Retail Prices in a City

Alon Eizenberg, Saul Lach and Merav Oren-Yiftach*

August 2, 2020

Abstract

This study examines grocery price differentials across neighborhoods in a large metropolitan area (the city of Jerusalem, Israel). Important variation in access to affordable grocery shopping is documented using CPI data on prices, and neighborhood-level credit card expenditure data. Residents of peripheral, non-affluent neighborhoods are charged some of the highest prices in the city, and yet display a low tendency to shop outside their neighborhood. In contrast, residents of affluent, centrally-located neighborhoods often benefit from lower grocery prices charged in their own neighborhood, while also displaying a high propensity to shop at the hard-discount grocers located in the city's commercial districts. The role of spatial frictions in shaping these patterns is studied within a structural model where households determine their shopping destination and retailers choose prices. The estimated model implies strong spatial segmentation in households' demand. Counterfactual analyses reveal that alleviating spatial frictions results in considerable benefits to the average resident of the peripheral neighborhoods. At the same time, it barely affects the equilibrium prices charged across the city, and so it does little to benefit households with limited mobility (e.g., the elderly).

*Eizenberg: The Hebrew University of Jerusalem & CEPR, alon.eizenberg@mail.huji.ac.il. Lach: The Hebrew University of Jerusalem & CEPR, saul.lach@mail.huji.ac.il. Oren-Yiftach: Israel Central Bureau of Statistics, meravo@cbs.gov.il. We are grateful to the editor and anonymous referees for comments and suggestions which greatly improved the paper. We thank Eyal Meharian and Irit Mishali for their invaluable help with collecting the price data and with the provision of the geographic (distance) data. We also wish to thank a credit card company for graciously providing the expenditure data. We are also grateful to Daniel Felsenstein for providing the housing price data, and to Elka Gotfryd for mapping zipcodes into statistical subquarters. We thank Steve Berry, Pierre Dubois, Phil Haile, JF Houde, Gaston Illanes, Volker Nocke, Kathleen Nosal, Mark Rysman, Katja Seim, Avi Simhon, Konrad Stahl, Yuya Takashi, Ali Yurukoglu and Christine Zulehner for helpful comments, as well as seminar participants at Carlos III, CEMFI, DIW Berlin, Frankfurt, Harvard, Johns Hopkins, Penn State, Universidad de Vigo, UVA, Yale and Wharton, and participants at the Israeli IO day (2014), EARIE (2014), the Economic Workshop at IDC (2015), UTDT Conference (2016), CEPR-JIE IO Conference (2017), and IIOC (2017). This project was supported by the Israeli Science Foundation (ISF) grant 858/11, by the Wolfson Family Charitable Trust, and by the Maurice Falk Institue for Economic Research in Israel.

1 Introduction

In January 2014, residents of Qiryat HaYovel, a residential neighborhood in the city of Jerusalem, Israel, initiated a consumer boycott against a neighborhood supermarket. They claimed that prices at this supermarket were much higher than those charged at branches of the same chain operating elsewhere in the city, and that such alternative shopping destinations were not readily accessible to them: "Young families will not travel to Talpiot or Givat Shaul (two commercial districts with hard discount supermarkets) to shop and, instead, shop in the neighborhood for lack of time."[1] Senior residents of the neighborhood were mentioned as another affected demographic group. The boycott ended after the chain agreed to lower the cost of a basket of goods by 14 percent, according to the organizers. In June 2017, a similar consumer boycott in the Jerusalem neighborhood of Gilo was reported to have achieved its goals: the management agreed to equate prices there to those charged by the same chain at one of the city's large commercial districts.[2]

The "boycotting" neighborhoods share important characteristics: both are non-affluent neighborhoods located in the periphery of the city, at considerable distance from the main shopping districts. Access to affordable grocery shopping may therefore depend on both socioeconomic factors and geographic location. In this paper we systematically explore these relationships. Using price data collected in Jerusalem by the Central Bureau of Statistics (CBS), we document considerable variation in grocery prices charged across neighborhoods. We also explore households' shopping patterns via data on aggregate grocery expenditure flows between neighborhoods, obtained from a credit card company.

The combined message from these data is striking: residents of non-affluent, peripheral neighborhoods are charged very high prices in their own neighborhood but, despite this fact, have a high tendency to shop in it. In contrast, residents of more affluent, centrally-located neighborhoods are often charged lower prices in their neighborhood, and also display a higher tendency to shop at even cheaper locations outside their neighborhood. Both empirical facts are consistent with activists' complaints regarding the lack of access to affordable shopping locations. In equilibrium, this lack of access enhances the market power enjoyed by retailers operating in the peripheral neighborhoods.[3]

We explore these equilibrium forces by estimating an empirical model of demand and supply for groceries. Our demand model quantifies households' price-distance trade-off which is substantial: "pushing" a destination neighborhood 1 km further away from an origin neighborhood reduces demand from that origin by about 35 percent. The supply model describes retailers'

[1] "Qiyat HaYovel: the residents' battle against 'My Shufersal'," Ynet (an Israeli news outlet), January 2014.

[2] The organizers said that the chain insisted on maintaining pricing flexibility on 20 "non-essential" products. "Shufersal Deal Gilo announced it will set the same prices as charged at the Talpiot branch," Kol Hair (a Jerusalem local newspaper), June 2017.

[3] Such market power could attract additional supermarket entry into these neighborhoods. However, barriers to supermarket entry in residential neighborhoods are substantial owing to space constraints and zoning restrictions. For tractability reasons, in this paper we treat the entry decisions of supermarkets as given.

pricing decisions given the estimated demand elasticities.

We use the model in counterfactual analysis to tease out the impact of spatial frictions. We investigate the effects of reducing these frictions by changing parameters that capture households' implicit cost of travel. We find that the price level charged in residential neighborhoods is only mildly reduced, and, in certain peripheral neighborhoods, prices even go up slightly. This seemingly-surprising result captures the effect of two conflicting forces: increased competition exerts a downward pressure on prices but, at the same time, the average demand elasticity faced by residential neighborhood retailers becomes lower, prompting them to raise prices.

Whereas prices barely respond, consumer behavior does change when spatial frictions are alleviated: a much larger fraction of residents now travels to the hard discount supermarkets in the commercial districts. As a consequence, the average price paid by residents of a given neighborhood declines significantly, and this effect is particularly strong in the peripheral, non-affluent neighborhoods. An important lesson from this finding is that reductions in spatial frictions can have important policy effects that would be entirely missed by a statistical analysis of prices alone. This motivates our joint analysis of both prices and shopping patterns.

Another lesson from this analysis is that, while the average resident of the peripheral, non-affluent neighborhoods gains substantially when spatial frictions are alleviated, residents with reduced mobility (e.g., the elderly) gain little or not at all, as equilibrium prices charged in those neighborhoods remain high. Relief to such residents will likely require targeted policies.

Our analysis reveals the connections between spatial frictions and the distribution of grocery prices, and the manner by which different groups of households may be affected by potential changes to such frictions. While our counterfactual exercises are conceptual and not intended to simulate concrete policies, they do have some policy relevance. In the case of Jerusalem, the city plans to improve access to its main shopping district both via the extension of the light rail system and via improvements to its internal organization, actions that mimic our counterfactuals.[4]

We next explain how we differ from previous analyses of spatial price equilibria within urban settings. Following this literature review, the paper proceeds as follows: Section 2 presents our data, Section 3 presents the demand model and its estimation, and Section 4 describes our pricing model. Section 5 presents our counterfactual experiments, and Section 6 concludes.

Related literature. Retail price differentials across neighborhoods have attracted considerable attention in the economic literature. Such differentials suggest that standard measures of inequality, based on nominal wages, may be biased (as in Moretti 2013). A vast literature, starting with Caplovitz (1963), has attempted to measure such price differentials to understand whether "the poor pay more," producing mixed empirical findings.[5] This literature focuses on

[4] "The plan: the Talpiot industrial zone to undergo a revolution in the next decade," Kol Hair, April 2016.

[5] MacDonald and Nelson (1991), for example, compared the price of a fixed basket of goods across 322 supermarkets in 10 metropolitan areas in the US, revealing that prices in suburban locations were about 4 percent lower than in central city stores where poorer population lived. Chung and Myers (1999) similarly report that the price of a weekly home food plan was higher in poorer neighborhoods of the Twin Cities

statistical comparisons of retail prices across affluent and non-affluent neighborhoods. With relatively few exceptions, these studies have abstracted from the issue of cross-neighborhood shopping (i.e., shopping outside one's neighborhood of residence).[6] We differ from this literature in two ways: we combine the typical neighborhood-level price data with less typical data on cross-neighborhood shopping flows, and move beyond statistical comparisons by presenting a structural model of the equilibrium forces driving the observed price differentials.

A vast literature on spatial frictions includes classic theoretical contributions by Hotelling (1929) and Salop (1979). Several recent empirical papers have taken a structural approach to study spatial competition in various industries, including Adams and Williams (2019), Miller and Osborne (2012), Thomadsen (2005), Davis (2006), McManus (2007), Houde (2012) and Davis, Dingel, Monras and Morales (2019). Substantial empirical work has considered spatial competition among supermarkets (for example, Chintagunta, Dubé, and Singh 2003, and Smith 2004). Our work shares several features with Dubois and Jódar-Rosell (2010) who study supermarket competition: we also estimate a discrete-continuous demand model, use a supply-side model to identify heterogeneous marginal costs, and consider a counterfactual analysis in which travel costs are reduced (see also Figurelli 2013 and Ellickson, Grieco, and Khvastunov 2016).

