (%)	−0.0679***	0.0339***	−0.0473**	−0.00427	−0.0182*	0.00184
	(0.0226)	(0.0113)	(0.0184)	(0.00387)	(0.00956)	(0.00148)
Std. lag foreign born (%)	−0.0123	0.00617	−0.0155*	−0.00175	0.00452	0.000423
	(0.00960)	(0.00480)	(0.00814)	(0.00109)	(0.00427)	(0.000639)
Std. lag poverty (%)	−0.00386	0.00193	−0.0123	0.00175	0.00501	0.00169
	(0.0116)	(0.00581)	(0.0101)	(0.00135)	(0.00555)	(0.00125)
Std. lag real med. hhld. inc.	−0.0181*	0.00904*	−0.000958	0.000244	−0.0172***	−0.000205
	(0.0108)	(0.00539)	(0.0101)	(0.00129)	(0.00603)	(0.00141)
Std. lag ratio avg./med. hhld. inc.	−0.0351*	0.0175*	−0.0358**	0.00197	−0.00413	0.00289
	(0.0197)	(0.00985)	(0.0175)	(0.00327)	(0.0114)	(0.00248)
Std. lag unemployment (%)	0.00299	−0.00149	0.00473	0.000196	−0.000456	−0.00148**
	(0.00752)	(0.00376)	(0.00637)	(0.000894)	(0.00432)	(0.000613)
Std. lag 4-year degree (%)	0.00236	−0.00118	−0.0165*	0.00365**	0.0147***	0.000496
	(0.00999)	(0.00500)	(0.00875)	(0.00167)	(0.00544)	(0.00145)
Std. lag real gross rent	0.00748	−0.00374	−0.000679	0.000372	0.00668	0.00111
	(0.00877)	(0.00439)	(0.00606)	(0.00116)	(0.00566)	(0.000783)

(continued)

Table 12. (Continued)

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag low price/unit	−0.00195	0.000976	0.00852*	−0.00172*	−0.00821**	−0.000540
	(0.00526)	(0.00263)	(0.00456)	(0.000926)	(0.00359)	(0.000680)
Std. lag low price/unit* rising price/unit	−0.00633*	0.00317*	−0.0173***	−1.11e−06	0.0108***	0.000195
	(0.00371)	(0.00185)	(0.00281)	(0.000515)	(0.00335)	(0.000497)
Std. 2005	−0.0237***	0.0118***	0.0191***	0.00461***	−0.0495***	0.00208***
	(0.00290)	(0.00145)	(0.00254)	(0.000699)	(0.00220)	(0.000358)
Std. 2010	−0.00605*	0.00303*	−0.0111***	−0.000185	0.00617**	−0.000925**
	(0.00359)	(0.00179)	(0.00274)	(0.000607)	(0.00238)	(0.000355)
Constant	0.891***	0.555***	0.528***	0.0155***	0.334***	0.0128***
	(0.00577)	(0.00288)	(0.00431)	(0.000698)	(0.00226)	(0.000522)
Observations	7168	7168	7168	7168	7168	7168
R-squared	0.035	0.035	0.044	0.034	0.131	0.013
Number of PUMAs	55	55	55	55	55	55

Notes: All outcome variables measure changes between t and $t-5$ and are expressed as shares of establishments over the average of t and $t-5$. All specifications include PUMA fixed effects. Population density is calculated per square mile of land. Vacancy refers to vacancy of housing units. Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

turnover is primarily driven by (i) an increase in births and move-induced entries in neighbourhoods with relatively homogenous commercial mixes and (ii) both openings and permanent closings in neighbourhoods with a higher share of retail more generally. On the other hand, churn goes down in neighbourhoods with bigger establishments; this is driven by a relatively larger drop in births and an increase in the share of businesses that stay (compared with a smaller increase in move-induced entries). In terms of magnitude, all of these coefficients are similar. Altogether, neighbourhoods with less (and more heterogeneous) retail concentration, as well as bigger businesses, are more stable.

Next, we turn to housing market metrics and specifically measures of baseline housing prices/rents and appreciation on the right-hand side. Controlling only for year and PUMA in Table 11, neither the baseline housing price/rent nor its appreciation affect retail churn. Births are more likely, and move-induced entries marginally less likely, in lower-price areas, but these effects are small and not meaningful enough to drive overall churn.

Next in Table 12, we add other household characteristics to the model, all of which could be correlated with the neighbourhood's economic and retail changes. In these augmented models, the coefficient on price appreciation becomes significant and now has a negative sign, suggesting the churn goes down in initially low-price neighbourhoods that experience price appreciation. Even though these appreciating neighbourhoods also experience more deaths (which is consistent with expectations of business displacement), the reduction in overall churn is driven by the simultaneous reduction in births (so there are not necessarily new businesses coming in to replace those lost services).[25] The other dimensions of consumer demand that seem to influence churn include, population density, household size and composition (the latter of which displays the largest standardised coefficient) and share of the population that identifies as white. All of these factors reduce churn, mostly owing to the fact that businesses are more likely to stay in place (with fewer instances of retail entry/birth or exit/death).[26]

In Table 13, we combine all of these variables into a single model and observe that the coefficients are generally unchanged, especially those capturing commercial factors. We note a few meaningful differences in the consumer-related coefficients. Controlling now for commercial factors, higher rates of population change and housing vacancies significantly increase churn, mostly because of elevated birth rates. In addition, the coefficients on household income, average-median income ratio and price appreciation are no longer significant for any turnover variables.