Although we share with the Industrial Organization (IO) literature the structural approach to estimate demand and supply primitives, our work is motivated by a perennial question in the urban economics literature. Our focus on the relationship between prices, location and consumer flows motivates our use of data sources different from the typical scanner data used in IO papers. Scanner data are ideal for uncovering rich preference structures, but they may be less useful for uncovering shopping patterns within the city. Instead we use a price index for a basket of grocery goods derived from the official statistical agency's methodology that is comparable across space and time, as well as credit card data that provide a systematic description of shopping flows across all neighborhoods. While it is possible to construct such neighborhood-level price indices and consumer flows from scanner data, it is not obvious that doing so will provide sufficient coverage in the context of our research question.[7] Overall, we view our approach as complementary to the established use of scanner data in studying supermarket demand.

---

metropolitan area. Recent work, in contrast, reports that prices in richer zip codes (Hayes, 2000) or prices paid by high income households (Aguiar and Hurst, 2007) are significantly higher.

[6]Kurtzon and McClelland (2010) study a Bureau of Labor Statistics survey in which respondents report their shopping destinations. They find that the "poor pay neither more nor less than the rich at the stores they shop at." See also Aguiar and Hurst (2007) and Griffith et al. (2009) for analyses of survey data where recorded prices correspond to prices actually paid by households.

[7]Even if the sample of households in the scanner data is random and representative of residents in each neighborhood, it need not adequately cover all origin-destination neighborhood pairs characterizing the shopping decisions. Our credit card data also suffer from selectivity bias, and we address this issue econometrically.

2 Data

We begin by describing Jerusalem's urban structure and its notable partition into distinct neighborhoods. Additional subsections describe the prices collected at retail locations throughout the city, and the data on consumer transaction flows across neighborhoods.

2.1 Jerusalem's urban structure: neighborhoods

Our analysis covers 46 neighborhoods in Western Jerusalem (see Online Appendix C for definitions). These are predominantly Jewish neighborhoods. The eastern part of the city has predominantly Arab neighborhoods which we do not include in our study because of significant differences in the basket of groceries purchased and in the extent of credit card usage across these populations. Moreover, residents of Western Jerusalem do not typically perform their weekly grocery shopping in Eastern Jerusalem and vice-versa.

Figure 1 displays Jerusalem's neighborhoods, highlighting those covered by our study in color. Neighborhoods developed historically along the roads radiating from the "Old City," as is typical in many ancient cities. Jerusalem's hilly topography resulted in geographically separated neighborhoods such that moving between them typically requires some mode of transportation. Most neighborhoods have a small commercial center with a small grocery store and other retail services while many, but not all, have one or two supermarkets. Hard discount (HD) supermarkets are located in well-defined commercial districts. Jerusalem does not have an important suburban ring surrounding it.

At the neighborhood level, we observe demographic variables from the 2008 Israel Census of Population (Israel Central Bureau of Statistics (2008)) that are likely to shift price and travel sensitivities: the fraction of the neighborhood's households that own a car, the fraction of residents above the age of 15 who drive to work, and the fraction of senior residents. We use the average price of housing per square meter in 2007-2008, obtained from the Tax Authority's records of real estate transactions, as a proxy for the neighborhood's wealth.[8] Table 1 reports descriptive statistics.

There are sharp socioeconomic differences across neighborhoods. For example, housing is more than twice as expensive in the central, affluent neighborhood of Rehavya than in the peripheral, non-affluent neighborhood of Neve Yaaqov. In our model, this variation will help identify price and distance sensitivities.

Spatial frictions are captured via a 46-by-46 matrix of distances between each pair of neighborhoods (Israel Central Bureau of Statistics (2012)).[9] Table 2 reports statistics regarding the

[8] Felsenstein (2007-2008).

[9] In the online Appendix C we explain that each neighborhood is comprised of several "statistical areas." The CBS provided us with the shortest road distance between the centroids of each pair of such areas. We then compute the distance d_{jn} between neighborhoods j and n as an average of the distances between each pair of statistical areas belonging to these neighborhoods. As some neighborhoods are quite large, we define neighborhood j's "own distance" d_{jj} as the mean distance between the centroids of each pair of statistical areas

distance to the city center and to the city's two prominent commercial districts (Talpiot and Givat Shaul, hereafter referred as CD1 and CD2, respectively). It also reports statistics on the average distance to all the other neighborhoods, a rough measure of how peripheral the neighborhood is. Table 2 indicates considerable variation: the maximum distances to the city center and to the commercial districts are about twice as large as the corresponding mean (or median) distance.

To facilitate the study of our research question we next identify several neighborhoods of interest that differ considerably in terms of their observed characteristics and location within the city. Three neighborhoods — Neve Yaaqov, Givat Shapira, and Qiryat HaYovel South — are both Non-Affluent and Peripherally located. We shall hereafter refer to them as NAP1, NAP2, and NAP3, respectively. Appendix Table C2 shows that housing prices in those neighborhoods are 9.5, 10.7, and 11.5 NIS, i.e., below the mean (median) of 13.4 (13.3) reported in Table 1. These neighborhoods' peripheral location is clearly indicated in Figure 1.

We also identify three other neighborhoods as Affluent and Centrally located: Rehavya, Qiryat Moshe - Bet Hakerem, and Baqa-Abu Tor-Yemin Moshe, denoted by AC1, AC2 and AC3, respectively. In these AC neighborhoods housing prices are 21.1, 15.8 and 15 NIS (Appendix Table C2), well above the mean price. Figure 1 shows that these neighborhoods are within close proximity to the city center, as well as to the CD1 and CD2 commercial districts.

In short, Jerusalem shares many of the characteristics of other large metropolitan areas: it features well-defined commercial districts, affluent and less affluent neighborhoods, and central and peripheral locations. It is therefore a useful laboratory to study the role of spatial frictions in generating price differentials across neighborhoods.

2.2 Price data

The price data (Israel Central Bureau of Satistics, 2007-2008) were collected by CBS personnel as part of their monthly computation of the Consumer Price Index (CPI), but the sample used in this research includes additional supermarkets, beyond those normally used in the CPI sample. We focus on 27 popular, everyday products consumed by most households. Each selected product is associated with a unique universal product code (UPC) and is therefore identical across sampled stores (e.g., the same brand, size, packaging, etc.).[10] Price observations were collected in three periods: in September and November 2007, and in November 2008. CBS personnel sampled 60 distinct stores in Jerusalem: about 55 percent of them were supermarkets, 20 percent were open market stalls and 15 percent were grocery stores. The sampled stores are present in 26 of our 46 neighborhoods. While this may appear as a major omission, we note that in the remaining 20 neighborhoods there are no important supermarkets and, typically,

included in it.

[10]Even among fruits and vegetables there are no noticeable quality differences across stores because the CBS collects prices of produce of a specific quality grade.

they only have a small grocery store.[11]

The list of products, their mean price and coefficients of variation are displayed in Online Appendix D. Fruits and vegetables usually exhibit higher price dispersion than other foodstuff. One possible explanation for this higher variance is their perishable nature: unsold stocks trigger price reductions, thereby generating a higher variance. We further note that an alternative composite good that excludes fruits and vegetables (as opposed to the one we use, defined below, that includes all products) displays higher price dispersion across neighborhoods. The dispersion in the prices of fruits and vegetables, while important, is therefore not the primary driver of the price variation that we study in the paper.

A composite good. We aggregate individual product items to a composite good whose price is measured at the neighborhood level. Our focus on a composite good is in line with the relevant urban economics literature (e.g., MacDonald and Nelson 1991), and makes particular sense in our application because we observe neighborhood-level expenditure flows.

Residential neighborhoods tend to be served by smaller, more expensive store formats, whereas commercial neighborhoods have larger HD stores. Most price variation is, therefore, between rather than within neighborhoods. Indeed, using the neighborhoods with at least two stores, we computed the between and within variance of price for each item and period separately. In 86 percent of the cases the between-neighborhood variance of prices is larger than the within variance and the median ratio of between to within variance is 3.2.

We define the price of the composite good charged in a given neighborhood as a weighted average of individual-item prices using the CPI expenditure weights. Letting ω_i be the weight of product i, $i = 1, \ldots, 27$, Ω_{nt} be the set of products observed in neighborhood n at time t, and p_{nit} the *average* price of product i in neighborhood n in period t (over all stores selling the product in the neighborhood), the price of a single unit of the composite good is

$$p_{nt} = \sum_{i \in \Omega_{nt}} \left(\frac{\omega_i}{\sum_{i \in \Omega_{nt}} \omega_i} \right) p_{nit}. \tag{1}$$

Missing price observations are typical of studies that construct indices from prices collected by official statistical agencies. Indeed, not all 60 stores are surveyed in each of the three periods, and not all products are surveyed in each store-period.[12] Statistics regarding the number of sampled stores and the prevalence of missing prices are presented in Table 3 (see Online Appendix Table D3 for neighborhood-specific values). On average, the 26 neighborhoods have two sampled stores, including one supermarket. CD1, the main commercial district, has 5 hard discount supermarkets. The average neighborhood has non-missing price data for 17-18

[11] In our econometric demand model, we address the presence of neighborhoods without sampled prices in an internally-consistent fashion by including such shopping destinations in the households' outside option.

[12] While our 27 items are popular products that should be available in all stores, recall that a product is defined by its unique UPC. Some stores may carry a different version of what is essentially the same product (e.g., differing in packaging), generating a missing price observation.

products, and the typical neighborhood (see bottom panel) has observed prices for most of the 27 products.

Our goal is to define a composite good that would be as homogeneous as possible without reducing the sample size too much. We therefore pursue a leading specification that computes the index in 15 neighborhoods (including four commercial districts) where at least 21 of the 27 items have a non-missing price observation. We treat prices in the remaining neighborhoods as unobserved, a feature that will be consistent with our econometric model.[13]

One concern is that changes over time in the identity of the products in the composite good can generate spurious price variation over time. Note, however, that in 8 out these 15 neighborhoods, we observe at least 26 of the 27 products in all three periods. In fact, most of the products appear in all three periods in most neighborhoods. Another concern is that price differences across neighborhoods reflect differences in the components of the composite good. There is, however, considerable overlap in the basket of goods across neighborhoods.[14]

Nonetheless, to ensure that our results are not driven by spurious variation due to missing data we projected, for each product separately, the observed prices on a large set of demographic variables and used the estimated coefficients to impute prices in the neighborhoods with missing prices. As reported below, the demand estimates using the imputed prices are qualitatively the same as those in our leading specification. Thus, the limited changes in the composition of the basket of goods over time or across neighborhoods do not appear to be driving our results.

Beyond this imputation exercise, we checked the robustness of our results to alternative methods of computing the price of the composite good. We used a threshold lower than 21 products, we used only fruits and vegetables, and restricted the sample to supermarkets only. The last two, in particular, ameliorate considerably the missing price problem. Reassuringly, the estimated demand patterns remain qualitatively the same.

Using identical CPI weights for different households is commonplace in the literature, but does not allow tastes to vary with income (Handbury 2013). The uniform weights result in a well-defined single price at each location charged to residents of all origins. This makes our counterfactual analyses more transparent. Nonetheless, we also computed a price index using CPI weights that vary by socioeconomic standing, provided by the CBS. We thus assign differential weights to different origin neighborhoods. This alternative price index has a simple correlation of 0.85 with our index in equation (1) and, not surprisingly, delivers similar demand estimates.[15]

Price differentials. Table 4 provides statistics for the price of the composite good. The

[13]The resulting subsample keeps essentially the same distribution of store formats as the 26 neighborhood sample (57 percent supermarkets, 21 percent market stalls and 12 percent grocery stores).

[14]See Online Appendix Tables D3-D5.

[15]All robustness checks are reported in online Appendix A. In our demand model we partially compensate for the uniform weights by allowing households to derive utility from unobserved aspects of the shopping destinations captured by fixed effects, and by interacting those with the origin's housing prices. We therefore allow the variety of additional products (i.e., beyond our 27) to be valued differently by households of different income levels.

time variation in our sampled prices appears to be in line with the CPI inflation rate.[16] Prices vary significantly across neighborhoods within both commercial and residential districts, with the maximum price being about 16-29 percent above the minimum price. The quantitative importance of cross-neighborhood price variation is manifested in the (gross) savings generated by shopping at the cheapest location in the city. Specifically, Panel C of Table 4 shows the distribution of these savings, $100 \times (p_{jt} - Min_n p_{nt})/p_{jt}$ for each of the 15 neighborhoods with valid prices over all three periods; mean savings are 13 percent.

Another message of Table 4 is that prices in the commercial districts are in general lower than in most residential neighborhoods; the mean price of the composite good is between 6-7 percent higher in the residential neighborhoods. Variation between residential neighborhoods is also considerable: panel B shows that prices in our three NAP (non-affluent peripheral) neighborhoods are typically ranked *above* prices in the three AC (affluent, central) neighborhoods.[17] This observation is central to our research question and so we explore it using two additional figures.

Figure 1 displays the composite good prices in November 2008. It corroborates the observation that some of the highest prices in the city are charged by retailers located in the peripheral, non-affluent neighborhoods. Neighborhoods such as AC2 or AC3 are much more affluent, yet retail prices charged there are lower than those charged at NAP1-NAP3. The figure illustrates that the AC neighborhoods are less isolated and, in fact, are located in the vicinity of the HD supermarkets in the major commercial districts CD1-CD2. Prices in these AC residential neighborhoods are likely disciplined by the lower prices in the commercial districts, whereas no such effect operates in the peripheral neighborhoods.[18]

Figure 2 plots composite good prices against housing prices, along with a fitted regression line.[19] Prices at AC1, for example, are very high — but are perfectly aligned with that neighborhood's affluence level. Prices at the NAP neighborhoods, in contrast, are considerably higher than what can be systematically associated with their affluence level. While the figure has only 15 data points, it highlights the message that these peripheral neighborhoods stand out in terms of the prices charged by their local retailers. In our structural model, we will link these findings to the presence and effects of spatial frictions.[20]

[16]The composite good's price increased by 10 percent between November 2007 and November 2008. For comparison, the CPI inflation for food between December 2007 and December 2008 was 8.3 percent.

[17]AC1, the most affluent neighborhood is, however, usually more expensive than the NAP neighborhoods.

[18]One may wonder why prices at NAP1 are not disciplined by the low prices available at the neighborhood that lies on its southern border (Pisgat Ze'ev North). Discussions with a resident of that area suggest that the supermarket at Pisgat Ze'ev North is not particularly attractive for a weekly shopping trip due to its small size. Such issues motivate the inclusion of destination fixed effects in our model of household preferences.

[19]Commercial districts have a small residential population and therefore we have housing prices there. The linear predicted line suggests a positive relationship between composite good and housing prices. But the small number of data points (15 observations) and the lack of other controls preclude us from reaching general conclusions as to whether "the poor pay more."

[20]While our framework emphasizes differences in absolute prices charged across neighborhoods with different incomes, it is also possible to consider income-adjusted prices. For completeness, we also constructed the (G)EKS-Fisher multilateral price index presented in equation (5) in Deaton and Heston (2010). Reassuringly,

Finally, price rankings are quite persistent: the rank correlation of p_{nt} between September and November 2007 (November 2007 and November 2008) is 0.68 (0.57). This supports our focus on spatial, rather than informational frictions.[21]

2.3 Cross-neighborhood expenditure flow data

We obtained data on consumers' expenditures from a credit card company.[22] Institutional details suggest that customers of this company are not different from customers of other companies. While credit cards are used by 88 percent of the Jewish population, the use of debit cards is minimal in Israel.[23] Our data should therefore be representative of transactions performed via payment cards. Grocery shopping is, of course, also performed using cash and checks, and their use may be correlated with important household characteristics. Our econometric model addresses this measurement problem in detail. We defer discussion of this issue to Section 3.1.

We observe expenditures in supermarkets, grocery stores, bakeries, delicatessen, butcher stores, wine stores, fruits and vegetables stores and health stores — the type of stores where our 27 products are likely to be sold — in the same three periods covered by our price data. The data consist of total expenditures by residents of each origin neighborhood j performed at each destination neighborhood n where $j, n \in \{1, ..., 46\}$. This results in a 46 by 46 matrix of expenditure flows between each pair of neighborhoods for each period.

The data were constructed as follows: first, card holders' neighborhood of residence and their shopping destination neighborhood were identified via customers' and stores' zip codes.[24] The expenditure data were then aggregated to the neighborhood level matrix described above. To be clear, we do not observe data at the individual household or store level.[25]

Table 5 provides statistics regarding the expenditure data. The most popular commercial district is CD1 where, on average, 27 percent of expenditures are incurred. CD1 is the top

the correlation between the simple composite price index used in the paper and the GEKS-Fisher index is very high (the correlation coefficient is 0.89, 0.81 and 0.94 in periods 1-3, respectively).

[21] The rise in the average price in the commercial districts in the third period is entirely due to an unexplained jump in the composite good price at the Romema commercial district (Appendix Table D6). In a robustness check (not reported), we estimated the model without this commercial district, obtaining very similar results.

[22] Anonymous company (2012).

[23] Credit cards are also used by 80 percent of the ultra-orthodox Jewish population (https://www.themarker.com/advertising/1.2413558, in Hebrew) On the lack of use of debit cards, see http://www.antitrust.gov.il/yozma.aspx (in Hebrew).

[24] This required a nontrivial mapping between zipcodes and neighborhoods, where zipcodes can map into multiple neighborhoods. We employed a "majority rule": the zip code was mapped to the neighborhood with which it has the largest geographical overlap.

[25] We also observe total expenditures of each origin neighborhood at destinations outside the city. Jerusalem does not have a substantial ring of satellite cities providing attractive shopping opportunities. We therefore conjecture that much of the observed shopping outside the city corresponds to individuals who have a mailing address in Jerusalem but do not actually reside in it (e.g., students). We therefore do not use these data in our baseline analysis. Robustness checks (not reported) in which we added the expenditures incurred outside Jerusalem to our model's "outside option" (see below) yield remarkably close results to the ones reported in Section 3.2.

destination for residents of 16 to 20 of the 46 neighborhoods (depending on the period). CD2 is at a distant second place, although it is quite popular among nearby neighborhoods such as AC2 (see Appendix Table D7). Most expenditures are not incurred within the home neighborhood, yet home-neighborhood shopping is substantial capturing, on average, 22 percent of total expenditures. The home destination is the top destination in 12 to 17 cases (depending on the period).

As panel B indicates, residents of the non-affluent and peripheral neighborhoods, NAP1-NAP3, have a higher tendency to shop at their home neighborhood relative to the median neighborhood whose share of expenditures at home is 0.16. As we have seen in Section 2.2, this happens despite the fact that shopping "at home" is quite expensive for these residents. We interpret this as evidence for the importance of spatial frictions. We next explore these frictions within a structural model of demand for groceries across the city.