The multivariate analyses add to our understanding of neighbourhood-based retail turnover, in that most of the patterns that appeared significant from the cross-tab analysis do not persist once we simultaneously control for other commercial and consumer characteristics. The two notable exceptions are (i) the persistently significant coefficients on the share of white households in the neighbourhood and (ii) the reversed, i.e. now positive, sign on the coefficient for the neighbourhood's share retail. More telling, perhaps, is the fact that retail turnover is more strongly associated (both in terms of the standardised coefficients' magnitudes and collective significance) with the neighbourhood's consumer characteristics than the commercial ones. Specifically, the largest standardised coefficient comes from household size (where larger averages are associated with less churn); this suggests that bigger households, controlling for population overall, consume in such a way as to

Table 13. OLS regression of turnover variables on commercial, consumer and gentrification variables.

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag estab. density	−0.0232	0.0116	−0.0320*	−0.00228	0.0107	0.000367
	(0.0199)	(0.00997)	(0.0162)	(0.00207)	(0.00764)	(0.00189)
Std. lag rel. estab. % change	0.00200	−0.00100	−0.0221***	0.00171	0.0221***	0.000232
	(0.00647)	(0.00323)	(0.00518)	(0.00154)	(0.00322)	(0.000600)
Std. lag avg. employees/estab.	−0.0435***	0.0218***	−0.0431***	0.00387	−0.00402	−0.000270
	(0.0153)	(0.00763)	(0.0109)	(0.00244)	(0.00626)	(0.00170)
Std. lag retail (%)	0.0275***	−0.0138***	0.0150***	0.00432***	0.00835*	−0.000103
	(0.00910)	(0.00455)	(0.00535)	(0.00127)	(0.00424)	(0.00186)
Std. lag necessity (%)	−0.0127	0.00637	−0.00560	−0.00211	−0.00273	−0.00230*
	(0.00854)	(0.00427)	(0.00638)	(0.00127)	(0.00428)	(0.00121)
Std. lag independent (%)	−0.00380	0.00190	−0.00768	0.00167	0.00348	−0.00126
	(0.00667)	(0.00334)	(0.00527)	(0.00122)	(0.00430)	(0.000919)
Std. lag herfindahl index	0.0507***	−0.0254***	0.0776***	0.00545***	−0.0325***	0.000184
	(0.0117)	(0.00586)	(0.00846)	(0.00186)	(0.00482)	(0.00230)
Std. lag estab. density*herfindahl	−0.0328	0.0164	−0.0425*	−0.00285	0.00989	0.00263
	(0.0200)	(0.00998)	(0.0248)	(0.00227)	(0.0130)	(0.00262)
Std. lag log commercial area (2005)	−0.0126	0.00629	0.00549	−0.00261*	−0.0156***	0.000108
	(0.00885)	(0.00443)	(0.00676)	(0.00149)	(0.00361)	(0.000739)
Std. lag pop. density	−0.0238**	0.0119**	−0.00554	−0.00336**	−0.0127***	−0.00221***
	(0.00921)	(0.00460)	(0.00671)	(0.00134)	(0.00428)	(0.000738)
Std. lag population change (%)	0.0801***	−0.0401***	0.0706***	−0.00212	0.0123	−0.000691
	(0.0263)	(0.0132)	(0.0252)	(0.00350)	(0.0200)	(0.00313)
Std. lag vacancy (%)	0.0151**	−0.00753**	0.0152***	−0.000427	0.00106	−0.000822
	(0.00672)	(0.00336)	(0.00536)	(0.00116)	(0.00400)	(0.000589)
Std. lag homeownership (%)	0.00720	−0.00360	−0.00448	0.00111	0.0115*	−0.000978
	(0.0108)	(0.00538)	(0.00941)	(0.00158)	(0.00649)	(0.00155)
Std. lag non-family hhlds. (%)	−0.0289**	0.0145**	−0.0132	−0.000291	−0.0137	−0.00174
	(0.0143)	(0.00715)	(0.0103)	(0.00212)	(0.00957)	(0.00200)
Std. lag average persons/hhld.	−0.404***	0.202***	−0.230***	0.00333	−0.166**	−0.0108
	(0.127)	(0.0634)	(0.0816)	(0.0183)	(0.0768)	(0.0159)
Std. lag black pop. (%)	−0.0213	0.0107	−0.0223	−0.00785**	0.00605	0.00278*
	(0.0172)	(0.00862)	(0.0140)	(0.00318)	(0.00910)	(0.00150)
Std. lag Hispanic pop. (%)	−0.00469	0.00234	−0.00608	−0.00422*	0.00401	0.00159
	(0.0135)	(0.00673)	(0.0107)	(0.00223)	(0.00763)	(0.00122)

(continued)

Table 13. (Continued)