3 A structural model of demand in the city

The following subsections present a model of households' preferences, describe its estimation, and provide results on price and distance elasticities. We also use the model to calculate the "Average Price Paid" (hereafter, APP) for each neighborhood of residence. The APP takes into account the probabilities with which residents of origin j shop in each destination. It therefore allows for a comparison of the cost of grocery shopping across neighborhoods of residence, accounting for shopping outside the home neighborhood.

3.1 A model of household preferences

We posit a discrete choice model in which residents in each of the 46 neighborhoods choose where to perform their grocery shopping. The nested logit framework that we employ is quite standard. We therefore simply spell out our assumptions and the resulting estimation equation; a complete derivation of this equation is in Online Appendix F. That appendix also contains additional discussion and justification for some of our assumptions (in particular, we discuss our emphasis on spatial rather than information frictions, the single shopping trip assumption, and additional forms of unobserved taste heterogeneity). Our treatment of measurement error in the expenditure data, a problem often ignored in applied work, is presented in the text.

Household behavior. A household residing in one of the 46 origin neighborhoods may shop for the composite good in any one of the $n = 1, ..., 15$ destinations where composite good prices are observed in our data. The household may also choose the "outside option" $n = 0$: shopping in one of the remaining 31 neighborhoods. The household's choice maximizes its utility among these 16 options.

Omitting the time index, the (indirect) utility of household h residing in neighborhood j from buying the composite good at store s located in neighborhood n is given by

$$U_{hjsn} = \nu_c + \nu_j + \nu_n + hp_j \cdot \nu_n + \left(\gamma^{-1} \ln y_j - \ln p_{sn}\right) \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn} + \zeta_{hn}(\sigma) + (1-\sigma)\epsilon_{hjsn}$$

where ν_c is a constant that shifts the utility from all "inside options" relative to the utility from the outside option (given the standard normalization of the systematic utility of the outside option to zero). The origin neighborhood fixed effects ν_j capture utility differences across origins with respect to this outside option, while the destination fixed effects ν_n capture quality differences across destinations. Those may include amenities (parking space, opening hours) and differences in product variety (i.e., the availability of products other than our basic 27 items). The destination fixed effects are also interacted with the origin neighborhood's housing prices, hp_j, allowing residents of more affluent neighborhoods to value amenities differently.

The spatial friction — i.e., the utility cost of shopping far away from one's neighborhood — is captured via d_{jn}, the distance (km) between origin j and destination n, which is interacted with x_j, a vector of origin neighborhood characteristics such as the rate of car ownership. This friction is further reflected in h_{jn}, a "shopping at home" dummy variable taking the value 1 if $j = n$. The parameter κ therefore captures benefits of shopping in the home neighborhood on top of the implied savings of travel time (and direct travel costs), already captured by β. Put differently, κ captures a "fixed cost" of shopping outside one's home neighborhood, possibly related to the need to drive, or to give up a parking space near home.

Households' price sensitivity is introduced via the term $(\gamma^{-1} \ln y_j - \ln p_{sn}) \cdot x_j\alpha$ where y_j is the average income in origin neighborhood j and p_{sn} is the composite good price at store s located in destination neighborhood n. This functional form follows Björnerstedt and Verboven (2016) and implies that, conditional on buying at store s in destination n, the quantity (units of the composite good) demanded by household h residing in neighborhood j is $\gamma y_j/p_{sn}$, so that expenditure on the composite good is a constant fraction γ of the (representative) household's income. The fraction γ drops out of the estimation equation and therefore it could vary across origin neighborhoods. We note that more sophisticated discrete-continuous choice models are present in the literature (e.g., Smith 2004, Figurelli 2013). In the context of our aggregate (neighborhood-level) demand data, we favor this simpler modeling strategy. As with the distance sensitivity, the price sensitivity is also interacted with origin neighborhood characteristics.

Finally, the idiosyncratic term $\zeta_{hn}(\sigma) + (1 - \sigma)\epsilon_{hjsn}$, distributed Type-I Extreme Value, follows the representation of the nested logit model in Berry (1994). The nests are destination neighborhoods, allowing stores within a neighborhood to be closer substitutes than stores located in different neighborhoods. This substitution is governed by the parameter σ that takes values in the interval $[0, 1)$, with larger values implying stronger within-neighborhood substitutability. The shock ϵ_{hjsn} captures store-level random variation: for example, a household may particularly value shopping at store s if it is on the way home from work.

We next introduce an assumption, motivated by the nature of our data, regarding symmetry in the systematic utility provided by stores operating in the same neighborhood:

Assumption 1 *Denote by* $\delta_{jsn} = \nu_c + \nu_j + \nu_n + hp_j \cdot \nu_n - \ln p_{sn} \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn}$ *the mean utility level, common to all origin* j *residents who shop at store* s *in destination* n*. Stores within a neighborhood offer identical mean utility levels across households, i.e.,* $\delta_{jsn} = \delta_{jn}$ *for every* j, s, n*.*

Since the only element of δ_{jsn} that depends on the store index s is p_{sn}, this symmetry assumption is consistent with a symmetric (within-neighborhood) price equilibrium, i.e., $p_{sn} = p_n$ for every store s in neighborhood n. This price symmetry will be consistent with the pricing model introduced in Section 4.

Assumption 1 implies that stores within a neighborhood are symmetrically differentiated: they have identical mean utility levels, but offer distinct benefits to individual households via the idiosyncratic error ϵ_{hjsn}. This assumption therefore allows households who reside in different streets within the neighborhood to favor the store nearest to them. It also allows for any other type of horizontal differentiation among stores that is valued idiosyncratically by the neighborhood's residents.

The symmetry assumption accommodates the limitation that we observe expenditures at the neighborhood level rather than at the store level. Tackling such data limitations via a symmetric differentiation assumption is a familiar strategy in the literature (e.g., Berry and Waldfogel 1999). This assumption is not very restrictive in our case because stores within a neighborhood are typically of the same type (e.g., hard discount supermarkets in commercial districts) and, consequently, as shown in Section 2.2, most of the price variation is across, rather than within, neighborhoods.[26]

In Online Appendix F we show that the assumptions spelled out above regarding households' preferences deliver the following linear equation:

$$(2)\quad \ln\left(\frac{E_{jnt}}{E_{j0t}}\right) = \nu_c + \nu_j + (\nu_n + (1-\sigma)\ln L_n) + hp_j \cdot \nu_n + \nu_t - \ln p_{nt} \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn},$$

where E_{jnt} (E_{j0t}) are total expenditures incurred by residents of origin neighborhood j in destination n (in the outside option) at time t. L_n is the number of competitors in destination n which is constant over time.

The left-hand side of (2) contains expenditure shares that are implied by the model but are measured with error in the data. This error stems from two sources: first, observed prices

[26] The symmetry assumption would not be needed if we were to use scanner data, since then we would obtain both price and quantity data at the establishment level. We discussed above, however, the advantages of our approach in which we combine establishment-level price data from the statistical authority — that are easily comparable across space and time — with systematic cross-neighborhood credit card expenditure data. The latter provide an efficient coverage of the shopping probabilities characterizing *all* origin-destination pairs.

pertain to (at most) 27 products, whereas observed expenditures correspond to purchases of many additional products. Second, we observe credit card expenditures rather than total expenditures.

Let $\tilde{E}_{jnt}$ denote expenditures using any payment means on *all* products sold at the relevant establishments. Without loss of generality, we can always express expenditures on the 27 products using any payment means, denoted by E_{jnt}, as a proportion of $\tilde{E}_{jnt}$, $E_{jnt} = \lambda_{jnt}\tilde{E}_{jnt}$, where $0 \leq \lambda_{jnt} \leq 1$. Similarly, our observed *credit-card* expenditures on *all* products, denoted by $\tilde{E}^{cc}_{jnt}$, can also be expressed as a proportion of $\tilde{E}_{jnt}$, $\tilde{E}^{cc}_{jnt} = \tau_{jnt}\tilde{E}_{jnt}$, with $0 \leq \tau_{jnt} \leq 1$. These definitions allow us to map E_{jnt}, which is derived from the model, into the observed expenditures $\tilde{E}^{cc}_{jnt}$ via $\tilde{E}^{cc}_{jnt} = (\tau_{jnt}/\lambda_{jnt})\, E_{jnt}$.

Thus, adding $w_{jnt} = \ln\left(\frac{\tau_{jnt}}{\lambda_{jnt}}\frac{\lambda_{j0t}}{\tau_{j0t}}\right)$ to the right hand side of (2) allows us to use the observed $\ln\left(\tilde{E}^{cc}_{jnt}/\tilde{E}^{cc}_{j0t}\right)$ as the dependent variable. The error term w_{jnt}, generated by the mismeasurement of expenditures, presents an identification challenge because the proportionality factors λ and τ are likely to be correlated with origin and destination neighborhood characteristics. For example, residents of less affluent origin neighborhoods may have a higher than average tendency to use cash, and the use of cash may also be more prevalent in certain destinations (e.g., open fresh produce market). Our strategy for dealing with this endogeneity is to soak up such tendencies into origin and destination fixed effects via the following assumption:

Assumption 2 *Conditional on origin, destination and time fixed effects, w_{jnt} is uncorrelated with prices and distances.*

This assumption implies that the proportionality factors may depend on fixed neighborhood characteristics but can not depend on prices and distance, given these characteristics. It therefore allows for tendencies of residents of particular neighborhoods to use more or less cash in certain destinations but assumes that such tendencies are accounted for by the fixed effects. The panel structure of the data indeed allows us to control for such fixed effects.

Let u_{jnt} be the error from linearly projecting w_{jnt} on a set of origin, destination and time fixed effects. Assumption 2 allows us to rewrite (2) as

$$(3) \qquad \ln(\tilde{E}^{cc}_{jnt}/\tilde{E}^{cc}_{j0t}) = \phi_c + \phi_j + \phi_n + \phi_t + hp_j \cdot v_n - \ln p_{nt} \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn} + u_{jnt}$$

where the $\phi' s$ are fixed effects. Given the above assumptions, equation (3) is amenable to consistent estimation via OLS. The observations used consist of all triplets (j, n, t) pertaining to origin neighborhood j, destination neighborhood n and time period t. The sample size should therefore equal $46 \times 15 \times 3 = 2,070$ observations, corresponding to expenditure data of residents of 46 origin neighborhoods at 15 destination neighborhoods over the three time periods. While the model predicts a positive expenditure share by residents of any origin j at any destination n, observed expenditures $\tilde{E}^{cc}_{jnt}$ are zero in about 12 percent of all potential observations. We drop such observations from the sample, reducing its size to $1,819$ observations. The results are qualitatively robust to substituting a very small number for $\tilde{E}^{cc}_{jnt}$ (online Appendix A).[27]

Some of the model's parameters are not identified. First, the origin, destination and time dummies (ϕ_j, ϕ_n, ϕ_t) do not identify the *utility effects* (v_j, v_n, v_t) but rather confound them with that part of the measurement error w_{jnt} that is correlated with the dummies. An additional assumption will therefore be required for the computation of elasticities and other quantities of interest. The consistent estimation of (α, β, κ), however, only requires the assumptions stated above.

Second, the parameter σ that captures the degree of within-neighborhood competition is unidentified absent time series variation in the number of competitors in destination n, L_n. This happens because the fixed effect ϕ_n captures the sum of the utility terms $v_n + (1 - \sigma) \ln L_n$ as well as the linear projection of w_{jnt} on the destination dummy variable. Our practical solution is to calibrate σ so that it generates reasonable markups given the identified parameters.[28]

Based on conversations with people familiar with the industry, retail markups of 15-25 percent are reasonable for the type of products studied in this paper.[29] Section 4, where the pricing model is introduced, shows that setting $\sigma = 0.7$ yields an average (median) markup of 22 (20) percent and therefore this is the value chosen for σ. Setting this value to 0.8 or 0.9 instead makes no difference for the qualitative findings in this paper.[30] Online Appendix F provides additional discussion of identification and of our assumptions.

3.2 Estimation results

Table 6 shows OLS estimates of equation (3). We employ 2-way clustering of standard errors at the origin and destination level, allowing for arbitrary correlation across observations sharing an origin or a destination. The different specifications control for different sets of fixed effects and socioeconomic interactions. Across all specifications, the coefficients have the expected signs. Coefficients on log price and distance (which we entered with a negative sign) are positive, and so is the coefficient on the "shopping at home" dummy variable, consistent with the high tendency towards home neighborhood shopping observed in the data (Table 5).

Column (4) includes the full set of origin, destination and period dummies required by

[27] This is not a formally valid correction but one often used in practice (see Gandhi, Lu and Shi 2013 for a partial identification approach).

[28] Such an approach has some precedence in the literature. For example, Björnerstedt and Verboven (2016) calibrate a conduct parameter to generate reasonable markups. We could alternatively pin this parameter down by incorporating supply-side restrictions into the estimation procedure. We favor the calibration as it eliminates the need to rely on our pricing model in generating the demand estimates.

[29] Note that these are markups above marginal cost. They are, therefore, higher than markups over average costs, the latter often approximated using information from retailers' financial reports.

[30] We also regressed the estimated fixed effects $\hat{\phi}_n$ on $\ln L_n$ to estimate $1 - \sigma$. This yields an (imprecisely estimated because of the small number of observations) estimate of $\hat{\sigma} = 0.81$. This estimate is likely to be biased since v_n and the projection of w_{jnt} on v_n are likely to be correlated with L_n. Nevertheless, it is somewhat comforting that the calibrated and estimated values are similar in magnitude.

our theory, but without interacting the main regressors with demographics. Both the price and distance effects have the expected sign and are statistically significant. The inclusion of destination fixed effects substantially increases the regression's goodness of fit from 0.38 in column (2) to 0.66 - 0.78 in columns (3)-(8). This shows the importance of controlling for unobserved amenities (e.g., availability of parking, opening hours, product variety etc.).

Columns (5)-(8) then allow for interactions of the price and distance sensitivities with characteristics of the neighborhood of origin. Households in richer neighborhoods, as proxied by housing prices, are significantly less sensitive to prices. Distance sensitivity is quite robust to the inclusion of additional regressors. It is higher in neighborhoods with a large fraction of elderly residents, though this interaction is not statistically significant. Senior individuals may face a lower cost of time, but, on the other hand, may find shopping at other neighborhoods more challenging. The distance sensitivity is smaller in neighborhoods where the share of residents who own a car, or drive to work, is higher, but these effects are not statistically significant.

In column (6), we add the interaction of origin housing prices with destination dummies, allowing for different valuations of unobserved amenities across socioeconomic classes. The estimated price coefficient is mildly reduced (from 5.1 to 4.7). Distance coefficients are also only minimally affected except for the interaction with the percentage of senior citizens.

We adopt column (6) as our baseline specification since it fully and most flexibly controls for effects that we expect to be important in the household's decision problem. In particular, the added interaction term between origin housing prices and destination fixed effects controls for a broad range of scenarios involving unobserved heterogeneity that could potentially bias our results (see the discussion in Online appendix F). Overall, estimates in columns (7) – (8) are very close to the baseline specification in column (6).[31]

Elasticities. The economic implications of these estimates are captured in price and distance elasticities. The own-price elasticity faced by store s in neighborhood n is (Online Appendix F):

$$\eta_{sn,p} = \frac{p_{sn}}{Q_{sn}}\frac{\partial Q_{sn}}{\partial p_{sn}} = -\sum_{j=1}^{J}\frac{Q_n^j}{Q_n}\left[1 + x_j\alpha\left(\frac{1}{1-\sigma} - \frac{\sigma}{(1-\sigma)\,L_n} - \frac{\pi_{jn}}{L_n}\right)\right] \quad (4)$$

where Q_{sn} is the total demand at store s located in neighborhood n. π_{jn} is the probability that a resident from origin j shops in neighborhood n. The total demand faced by all retailers in neighborhood n is denoted by Q_n, whereas Q_n^j is the part of this demand generated by residents of origin j. The demand elasticity faced by the store is therefore a quantity-weighted

[31] In column (7), the price coefficient changes because its interaction with housing prices is omitted. However, the own price elasticities implied by columns (6) and (7) are nearly identical, differing by 5% on average. Moreover, interacting price with family size (not reported) yields an insignificant effect and does not alter the other coefficients.

average of origin-specific elasticities, where the weights depend on the fraction of the retailers' demand generated by residents of those origins.

The semi-elasticity of Q_n^j with respect to the distance between j and n is

$$\eta_{jn,d} = \frac{1}{Q_n^j} \frac{\partial Q_n^j}{\partial d_{jn}} = -x_j \beta \left(1 - \pi_{jn}\right),$$

measuring the percentage change in demand from residents of neighborhood j at destination $n \neq j$ in response to a 1 km increase in the distance between these neighborhoods.

Estimating the elasticities requires the estimated parameters obtained above (including the calibrated σ), an estimate of the choice probabilities π_{jn}, and data on the number of stores in destination n, L_n. We set L_n equal to the number of the neighborhood's supermarkets in 2008.[32] We do not directly count grocery stores, or other non-supermarket retail establishments, as these are not close substitutes to supermarkets for the purpose of the weekly grocery shopping (e.g., because of limited availability of items). Nonetheless, to partially take these additional store formats into account, we add the value of one to L_n in residential neighborhoods, while keeping it equal to the number of supermarkets in the commercial districts. This modification results in more reasonable estimated margins and has a negligible effect on the qualitative findings of the counterfactual analyses reported in Section 5.

Estimation of the choice probabilities π_{jn} is complicated by the fact that these depend on the mean utility levels and, therefore, on the utility fixed effects (ν) which are not identified. Appendix F shows that the mean utility levels are identified under the following assumption:

Assumption 3 *For each origin j, the ratio $\frac{\tau_{jn}}{\lambda_{jn}}$ is identical for all destinations n.*

This assumption implies that choice probabilities are equal to the observed expenditure shares. This is clearly weaker than simply assuming their equality which amounts to ignoring the measurement error altogether. We stress that Assumption 3 is not required for the consistent estimation of the parameters α, β, κ. Our framework therefore clarifies the different sets of assumptions that can be used to accomplish different goals in the presence of measurement error in the expenditure data.

Employing the leading specification (column 6 of Table 6) we estimate price elasticities for each destination, and distance semi-elasticities for each origin-destination pair. Table 7 presents estimates for the last period, November 2008, which are nearly identical to the average over the three periods. The average (median) store-level own *price elasticity* $\eta_{sn,p}$ is 4.82 (4.95) in absolute value. The individual estimates are tightly distributed around the mean. Recalling that close substitutes are often available in the form of other stores within the same

[32]The number of supermarkets is shown in the last column of Online Appendix Table C2. This value includes all supermarkets, not just those where prices were sampled. A specific issue arises with respect to the open market of Mahane Yehuda where many small sellers – open stalls – are present. To retain internal consistency, we set $L_n = 2$ in that location (because there is a small supermarket in the neighborhood). Using different values affects the margins for retailers in this specific neighborhood, but does not affect the qualitative findings.

neighborhood, this relatively-elastic demand seems reasonable. Increasing σ to 0.8 generates a higher mean price elasticity of 6.43 but, as reported in online Appendix B, it makes no difference in terms of our qualitative conclusions.[33] Online Appendix A presents the robustness of the estimated elasticities to alternative computations of the price of the composite good discussed in Section 2.2.