	(1) Churn	(2) Stay	(3) Birth	(4) Enter	(5) Death	(6) Exit
Std. lag white pop. (%)	−0.0867***	0.0433***	−0.0611***	−0.00514	−0.0236*	0.00314
	(0.0241)	(0.0120)	(0.0187)	(0.00416)	(0.0120)	(0.00188)
Std. lag foreign born (%)	0.000187	−9.36e−05	0.0118	−0.00257*	−0.00908**	−1.34e−05
	(0.00972)	(0.00486)	(0.00770)	(0.00135)	(0.00418)	(0.000875)
Std. lag poverty (%)	0.00497	−0.00248	−0.00100	−0.000207	0.00435	0.00183
	(0.0128)	(0.00639)	(0.00901)	(0.00164)	(0.00686)	(0.00166)
Std. lag real med. hhld. inc.	−0.0207	0.0103	−0.00764	0.00149	−0.0134	−0.00114
	(0.0141)	(0.00706)	(0.0111)	(0.00169)	(0.0101)	(0.00135)
Std. lag ratio avg./med. hhld. inc.	−0.0117	0.00587	0.000407	0.00573	−0.0187	0.000779
	(0.0206)	(0.0103)	(0.0159)	(0.00388)	(0.0141)	(0.00287)
Std. lag unemployment (%)	0.00197	−0.000986	0.00170	0.00105	0.00141	−0.00219***
	(0.00816)	(0.00408)	(0.00746)	(0.00110)	(0.00497)	(0.000688)
Std. lag 4-year degree (%)	0.00973	−0.00486	0.00183	0.00158	0.00618	0.000145
	(0.0120)	(0.00602)	(0.00942)	(0.00202)	(0.00875)	(0.00171)
Std. lag real gross rent	0.0170*	−0.00848*	0.00723	0.000854	0.00775	0.00114
	(0.00891)	(0.00445)	(0.00555)	(0.00168)	(0.00584)	(0.00104)
Stg. lag low price/unit	−0.00402	0.00201	−0.00263	−0.00125	0.000335	−0.000482
	(0.00541)	(0.00270)	(0.00441)	(0.00112)	(0.00355)	(0.000844)
Std. lag low price/unit*rising price/unit	0.000637	−0.000318	−0.000670	−0.000297	0.00168	−7.07e−05
	(0.00397)	(0.00199)	(0.00319)	(0.000660)	(0.00289)	(0.000577)
Std. 2005	−0.00466	0.00233	0.0537***	0.00419***	−0.0640***	0.00140***
	(0.00324)	(0.00162)	(0.00323)	(0.000668)	(0.00288)	(0.000423)
Std. 2010	0.0136***	−0.00682***	0.0267***	−0.000244	−0.0113***	−0.00149***
	(0.00418)	(0.00209)	(0.00405)	(0.000702)	(0.00316)	(0.000469)
Constant	0.843***	0.578***	0.442***	0.0189***	0.368***	0.0144***
	(0.00620)	(0.00310)	(0.00569)	(0.00102)	(0.00409)	(0.000799)
Observations	5463	5463	5463	5463	5463	5463
R-squared	0.083	0.083	0.214	0.060	0.243	0.019
Number of PUMAs	55	55	55	55	55	55

Notes: All outcome variables measure changes between t and $t-5$ and are expressed as shares of establishments over the average of t and $t-5$. All specifications include PUMA fixed effects. 'Rel. estab. % change' refers to percent change in number of establishments in a tract between time t and $t-5$ relative to the change citywide. Establishment and population density are calculated per square mile of land. Vacancy refers to vacancy of housing units. Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

mitigate against turnover (or certain, more stable establishments locate closer to larger households). The two covariates with the second-largest standardised coefficients (albeit by about one-quarter the magnitude) are the percentage change in population (which increases churn) and the share of white households (which consistently decreases churn). The persistent significance of race (specifically the concentration of white households), even while controlling for a host of other commercial and consumer characteristics, is particularly stark. Again, this pattern could indicate something either about the consumption behaviour of white households that tends to stabilise local businesses or the fact that more stable establishments tend to locate in neighbourhoods with a higher share of white households.[27]

Origin-destination neighbourhoods

Finally, we pull out the businesses that relocate within New York City, since we can document both their origin and destination neighbourhoods. This analysis gives us a sense of the businesses' neighbourhood trajectories and what factors might influence their decisions to relocate. For every business that exits a neighbourhood, we document the social and economic characteristics of its origin neighbourhood and compare them with the same characteristics of its destination neighbourhood. The results of this analysis are displayed in Table 14. There are significant differences between the origin and destination neighbourhoods. Most noticeably, businesses tend to relocate to neighbourhoods that have experienced new housing investment and recent growth in retail. Overall, the differences in socioeconomic characteristics do not differ significantly across the entry and exit neighbourhoods. Together these patterns suggest that businesses are perhaps more motivated to move by different infrastructure opportunities (or possibly lower rents in the emerging retail areas – the entry neighbourhoods do tend to be less commercially dense at the time of the move as well). The data, however, do not support the claim that businesses are moving under the allure of a different consumer base (in fact, they are perhaps making sure to keep that factor constant in their decisions). This is consistent with the consumer-driven turnover results from the multivariate analysis, which are most pronounced for new births (versus relocations).

Conclusion and policy implications

Urban neighbourhoods thrive on their mixed-use character – their vitality is dependent on the coexistence of residential and commercial activity. Indeed, this interaction generates both economic and cultural value for the neighbourhood, a phenomenon that has been documented in local shopping districts in cities across the globe (Zukin et al., 2015). We understand less, however, about the prevalence and conditions of retail turnover across multiple neighbourhoods in a single municipality, a process that could both threaten and invigorate communities. In this paper, we document retail turnover in a large, dense municipality, New York, and test, at a scale not done before, whether or not it varies depending on neighbourhood characteristics and circumstances.

We find that there are meaningful differences in retail turnover across neighbourhoods, related to business activity, commercial infrastructure and the neighbourhood's consumer profile. However, when all three sets of factors are considered simultaneously, consumer-related characteristics seem to explain turnover more than those related to the local commercial environment. Food establishments (i.e. restaurants) tend to be a more stabilising presence

Table 14. Origin compared with destination neighbourhoods.

	Tract entered		Tract exited		Difference	
	N	Mean	*N*	Mean		Sig. diff.
Estab. density	8563	2118	8563	2327	−210	***
Emp. density	8563	25,485	8563	27,168	−1683	**
Rel. estab. % change	6992	1.302	6998	1.058	0.244	***
Herfindahl index	8563	0.159	8563	0.158	0.001	
Commercial building area > 75th percentile of tracts	4298	0.579	4298	0.633	−0.054	***
Commercial building area (sq. ft.)	4298	3,655,000	4298	3,844,000	−188,869	
Homeownership (%)	8563	0.282	8515	0.254	0.028	***
Rental housing units built before 1970 (%)	8558	0.839	8513	0.851	−0.014	**
Black pop. (%)	8563	0.123	8563	0.123	0.000	
Hispanic pop. (%)	8563	0.171	8563	0.173	−0.002	
White pop. (%)	8563	0.589	8563	0.584	0.004	
Foreign born (%)	8563	0.299	8563	0.299	−0.001	
Poverty (%)	8563	0.160	8563	0.163	−0.003	
Median household income	8563	$70,314	8563	$70,330	−$17	
Unemployment (%)	8563	0.081	8563	0.080	0.001	
Receiving public assistance (%)	8563	0.096	8563	0.096	0.000	
Less than high school diploma (%)	8563	0.206	8563	0.208	−0.001	
High school diploma (%)	8563	0.196	8563	0.187	0.009	***
Some college (%)	8563	0.189	8563	0.182	0.007	***
4-year degree or more (%)	8563	0.409	8563	0.420	−0.011	***
Tract median rent > median of all tracts	8563	0.774	8562	0.766	0.008	
Median price per residential unit	7876	$1,524,000	7916	$1,204,000	$326,339	
Median price per unit relative to borough median	7876	2.941	7916	2.392	0.554	
Median price per unit relative to city median	7876	4.904	7916	4.066	0.857	
Units authorised by new residential bldg. permits	8225	15.210	8225	15.660	−0.446	
Units authorised in last 5 years > median of tracts	8225	0.150	8225	0.114	0.036	***

Notes: As an example, we interpret some of the cells: of establishments that moved, the average establishment's destination tract had 210 fewer establishments per mile and a homeownership rate 2.8 percentage-points higher than its origin tract. 'Rel. estab. % change' refers to percent change in number of establishments in a tract between time t and $t-5$ relative to the change citywide. Establishment density is calculated per square mile of land.

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

in neighbourhoods over time, and businesses that provide more frequently consumed necessity goods and services are more likely to stay in place. Chain establishments are less likely to open up brand new establishments in New York City, and, when they do open, are more likely to enter neighbourhoods with more commercial space, lower vacancy rates, lower housing prices, more affluent households, and fewer owner-occupied and college-educated households. Overall, neighbourhoods with less (and more heterogeneous) general retail (as opposed to food service) concentration, as well as bigger businesses, are more stable. Most significantly, bigger households and higher shares of white residents are most strongly associated with less retail churn and population growth is the strongest predictor of more turnover.

Our results generally support the expectation that retail turnover should be a function of both production-related (for example, commercial space and robust existing markets) and consumer-related (for example, race and household size) factors. However, the consumer characteristics are more pronounced and tend to particularly influence the first-time entry of chains into the market; relocations of businesses are driven more by characteristics of the commercial environment, i.e. moving towards more/better space. These findings are consistent with previous studies of firm locations (i.e. Waldfogel, 2008) and with qualitative case studies of shopping districts in cities across the world that reveal the business' dependence on local consumers (both affluent and poor) in sustaining their small stores (Zukin et al., 2015). We also see that the nature of retail turnover matters – it is an incomplete metric to solely look at net changes in retail (which is what most public data make available). Our findings show that instances of increased retail churn are more often than not driven by more births or entries from other neighbourhoods in the city (rather than deaths or exits). This potentially sheds a more positive light on retail turnover, if it indeed brings in new services that were previously underprovided.[28] While we do not observe here the exact services and goods provided by those new businesses, they are not overwhelmingly emerging at the expense of other incumbent businesses.[29]

This raises an important qualification about retail stability and whether or not it is universally beneficial for the neighbourhood. There could be circumstances where the introduction of new services improves residents' quality of life, and that any shock to local comfort levels could be mitigated by such gains. This is a challenging, yet critical, balance to achieve: one that maintains a retail environment with some familiarity, as to not alienate incumbent residents and at the same time capitalises on increased local investment. It is also an opportunity for government to get involved, and to help the neighbourhoods to think holistically about their assets, deficits and risks. If commercial infrastructure matters, not only in terms of physically appropriate spaces, but also economically developed retail markets, then local governments can, through local zoning ordinances, allow, incentivise or even mandate the build-out of commercial spaces. They can go further and think about what kinds of businesses they hope to attract to those spaces. Independently owned establishments might be more likely to have ties with (and redistribute benefits to) the community; chain retailers could bring more selection and possibly lower prices and, according to our analysis, pose no significant threat of increased turnover; businesses that provide necessity services exhibit more stability and also meet more immediate needs; but a diversity of services also helps with stability.

Again, local government can mandate or nudge in order to motivate landlords to rent

to certain kinds of businesses (especially those that are typically viewed as more risky tenants). And the scale of intervention can vary as well, depending on the local government's stake in and vision for redevelopment. For example, in Shanghai, the local government has encouraged large-scale development at the expense of small shop owners; Amsterdam's government has played a less active role and market forces have taken hold; Toronto has relied heavily on public–private institutions in the form of business improvement districts to stimulate and integrate local retail development (Zukin et al., 2015). Therefore, the local government can play a role in not only diversifying or augmenting existing retail corridors, but also encouraging the growth of nascent ones. We also find that lower churn is typically accompanied by higher shares of businesses that stay in place – it is therefore important to think about policies that help these establishments stay open, especially in areas that otherwise demonstrate high turnover.

Finally, our findings suggest that accurate and accessible information is critical in achieving any of these policy goals. First, understanding the nuances of retail instability requires tracking and monitoring the flow of businesses into and out of neighbourhoods. This would ideally entail not only accessing microdata on business activity (which several government agencies should already possess), but also interacting with local community organisations and business improvement districts on the ground who can better speak to the quality and patronage of the services that come and go. Second, the persistent effect of race (that whiter neighbourhoods tend to experience less churn, even controlling for income) is consistent with prior research that finds race-based bias in business' location decisions (Helling and Sawicki, 2003). If businesses are using race, an observable feature of a neighbourhood, as a proxy for other, less accessible indicators of economic viability (i.e. safety, purchasing power), then local government could assist in the dissemination of accurate data on local consumer dynamics. Indeed, providing businesses with more complete information and refined tools to read local markets could better inform start-up business decisions, support in-place business sustainability, and ultimately better satisfy local service needs.