The average (and median) *distance semi-elasticity* $\eta_{jn,d}$ is 0.35 in absolute value implying that a 1 km increase in the distance between an origin j and a destination n decreases demand by residents from j at n by 35 percent. Spatial frictions are, therefore, a first-order consideration affecting households' choices, consistent with the anecdotal evidence surveyed in the Introduction.

To assess the price-distance trade-off, we consider residents of location j who shop at destination n. The maximum percentage price increase these consumers are willing to accept for destination n to become one kilometer closer to them is $100\,(\exp\left(x_j\beta/x_j\alpha\right)) - 1)$. The median of this estimated quantity over the 46 origin neighborhoods is 24.5 percent indicative, again, of a substantial spatial dimension in households' preferences.

3.3 Retail price differences revisited: the *Average Price Paid*

In Section 2.2 we showed that prices in non-affluent, peripheral neighborhoods (NAP) were often higher than prices in more affluent but centrally located neighborhoods. This, however, does not necessarily imply that households in NAP neighborhoods pay more for groceries because they may not shop in their neighborhood of residence. Our estimated model allows us to revisit this issue.

Recall that, given Assumption 3, the probability that a resident from neighborhood j buys the composite good in neighborhood n, π_{jn}, can be computed directly from the expenditure data. We can therefore compute the *Average Price Paid* (APP) for residents of neighborhood j: $p_j^A \equiv \sum_{m=0}^{N} \pi_{jm} p_m$. This weighted average takes into account the probabilities with which residents of origin j shop in each destination n. It therefore allows a comparison of the cost incurred by residents of different origin neighborhoods, taking into account their shopping behavior.[34]

Figure 3 plots the Average Price Paid against housing prices in each of the 46 neighborhoods in November 2008 (along with a linear predicted line). Interestingly, the centrally-located, affluent neighborhood AC1 has an APP which is perfectly explained by its affluence level. This was also true for the price charged by retailers operating in this neighborhood, as we saw in

[33] Further, increasing σ to 0.9 makes demand even more elastic, which is intuitive, but once again does not affect our qualitative conclusions. Details are available from the authors upon request.

[34] This requires also an estimate of the price of the composite good at the outside option neighborhoods, p_0, which is unobserved. These outside option neighborhoods are residential neighborhoods where we believe most shopping opportunities are at expensive grocery stores. We therefore set p_0 as the price charged in NAP3: the peripheral, non-affluent neighborhood that launched the consumer boycott in 2014.

Figure 2, which we refer to as the "posted" price. The three non-affluent, peripheral neighborhoods NAP1-NAP3 have an APP that lies above the predicted line, although this gap is less pronounced relative to the gap observed in Figure 2 for the posted prices charged in these neighborhoods. This suggests that despite the ability to shop outside the home neighborhood, residents of these NAP neighborhoods still pay more, on average, than what could be explained by their affluence level, consistent with the spatial friction. In 8 out of the 11 residential neighborhoods with valid prices, the APP is substantially lower than the observed price, reflecting the savings afforded to households by shopping outside their home neighborhood (when p_j^A is higher than p_j, the difference is small).

It is of interest to compare p_j^A with the minimum price across all 15 neighborhood, which would have been the price actually paid if households were to determine their shopping destination based on price only (ignoring equilibrium effects). The APP is, on average, 12.2 percent higher than this minimum price (the range being between 3.7 and 21.2 percent). This number reflects the monetary value of the spatial frictions faced by households (captured in the model via β and κ) as well as their preferences for specific shopping destinations (captured by v_n and the idiosyncratic terms). It also provides a rough indicator of the maximal extent to which prices can be expected to decline were these frictions removed.

Importantly, the Average Prices Paid at the peripheral, non-affluent neighborhoods NAP1-NAP3 are higher than those faced by residents of more affluent neighborhoods that are located closer to the commercial areas, AC2-AC3. Moreover, a strong, positive relationship is depicted in Figure 4 between the APP p_j^A and distance to the main commercial district CD1. This is yet another manifestation of the role played by spatial frictions in determining the variation in the cost of grocery shopping across households.

We next introduce a model of pricing decisions and employ it, along with the estimated demand system, to tease out the effect of spatial frictions on the cost of groceries for residents of neighborhoods across the city, and, specifically, in the NAP neighborhoods. A supply model is needed in order to allow retailers to adjust their pricing decisions in counterfactual scenarios. This will be necessary, for example, when calculating the response of equilibrium prices across the city to a city-wide reduction in the travel cost.

4 Retail supply: pricing decisions

Our model for the supply side of the market is summarized in the following assumption:

Assumption 4 *(i) Each neighborhood n features L_n retailers that share a constant-in-output, symmetric marginal cost c_n. (ii) Retailers across the city engage in Bertrand competition: each store s in each neighborhood n chooses its price p_{sn} simultaneously. (iii) A unique, interior Nash equilibrium in prices exists. (iv) Equilibrium prices can differ across neighborhoods, but satisfy within-neighborhood symmetry: $p_{sn} = p_n$ at each store s in each neighborhood n.*

Part (i) of Assumption 4, along with the symmetric differentiation assumption from the demand model (Assumption 1) naturally allow us to focus on the (within-neighborhood) symmetric price equilibria assumed in part (iv).

The standard first order conditions from this model (see online appendix F) imply that the margin $(p_n - c_n)/c_n$ is inversely related to the own-demand elasticity in (4). Margins are therefore intuitively tied down to the model's primitives. Specifically, the margin garnered by neighborhood-n retailers increases in π_{jn} because it reflects neighborhood-j residents' tendency for shopping at n. This effect is mediated via demographics: it is stronger, the lower is the sensitivity of residents of j to price, reflected in a high value of $x_j\alpha$. The effect also increases in the share of sales by neighborhood n retailers to households from neighborhood j, Q_n^j/Q_n. In a *residential* neighborhood n, the term Q_n^n/Q_n – the fraction of the sales by retailers located at n made to residents of the same neighborhood – is usually large and will be dominant in determining the margin at n. If n is a peripheral neighborhood, π_{nn} will be large (Table 5) — implying an inelastic demand working in the direction of increasing the retail margin.

Table 8 displays the estimated costs and margins by neighborhood in November 2008, the time period in which we conduct the counterfactual analyses. Very similar quantitative and qualitative patterns obtain when averaging over the three time periods. Using the baseline value $\sigma = 0.7$, the average (median) estimated margins are 22 (20) percent. Conversations with people familiar with the retail industry in Israel suggest that this is a reasonable margin given the type of products considered in this paper. Indeed, this value for σ was chosen precisely for this reason (see Section 3.1). We also compute margins assuming $\sigma = 0.8$, generating somewhat lower margins but the same qualitative counterfactual conclusions (online Appendix B).

Margins in residential neighborhoods are generally higher than those in the large commercial districts. Our model attributes this to both spatial frictions and to low within-neighborhood competition in residential areas. Furthermore, marginal costs at the non-affluent, peripheral neighborhoods NAP1-NAP3 are particularly high. This may reflect the cost of transporting goods into more remote locations, and the lower operational economies of scale obtained at the relatively smaller supermarkets located there.[35]

5 Counterfactuals: the impact of spatial frictions

We perform counterfactual analyses to assess the role of spatial frictions in generating the city's price equilibrium. In a first exercise we make travel less costly, and in a second one we increase the appeal of the commercial districts. Intuitively, both changes should make households more willing to shop in the commercial districts. In a third exercise, we artificially increase the number of retailers in residential neighborhoods. We examine the impact of these changes on

[35] We do not consider multi-store pricing by chains that operate supermarkets in both residential and commercial districts. Because these chains' pricing in the commercial districts is strongly constrained by the presence of hard discounters that operate only there, we expect this issue to have little impact on our findings.

(i) the equilibrium posted prices, i.e., the prices charged by retailers in each neighborhood (ii) the probabilities with which residents of each origin neighborhood shop at each destination, and (iii) the Average Price Paid (APP) by residents of each neighborhood taking these probabilities into account.[36]

We use the baseline estimates from column 6 in Table 6 and $\sigma = 0.7$. Online Appendix B reports counterfactual results using the value $\sigma = 0.8$, delivering very similar results (as is also the case for $\sigma = 0.9$, with details available from the authors upon request). Online Appendix G provides technical details on the computation of the counterfactual equilibria.

Table 9 summarizes the impact on the posted prices charged by retailers across the city.[37] The first column corresponds to the observed posted prices whereas the other columns report the predicted percentage changes to those prices in each counterfactual experiment.

The first experiment reduces the distance disutility by half: that is, we add $0.5d_{jn}x_j\beta$ to the utility garnered by residents of each origin j from shopping in each destination n. The second experiment reduces spatial frictions even further by reducing, in addition, the preference for shopping at home parameter κ by 50 percent. As Table 9 shows, these two experiments reduce the median (across the 15 neighborhoods with observed prices) prices by 0.7 to 1 percent. Median posted prices within the 11 residential neighborhoods with observed prices are reduced by only one half of a percent. These are quite modest declines. Examining our non affluent, peripheral neighborhoods, NAP1-NAP3, we see that prices are reduced by 0.5 percent in NAP3 (where the boycott took place), but actually *increase* in NAP1 and NAP2.