Acknowledgements

The views expressed in the article are not necessarily those of the Department of Housing Preservation and Development or the City of New York. We thank Pooya Ghorbani and Colin Bottles for excellent research assistance, and the Furman Center for Real Estate and Urban Policy at New York University for generously sharing their data.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. See also Sutton (2010) for a broader summary of this argument.
2. Note that the transportation costs are less of a concern in this framing, since we are dealing with local neighbourhood services that do not tend to vary much with respect to transportation costs (they are all located within walking distance of nearby residents and employees). Furthermore, transportation infrastructure is generally fixed in New York City over the course of the study period and so changes in proximate transportation options will not play a large role in retail changes over time.
3. Here we treat information collection about the localised market as an up-front cost for setting up a business.
4. In the case of large 'big box' chains, the fixed costs associated with larger spaces might also be higher than smaller

independent establishments; this kind of investment will also trigger a higher threshold for closure or relocation (since the costs of moving will reflect a new set of sizable fixed costs).

5. ZIP codes are geographies in the USA that relate to postal service and were created to expedite the delivery of mail and packages. Their boundaries can change over time and do not necessarily have any contextual meanings other than what is a convenient route for postal delivery. Important for the current analysis, they are typically more than ten times bigger than the census tract, so likely obscure a good amount of business turnover.
6. Ege (2009) critiques NETS' coverage of small science and technology sector firms, saying it is thin; this discrepancy, if it exists, should not affect our analysis as we focus on neighbourhood services that are largely outside these sectors. Other critiques of NETS (see Davis and Haltiwanger, 1998; Davis et al., 1996) focus largely on discrepancies in employment counts and issues with reporting in the earlier iterations of Dunn & Bradstreet (from which NETS is derived); most of the concerns have been addressed with newer, more robust collection efforts by Dunn & Bradstreet (Neumark et al., 2005). In addition, since we focus primarily on firm counts rather than employment numbers, concerns over the latter are less threatening in the current analysis.
7. We make a number of cuts, including removing outliers and, in some cases, limiting tract-years to those corresponding with decennial Census years. In addition, since we rely on the number of establishments from the previous decade as the denominator in our primary calculation of business moves (detailed below), we also need to drop tract-years from 1990. We also note that there are four important variations in the sample, depending on the type of analyses. First, the final sample used in our core descriptive analyses contains 3852 tract-year records observed in 2000 and 2010. Second, in analyses of neighbourhoods where chains appear for the first time, we do not drop observations occurring in non-decennial Census years but drop tracts that have their first-chain appearance/entry/birth before 1996; after removing outliers, we are left with a sample size of 1015 tract-years. Third, in our analysis of inter-neighbourhood moves, there are 8563 tract-years that experience an inter-neighbourhood move during the study period. Finally, in multivariate analyses, we retain tracts observed in 1995 and 2005, which yields a sample size of 7393 tract-years.
8. The median New York City tract covers 0.0693 square miles of land (roughly 44 acres). A perfectly square tract of this size would have sides a little longer than a quarter of a mile, which would take five minutes to walk along at an average speed of three miles per hour. If the tract were rectangular, measuring about 0.5 by 0.14 miles, it would take about 10 minutes to walk along the longer side.
9. We replicate all of the analyses using a standard change calculation (where the denominator is the total number of establishments at time $t-k$). The results are generally the same; if anything, the midpoint method tends to produce lower share estimates.
10. These, along with other basic descriptives, are displayed in Table 1.
11. The discrepancies across typologies also allow for flexibility in how specific businesses are assigned. For example, many of the subsectors that could be considered discretionary, but are also frequently consumed (and therefore more necessity for some), would be captured in the discretionary-frequent classification. We do not display the results for all of the typology iterations, but these are available from the authors upon request. We note that the reported results are overwhelmingly representative of the universe of typology crosstabs.
12. We recognise that there could be a number of factors that mediate the effect of these broadly characterised dimensions, such as government interventions, social norms or place identity. The current analysis exploits a large-N sample to observe broader patterns (versus a case-study approach, which

might be able to delve more deeply into these myriad factors). We focus on documenting the magnitude of turnover disparities and identifying where we would expect to find more or less commercially stable communities; questions of why or how these disparities emerge or exacerbate are crucial, but outside the scope of the current analysis. We argue, however, that the systematic assessment (and a documentation of how turnover actually manifests itself) in our current analysis is critical to accurately testing the impact of other interventions or mediators in future research.