At first glance these increases appear puzzling since a reduction in spatial frictions should enhance competition across the city, exerting a substantial downward pressure on retail prices in the residential neighborhoods, and certainly in the peripheral ones. Why, then, do prices decline only mildly or even increase? The answer lies in the changes in the composition of demand faced by retailers in these neighborhoods. As travel becomes less costly, households that continue shopping in the expensive residential neighborhoods are those with very large idiosyncratic shocks favoring shopping there. Retailers in these neighborhoods therefore face a less elastic demand prompting them to raise prices, providing a countervailing force that diminishes, and sometimes even offsets, the competitive force.[38]

A similar picture arises when improving the amenities in the city's main commercial district, CD1, and at the two major ones, CD1-CD2, respectively. This is performed by increasing the destination fixed effects ν_n associated with each such district by one standard deviation.[39] This

[36]The model allows us to compute the impact on welfare but as our exercises directly affect utility parameters we find this less appealing.

[37]Online Appendix E provides complete neighborhood-specific results for all the counterfactuals.

[38]This is similar to the observed "generic drug paradox" that occurs when many, but not all, consumers switch to newly available generic drugs but prices among the incumbent (brand) drugs do not decline and even rise (Griliches and Cockburn, 1994).

[39]Given that ν_n is unidentified due to measurement error, we use the standard deviation of ϕ_n, the fixed effect that confounds ν_n with the measurement error effect, instead. One standard deviation of the distribution of ϕ_n may be greater than one standard deviation of the distribution of ν_n. This issue, however, does not drive our

may correspond to various improvements in the shopping experience in these districts: for example, the city may improve the physical infrastructure by setting up large parking spaces at the entry points to the commercial district with a convenient shuttle service.[40] Boosting the utility of shopping at n can make the citywide grocery market more competitive. But the same countervailing force applies here so, again, only mild price reductions are observed (median price declines of 0.1 – 0.8 percent in residential neighborhoods, and of 0.1 – 0.9 percent in NAP1-NAP3).

Finally, the last column considers the effect of exogenously increasing the number of competitors in each residential neighborhood n by 1. The median posted price decline across residential neighborhoods is 3.4 percent, whereas the price declines in the NAP1-NAP3 neighborhoods are 1.3 – 3.5 percent. Such additional entry, however, may be associated with substantial social opportunity costs due to zoning restrictions and lack of space. A price reduction of about 3.5 percent may not be large enough to justify such costs.

Whereas equilibrium posted prices respond very mildly, households' shopping behavior changes markedly: many households switch to shopping in the affordable commercial districts. This is captured by the APP, the weighted average of prices paid by residents of each neighborhood, taking into account their probability of shopping in the various destinations. Table 10 summarizes the counterfactual changes to the APP faced by residents of the eleven residential neighborhoods where prices are observed.[41]

The first column of Table 10 reports the APP faced by residential neighborhoods in the observed equilibrium. The percentage reduction in APP faced by neighborhood residents following a reduction in the spatial friction is much larger than the corresponding percentage reduction in *prices charged in* the neighborhood as reported in Table 9. As the top row of Table 10 indicates, in the first four experiments the median (over residential neighborhoods) APP falls by 1.8 - 5.6 percent, versus the 0.1 - 0.8 percent fall in median posted prices shown in Table 9.

The Average Price Paid, therefore, responds much more strongly than the posted price to changes in spatial frictions. A median average reduction of 5.6 percent is quite substantial because, as remarked in Section 3.3, the average difference between the APP and the minimum price in the observed equilibrium is about 12.2 percent which can be taken as an upper bound to the price effect in our counterfactual exercises.

In short, while Table 9 considers only the impact on equilibrium prices charged in different locations, the Average Prices Paid in Table 10 take also into account the changes in shopping

findings. We obtain very similar qualitative findings by adding one half of a standard deviation of ϕ_n instead.

[40] Interestingly, the city of Jerusalem recently announced plans to improve the main commercial district, CD1, exactly along these lines: "The plan: the Talpiot industrial zone expected to undergo a revolution over the next decade," Kol Hair (April 15, 2016).

[41] For completeness, Appendix Table E2 shows the impact on the APP for all 46 neighborhoods, delivering the same qualitative conclusions. We favor presenting here results for the 11 residential neighborhoods where prices are observed to facilitate comparison with the impact on posted prices displayed in Table 9.

patterns induced by the changes in parameters. This is evident in Figure 5 that compares the probability of shopping at CD1, the city's most important commercial district, in the observed equilibrium, to the same probability under the counterfactual that improves amenities in that district. The probability of shopping at CD1 increases for residents of all neighborhoods, and substantially more for those located in the periphery. While the price charged at CD1 increases slightly, it is still low, and, as a consequence, the Average Prices Paid decline considerably.

Viewed through the lens of its effect on the APP, the benefits from decreasing spatial frictions to the *average resident* of the three disadvantaged neighborhoods NAP1-NAP3 are substantial. When amenities at CD1 are improved, the APP incurred by residents of NAP3 — the neighborhood where the first boycott took place — drops by a substantial 7 percent (Table 10), whereas the posted price charged by the retailers at NAP3 dropped by 0.6 percent only (Table 9). Average Prices Paid at NAP1 and NAP2 drop by 2.2 and 6.6 percent, respectively, whereas posted prices charged by the retailers in both of these neighborhoods only drop by at most 0.1 percent.

Discussion. Evaluating the impact of a reduction in spatial frictions by considering only the effect on posted prices would be misleading: it would suggest very mild benefits, if at all. In contrast, the analysis that also considers the impact on shopping probabilities, embedded into the computation of the Average Price Paid, suggests substantial reductions in the cost of grocery shopping. This point applies to residential neighborhoods in general, and not only to the disadvantaged ones.

Reducing spatial barriers or improving amenities at commercial districts may confer substantial benefits to the average resident of the peripheral, non-affluent neighborhoods.[42] However, because prices charged at the NAP neighborhoods barely change, or even increase, a reduction in spatial frictions does little to benefit residents with limited mobility (e.g., the young families described as having no time to shop, or the elderly). Those residents will continue to pay the expensive prices in their home neighborhood. Importantly, these are precisely the household segments identified by the boycott organizers as experiencing the most harm from the price differentials in the first place. Alleviating the cost of living for such individuals may therefore require more targeted relief programs.[43] Increased competition, in the form of gradual improvements to the city's infrastructure, will fail to provide relief to such populations.

6 Summary and conclusions

This paper uses a unique dataset on prices in spatially-differentiated neighborhoods within a large metropolitan area, and on the distribution of expenditures across these neighborhoods,

[42] A caveat to this statement is that the attractiveness of a location v_n is probably endogenous and might change if it experiences a substantial increase in shopping activity. Addressing this would require a model of retailers' choice of amenities which is beyond the scope of this paper.

[43] One possibility would be to add location and mobility, in addition to income, to the criteria determining eligibility into welfare programs, e.g., food stamps.

to explore the determinants of price differentials and shopping patterns within the city. We document that retailers at several peripheral, non-affluent neighborhoods often charge higher prices than retailers located in more centrally located, affluent neighborhoods.

Using an estimated structural model of demand and supply, we establish that spatial frictions play an important role in generating these patterns. Our counterfactual analysis reveals that alleviating spatial frictions brings substantial benefits to the average resident of peripheral neighborhoods. This operates by improving access to hard discount supermarkets located in the commercial districts. The prices charged in the residential neighborhoods themselves, however, do not decline much, and sometimes even increase further exacerbating the cost of living conditions for households with limited mobility.

Our simple model can be extended in future work to accommodate multi-store pricing by retail chains, or more complicated demand systems. The parsimony of the model presented here has the important benefit that the demand model can be estimated via linear regressions. The model is capable of producing reasonable predictions that are consistent with institutional details and anecdotal evidence regarding the nature of retail spatial competition within an urban setting. We view the paper as a step toward a better understanding of the role played by spatial frictions in determining the cost of grocery shopping across a city.

Our analysis reveals the benefits from enhanced household mobility. It also reveals that better mobility, while strengthening competition across retailers, will not necessarily lead to lower prices in periphery neighborhoods, perhaps motivating more targeted policies aimed at alleviating the cost of living for low-mobility households residing there.

References

[1] Adams, B. and K.R. Williams (2019): "Zone Pricing in Retail Oligopoly," *American Economic Journal: Microeconomics*, 11(1), pages 124-156.

[2] Aguiar, Mark and Erik Hurst (2007) "Life-Cycle Prices and Production," *American Economic Review*, American Economic Association, vol. 97(5), pages 1533-1559

[3] Aguirregabiria, V. and J. Suzuki (2015): "Empirical Games of Market Entry and Spatial Competition in Retail Industries," *CEPR Discussion Paper* No. 10410

[4] Anonymous Company 2007-2008. "Expenditure data [database]". (accessed on 17-10-2012)

[5] Baye, M.R., J. Morgan, and P. Scholten (2006): "Information, Search, and Price Dispersion," in: Terrence Hendershot (Editor), *Handbook on Economics and Information Systems*, Elsevier.

[6] Berry, S.T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2): 242-262

[7] Berry, S., Levinsohn, J., and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4): 841-90

[8] Björnerstedt, J., and F. Verboven (forthcoming in 2016): "Does Merger Simulation Work? Evidence from the Swedish Analgesics Market," *American Economic Journal: Applied Economics.*

[9] Caplin, A., and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1): 25-59

[10] Caplovitz, D. (1963). *The poor pay more*. New York: Free Press of Glencoe.

[11] Chintagunta, P., J.-P. Dubé, and V. Singh (2003): "Balancing Profitability and Customer Welfare in a Supermarket Chain," Quantitative Marketing and Economics, 1(1): 111–147.