13. We also replicate all of these descriptives for three- and ten-year intervals; these are not displayed as they generally reinforce the results for the five-year intervals.
14. The hazard rate at time t is understood as the unobserved rate at which an event occurs, in this case, the entry of the first chain and is the expected duration of time (using the origin of the study period, 1990, as a starting point) until the event occurs. The partial likelihood of the Cox model is a flexible estimation option, because it allows for an unspecified form for the underlying survivor function. See Allison (1984, 1995) for a detailed description of using Cox regressions models in survival analysis. We extend the Cox proportional hazards model to include time-varying covariates; other than additional computational complexity, the partial likelihood estimation is robust to this specification (see Allison, 1995; Grambsch and Therneu, 1994). The presence (and significance) of time-varying covariates by definition violates the proportionality assumption of the proportional Cox model, but is also the choice method to address variation in the hazard over time (see Allison, 1995). Time-weighted scaled Schoenfeld residuals plotted against time indicate that proportionality is upheld in most cases (see Grambsch and Therneu, 1994) and supports linear non-proportionality in the cases where time-varying covariates are used.
15. We also replicate the model stratifying by PUMAs, which are Census-defined geographic boundaries that cover many tracts; the results are largely consistent, with the exception of the % necessity variable, which is no longer significant. We prefer the unstratified model, because the coefficient on % white population blows up tenfold in the stratified model (and becomes more significant). Therefore, the results displayed show a more conservative estimate of the effect of % white population on chain entry.
16. This effect, however, goes away if we stratify by PUMA.
17. The results from the hazard analysis are consistent with simple comparative statistics across the same subset of neighbourhoods. The main findings are also robust to more parsimonious models that eliminate moderately correlated covariates.
18. 'High' poverty is defined as those neighbourhoods with a poverty rate in the top 25th percentile of the distribution of all neighbourhood-year poverty rates. In 2000, this amount was set at approximately 28.3%.
19. We replicate the crosstabs separating out the 'stable' from the 'declining' neighbourhoods and we note one distinct pattern: (i) declining neighbourhoods exhibit more churn than stable neighbourhoods, close to levels of appreciating neighbourhoods, and this is largely because of a lower rate of businesses that stay in place. For purposes of brevity, and the fact that our analysis focuses on outcomes for gentrifying neighbourhoods, we focus on the dichotomous classification of increasing against non-increasing.
20. This is supported by a bivariate analysis of commercial space and permit activity; there is relatively more permit activity in areas with less commercial space, compared with areas with more commercial space. This permit activity also tends to take place in price-appreciating neighbourhoods.
21. These results are not displayed here, but are available from the authors upon request.
22. See Ellen and O'Regan (2008), McKinnish et al. (2010) and Meltzer and Schuetz (2012) for other examples of using prices/rents to identify neighbourhoods upgrading or 'gentrification'.
23. Used primarily in Census microdata products, PUMAs are geographic boundaries

that cover many tracts and allow us to control for unobservable neighbourhood characteristics. The boundaries of New York City's 55 PUMAs roughly approximate community districts, the areas served by an advisory governmental entity known as a community board.

24. The unadjusted coefficients are available from the authors upon request. We include a table of the variable means and standard deviations in Appendix B.
25. This is consistent with Meltzer (forthcoming), who finds increased displacement without replacement in gentrifying neighbourhoods in New York City during the 2000s.
26. The results for the augmented model including rent-based appreciation variables is not shown, but the results for the household covariates are substantively the same as those reported for the price-based models. The coefficients on the rent-based variables, however, are not significant.
27. In order to test for the sensitivity of our results to our definition of retail (and specifically to test it against a much more conservative definition), we replicate the full specification for frequently consumed necessity turnover metrics only (as the dependent variables). The consumer-related characteristics are all unchanged, with the exception of average household size, which goes down in magnitude and loses significance across the board. While the commercial variables are all consistent in terms of sign and magnitude, there are some differences with respect to significance levels. The establishment size variable loses significance for the churn model (which coincides with a now insignificant coefficient in the stay model). And, whereas the retail share was an important correlate with turnover in the model using the broader retail definition, the necessity share variable is now instead significantly associated with lower churn (driven by higher stay rates and lower birth/entry rates).
28. It does, however, suggests some re-sorting of existing businesses across neighbourhoods, rather than the formation of new businesses.
29. Meltzer (forthcoming) finds that the tradeoff between commercial displacement and the entry of new services varies depending on the intensity and nature of the neighbourhood's gentrification. New services can emerge in the context of both high and low displacement rates.

References

Allison PD (1984) *Event History Analysis: Regression for Longitudinal Data*. Newbury Park, CA: Sage Publications.

Allison PD (1995) *Survival Analysis Using the SAS System: A Practical Guide*. Cary, NC: SAS Institute Inc.

Alwitt L and Donley T (1997) Retail stores in poor neighborhoods. *Journal of Consumer Affairs* 31: 139–164.

Bates T and Robb A (2008) Analysis of young neighborhood firms serving urban minority clients. *Journal of Economics and Business* 60(1): 139–148.

Bates T and Robb A (2014) Small-business viability in America's urban minority communities. *Urban Studies* 51(13): 2844–2862.

Been V, Capperis S, De la Roca J, et al. (2013) *State of New York City's Housing and Neighborhoods in 2012*. New York: The Furman Center.

Berry ST and Waldfogel J (2003) *Product quality and market size*. NBER working paper #9675.

Bingham RD and Zhang Z (1997) Poverty and economic morphology of Ohio central-city neighborhoods. *Urban Affairs Review* 32(6): 766–796.

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

Bridge G and Dowling R (2001) Microgeographies of retailing and gentrification. *Australian Geographer* 32(1): 93–107.

Brueckner JK and Rosenthal SS (2009) Gentrification and neighborhood housing cycles: Will America's future downtowns be rich? *The Review of Economics and Statistics* 91(4): 725–743.

Capperis S, De la Roca J, Ellen GI, et al. (2015) *State of New York City's Housing and Neighborhoods in 2014*. New York: The Furman Center.

Capperis S, De la Roca J, Findlan K, et al. (2014) *State of New York City's Housing and Neighborhoods in 2013*. New York: The Furman Center.

Carree M and Thurik R (1996) Entry and exit in retailing: Incentives, barriers, displacement and replacement. *Review of Industrial Organization* 11(2): 155–172.

Chapple K and Jacobus R (2009) Retail trade as a route to neighborhood revitalization. In: Wial H, Pindus N and Wolman H (eds) *Urban and Regional Policy and its Effects*. Washington DC: Brookings Institution-Urban Institute, pp. 19–68.

Davis P (2006) Spatial competition in retail markets: Movie theaters. *The RAND Journal of Economics* 37(4): 964–982.

Davis SJ and Haltiwanger J (1998) Measuring gross worker and job flows. In: Haltiwanger J, Manser ME and Topel RH (eds) *Labor Statistics Measurement Issues*. Chicago, IL: University of Chicago Press, pp. 77–122.

Davis SJ, Haltiwanger J and Schuh S (1996) Small business and job creation: Dissecting the myth and reassessing the facts. *Small Business Economics* 8(4): 297–315.

Deener A (2007) Commerce as the structure and symbol of neighborhood life: Reshaping the meaning of community in Venice, California. *City & Community* 6(4): 291–314.

Dinlersoz EM (2004) Firm organization and the structure of retail markets. *Journal of Economics and Management Strategy* 13(2): 207–240.

Ege M (2009) *How do grants influence firm performance? An econometric evaluation of the SBIR Program at NIH*. Thesis submitted to the Graduate School-New Brunswick, Rutgers, The State University of New Jersey. Available at: http://mss3.libraries.rutgers.edu/dlr/outputds.php?pid=rutgers-lib:26471.

Ehrenhalt A (1999) Community and the corner store: Retrieving human-scale commerce. *Responsive Community* 9: 30–40.

Ellen IG and O'Regan K (2008) Reversal of fortunes? Low-income urban neighborhoods in the 1990s. *Urban Studies* 45(4): 845–869.

Freeman L (2005) Displacement or succession? Residential mobility in gentrifying neighborhoods. *Urban Affairs Review* 40(4): 463–491.

Freeman L and Braconi F (2004) Gentrification and displacement New York City in the 1990s. *Journal of the American Planning Association* 70(1): 39–52.

Geolytics (2003) *CensusCD Neighborhood Change Database*. Somerville, NJ: Geolytics, Inc.

Hamnett C (1991) The blind men and the elephant: The explanation of gentrification. *Transactions of the Institute of British Geographers* 16(2): 173–189.

Healey J (2012) Wal-Mart and Chinatown: Who speaks for the community? *Los Angeles Times*, 13 August.

Helling A and Sawicki DS (2003) Race and residential accessibility to shopping and services. *Housing Policy Debate* 14(1–2): 69–101.

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

Hyra DS (2008) *The New Urban Renewal: The Economic Transformation of Harlem and Bronzeville*. Chicago, IL: University of Chicago Press.

Immergluck D (1999) Neighborhoods, race, and capital The effects of residential change on commercial investment patterns. *Urban Affairs Review* 34(3): 397–411.

Jacobs J (1961) *The Death and Life of Great American Cities*. New York: Vintage.

Kaufman TK, Sheehan DM, Rundle A, et al. (2015) Measuring health-relevant businesses over 21 years: Refining the National Establishment Time-Series (NETS), a dynamic longitudinal data set. *BMC Research Notes* 8(1): 507.

Kwate NOA, Loh JM, White K, et al (2013). Retail redlining in New York City: racialized access to day-to-day retail resources. *Journal of Urban Health*, 90(4), 632–652.

Lees L (1994) Gentrification in London and New York: An Atlantic gap? *Housing Studies* 9(2): 199–217.

Ley D (1986) Alternative explanations for inner-city gentrification: A Canadian assessment. *Annals of the Association of American Geographers* 76(4): 521–535.

McKinnish T, Walsh R and White K (2010) Who gentrifies low income neighborhoods? *Journal of Urban Economics* 67(2): 180–193.

Meltzer R (forthcoming) Gentrification and small business: Threat or opportunity? *Cityscape*.

Meltzer R and Schuetz J (2012) Bodegas or bagel shops? Neighborhood differences in retail & household services. *Economic Development Quarterly* 26(1): 73–94.

Neumark D, Zhang J and Wall B (2005) *Employment Dynamics and Business Relocation: New Evidence from the National Establishment Time Series* (No. w11647). National Bureau of Economic Research.

Patch J (2008) Ladies and gentrification: New stores, residents, and relationships in neighborhood change. *Gender in an Urban World, Research in Urban Sociology* 9: 103–126.

Powell L, Slater S, Mirtcheva D, et al. (2007) Food store availability and neighborhood characteristics in the United States. *Preventive Medicine* 44(3): 189–195.

Pristin T (2009) A difficult birth for East Harlem Mall. *The New York Times*, 20 October.

Ross G (1998) *Are black communities truly underserved?: An analysis of retail and professional services in the Atlanta Metropolitan Area*. Ph.D. dissertation, Georgia State University.

Schuetz J (2015) Why are Walmart and Target next-door neighbors? *Regional Science and Urban Economics* 54: 38–48.

Schuetz J, Kolko J and Meltzer R (2012) Are poor neighborhoods 'retail deserts'? *Regional Science and Urban Economics* 42(1): 269–285.

Sloane DC, Lewis LB and Nascimento LM (2005) Assessing healthy food options in south Los Angeles restaurants. *American Journal of Public Health* 95(4): 668–673.

Smith N (1979) Toward a theory of gentrification a back to the city movement by capital, not people. *Journal of the American Planning Association* 45(4): 538–548.

Stanback TM, Bearse PJ, Noyelle TJ, et al. (1981) *Services, the New Economy*. Totowa, NJ: Allanheld, Osmun.

Sutton SA (2010) Rethinking commercial revitalization: A neighborhood small business perspective. *Economic Development Quarterly* 24(4): 352–371.

US Census Bureau (2012) Geographic Terms and Concepts – Census Tract. Available at: http://www.census.gov/geo/reference/gtc/gtc_ct.html (last modified 6 December 2012).

Waldfogel J (2008) The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics* 63: 567–582.

Walls and Associates (2012) *National Establishment Time-Series Database*. Oakland, CA: Walls and Associates.

Yang T and Aldrich HE (2012) Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing* 27(4): 477–492.

Zenk S, Schulz AJ, Israel BA, 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 S (2008) Consuming authenticity: From outposts of difference to means of exclusion. *Cultural Studies* 22(5): 724–748.

Zukin S, Kasinitz P and Chen X (2015) *Global Cities, Local Streets: Everyday Diversity from New York to Shanghai*. New York, NY: Routledge.

Zukin S, Trujillo V, Frase P, et al. (2009) New retail capital and neighborhood change: Boutiques and gentrification in New York City. *City and Community* 8(1): 47–64.

Appendix A. NAICS codes.

NAICS 2007	Description	Infrequent	Frequent	Necessity	Discretionary
441	Motor Vehicle and Parts Dealers	X			X
4421	Furniture Stores	X		X	
4422	Home Furnishings Stores	X			X
44311	Appliance, Television, and Other Electronics Stores	X		X	
44312	Computer and Software Stores	X			X
44313	Camera and Photographic Supplies Stores	X			X

(continued)

Appendix A. (Continued)

NAICS 2007	Description	Infrequent	Frequent	Necessity	Discretionary
444	Building Material and Garden Equipment and Supplies Dealers	X			X
44413	Hardware Stores[a]		X	X	
4451	Grocery Stores		X	X	
44521	Meat Markets		X	X	
44522	Fish and Seafood Markets		X	X	
44523	Fruit and Vegetable Markets		X	X	
44529	Other Specialty Food Stores		X		X
4453	Beer, Wine, and Liquor Stores		X		X
44611	Pharmacies and Drug Stores		X	X	
44612	Cosmetics, Beauty Supplies, and Perfume Stores		X		X
44613	Optical Goods Stores		X	X	
446191	Food (Health) Supplement Stores		X		X
446199	All Other Health and Personal Care Stores	X		X	
44711	Gasoline Stations with Convenience Stores		X	X	
44719	Other Gasoline Stations	X			X
4481	Clothing Stores	X		X	
4482	Shoe Stores	X		X	
4483	Jewelry, Luggage, and Leather Goods Stores	X			X
4511	Sporting Goods, Hobby, and Musical Instrument Stores	X			X
4512	Book, Periodical, and Music Stores		X		X
4521	Department Stores		X		X
4529	Other General Merchandise Stores		X	X	
4531	Florists		X		X
45321	Office Supplies and Stationery Stores		X		X
45322	Gift, Novelty, and Souvenir Stores		X		X
4533	Used Merchandise Stores	X			X
45391	Pet and Pet Supplies Stores		X		X
45392	Art Dealers	X			X
45393	Manufactured (Mobile) Home Dealers	X			X
45399	All Other Miscellaneous Store Retailers	X			X
52211	Commercial Banking		X	X	
52213	Credit Unions		X	X	
53223	Video Tape and Disc Rental		X		X
54194	Veterinary Services	X			X
71312	Amusement Arcades	X			X
71394	Fitness and Recreational Sports Centers		X		X
71395	Bowling Centers	X			X
7221	Full-Service Restaurants		X		X
7222	Limited-Service Eating Places		X		X
7224	Drinking Places (Alcoholic Beverages)		X		X

(continued)

Appendix A. (Continued)

NAICS 2007	Description	Infrequent	Frequent	Necessity	Discretionary
812111	Barber Shops		X	X	
812112	Beauty Salons		X	X	
812113	Nail Salons		X		X
812199	Other Personal Care Services		X		X
81231	Coin-Operated Laundries and Drycleaners		X	X	
81232	Drycleaning and Laundry Services (except Coin-Operated)		X	X	
81291	Pet Care (except Veterinary) Services	X			X

Notes: [a]Hardware stores are an exception within NAICS 44413. By searching for 'hardware' within the establishment name, we code these establishments as 'frequent' and 'necessities'. Otherwise, like the rest of NAICS 444, we code other establishments within 44413 (primarily tool retailers) as 'infrequent' and 'discretionary'.

Appendix B. Descriptive statistics of explanatory variables in regression sample.

Variable	*N*	Mean	SD
Estab. density (per sq. mi.)	7393	448.4	710.2
Rel. estab. % change	5603	1.2	6.7
Avg. employees/estab.	7393	5.1	4.5
Retail	7393	83.8%	13.0%
Necessity	7393	49.5%	18.6%
Independent	7393	93.8%	8.6%
Herfindahl index	7393	0.208	0.155
Estab. density*herfindahl	7393	71.07	116
Log commercial area (2005)	7385	12.5	1.3
Pop. density (per sq. mi.)	7393	47,657	34,582
Population change	7393	4.5%	15.2%
Vacancy	7393	5.9%	4.2%
Homeownership	7393	36.1%	23.5%
Non-family hhlds.	7393	32.8%	14.3%
Average persons/hhld.	7393	2.81	0.55
Black pop.	7393	25.4%	31.8%
Hispanic pop.	7393	23.6%	22.0%
White pop.	7393	40.4%	33.8%
Foreign born	7393	33.5%	15.8%
Poverty	7393	18.3%	12.9%
Real med. hhld. inc. ($)	7393	55,078	24,346
Ratio avg./med. hhld. inc.	7393	1.319	0.234
Unemployment	7393	9.5%	5.7%
4-year degree	7393	23.5%	17.4%
Real gross rent	7393	1012	285
Log median price/unit	7393	12.4	0.7

Notes: All changes are between time t and $t-5$. 'Rel. estab. % change' refers to percent change in number of establishments in a tract between time t and $t-5$ relative to the change citywide. Establishment and population densities are calculated per square mile of land. Vacancy refers to vacancy of housing units.