[12] Chung, C., and S.L. Myers (1999). "Do the poor pay more for food? An analysis of grocery store availability and food price disparities." *Journal of consumer affairs*, 33(2), 276-296.

[13] Davis, P. (2006): "Spatial competition in retail markets: movie theaters," *The RAND Journal of Economics*, 37(4), 964–982

[14] Davis, R.D., Dingel, J.I., Monras, J., and E. Morales (2019), "How Segregated is Urban Consumption?," *Journal of Political Economy*, 127(4), 1684-1738

[15] Dubois, P., and S. Jódar-Rosell (2010): "Price and Brand Competition between Differentiated Retailers: A Structural Econometric Model," *Mimeo*

[16] Ellickson, P.B., Grieco, P.L. and O. Khvastunov (2016), "Measuring Competition in Spatial Retail", *Mimeo*.

[17] Felsenstein, D. 2007-2008. Housing prices at the Jerusalem Subquarter level [database] (accessed 03-07-2012)

[18] Figurelli, L. (2013): "Store Choice in Spatially Differentiated Markets," *Mimeo*

[19] Gandhi, A., Z. Lu, and X. Shi (2013): "Estimating Demand for Differentiated Products with Error in Market Shares," *Mimeo*

[20] Griliches, Z., and I. Cockburn (1994), "Generics and New Goods in Pharmaceutical Price Indexes," *American Economic Review*, 84(5), 1213-32.

[21] Handbury, J. (2013) "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities", *Mimeo*.

[22] Hotelling, H. (1929): "Stability in competition," *Economic Journal* 39, 41-57

[23] Houde, J.-F. (2012): "Spatial differentiation and vertical contracts in retail markets for gasoline," *American Economic Review,* 102(5): 2147-82

[24] Kurtzon, G., and R. McClelland (2010): "Do the Poor Pay More Store-by-Store?," BLS Working Paper 437

[25] Israel Central Bureau of Statistics. 2007-2008. Grocery price data [database] (accessed 02-24-2012).

[26] Israel Central Bureau of Statistics. 2012. Distances between Geographic units [database] (accessed 02-15-2012).

[27] Israel Central Bureau of Statistics. 2008. "The 2008 Census of Population. https://www.cbs.gov.il/en/subjects/Pages/The-2008-Census-of-Population.aspx (accessed 20-07-2020).

[28] MacDonald, J. M., and P. E. Nelson (1991). "Do the poor still pay more? Food price variations in large metropolitan areas." *Journal of Urban Economics*, 30(3), 344-359.

[29] McManus, B. (2007): "Nonlinear pricing in an oligopoly market: The case of specialty coffee," *RAND Journal of Economics* 38, 512-532

[30] Miller, N.H. and M. Osborne (2014):"Spatial differentiation and price discrimination in the cement industry: evidence from a structural model." *The RAND Journal of Economics* 45(2): 221-247

[31] Moretti, E. (2013). "Real Wage Inequality," American Economic Journal: Applied Economics, 5(1): 65-103.

[32] Nocke, V., and N. Schutz (2016), "Multiproduct-Firm Oligopoly: An Aggregative games Approach," *CEPR DP* 11539m a structural model," *RAND Journal of Economics* 45(2): 221-247

[33] Pinske, J., M. E. Slade, and C. Brett (2002):"Spatial price competition: A semiparametric approach," *Econometrica* 70(3), 1111-1153

[34] Richburg-Hayes, L. (2000). "Do the poor pay more? An empirical investigation of price dispersion in food retailing". Princeton Dept. of Econ., Industrial Relations Working Paper, (446).

[35] Salop, S. (1979): "Monopolistic competition with outside goods," *Bell Journal of Economics* 10, 141-156

[36] Smith, H. (2004): "Supermarket Choice and Supermarket Competition in Market Equilibrium," *Review of Economic Studies*, 71(1), 235–263

[37] Thomadsen, R. (2005): "The effect of ownership structure on prices in geographically differentiated industries," *RAND Journal of Economics* 36, 908-929

7 Tables and Figures

Table 1: Distribution of demographics across Jerusalem neighborhoods

Variable	N	mean	sd	min	p25	p50	p75	max
Population (000s)	46	15.0	5.3	6.2	10.5	13.9	18.3	28.7
Households (000s)	46	4.4	1.6	2.1	3.3	4.2	5.3	8.8
Average household size	46	3.4	0.9	1.9	2.8	3.3	4.1	6.1
Housing prices (000s)	46	13.4	3.0	8.8	11.5	13.3	15.2	21.1
% Driving to work	46	39.7	18.6	7.5	23.8	47.2	55.3	68.1
% Car ownership	46	48.9	22.9	6.9	34.4	59.2	65.9	89.3
% Senior citizens	46	10.6	4.9	1.1	7.5	10.2	14.4	25.6

Notes: Housing prices = the 2007-2008 average price per square meter in thousands of NIS, Driving to work = percentage of individuals above 15 years of age who used a private car or a commercial vehicle (as a driver) as their main means of getting to work in the determinant week. Car ownership = percentage of households using at least one car. Senior citizens = percentage of individuals above age 65. Source: Central Bureau of Statistics (CBS).

Table 2: Distribution of distance between Jerusalem neighborhoods (in km)

Variable	N	mean	sd	min	median	max
Distance to City center	46	4.3	2.3	0.6	4.3	9.2
Distance to Commercial District 1 (CD1)	46	5.7	2.9	0.0	5.4	13.2
Distance to Commercial District 2 (CD2)	46	6.0	2.5	0.0	5.8	12.0
Mean distance to all other neighborhoods	46	6.1	1.6	4.2	5.8	10.8

Notes: Statistics of the distribution of distances in kilometers between each neighborhood and 1) the city center, 2) the two prominent commercial centers CD1 and CD2, and 3) all other neighborhoods. Source: CBS

Table 3: Number of sampled stores and of observed products

	# sampled stores			# observed products			# supermarkets
	Sep2007	Nov2007	Nov2008	Sep2007	Nov2007	Nov2008	
A. Statistics over all 26 neighborhoods where prices were collected							
Mean	2	2	2	18	17	17	1
Min	0	0	0	0	0	0	0
Max	10	10	9	27	27	27	5
Total	54	55	51				29
B. Particular neighborhoods of interest							
NAP1	1	1	1	27	27	27	1
NAP2	2	2	2	27	27	27	2
NAP3	3	2	2	27	26	26	1
AC1	2	2	2	24	25	24	1
AC2	3	3	3	27	27	27	2
AC3	1	1	1	26	25	23	1
CD1	7	7	7	27	27	27	5
CD2	3	3	3	27	27	26	3

Notes: Statistics regarding the number of stores and products sampled across the city. CD, NAP and AC correspond to Commercial Districts, Non-Affluent Peripheral, and Affluent Central neighborhoods (see text). The number of supermarkets (most-right column) includes all supermarkets in the neighborhood, not just those where prices were sampled. Some neighborhoods are not sampled in all three periods and therefore have zero stores in at least one period. Many of the ten stores sampled in the main open market are fresh produce stalls.

Table 4: Price of composite good across Jerusalem neighborhoods

A. Statistics over 15 neighborhoods with at least 21 observed price items

	Prices (NIS)		
	Sep-07	Nov-07	Nov-08
Mean residential (11)	7.40	7.19	7.92
Mean commercial (4)	6.91	6.81	7.46
Min residential (11)	6.23	6.56	7.36
Min commercial (4)	6.33	6.15	6.89
Max residential (11)	8.01	7.61	8.52
Max commercial (4)	7.45	7.30	8.69

B. Particular neighborhoods of interest

	Prices (NIS)			Price rank (lowest=1, highest=15)			
	Sep-07	Nov-07	Nov-08	Sep-07	Nov-07	Nov-08	Mean rank
NAP1	7.15	7.31	8.01	6	10	10	9
NAP2	7.54	7.39	8.14	10	14	11	12
NAP3	7.80	7.36	8.19	14	12	13	13
AC1	8.01	7.27	8.52	15	8	14	12
AC2	7.55	7.61	7.85	11	15	8	11
AC3	7.68	7.06	7.76	13	7	7	9
CD1	6.33	6.15	6.89	2	1	1	1
CD2	7.45	7.30	7.07	9	9	2	7

C. Gains from travel: distribution of savings (%) from shopping at the cheapest location

Percentile	10%	25%	50%	75%	90%
Savings	2.58	9.66	14.01	16.39	19.13

Notes: Panel A displays composite good price statistics over the 15 neighborhoods where the price could be computed using at least 21 observed products (see text). Panel B presents values for particular neighborhoods where CD, NAP and AC correspond to Commercial Districts, Non-Affluent Peripheral, and Affluent Central neighborhoods (see text). The last column in panel B presents the neighborhood's mean price rank over the three sample periods. For example, the CD1 commercial district has a mean rank of 1, i.e., it is, on average, the cheapest location. Panel C shows the distribution of savings, in percentage terms, from shopping at the cheapest location rather than at the home neighborhood using data for all 15 neighborhoods where the price could be computed using at least 21 observed products over the sample period. For example, the median savings are 14.01%.

Table 5: Credit card expenditure flows

A. Statistics over all 46 neighborhoods

	Fraction spent at Own neighborhood	CD1	CD2
Mean	0.22	0.27	0.06
Median	0.16	0.19	0.03
Min	0.00	0.01	0.01
Max	0.76	0.76	0.41

B. Neighborhoods of interest

	Fraction spent at Own neighborhood	CD1	CD2
NAP1	0.25	0.03	0.02
NAP2	0.42	0.18	0.04
NAP3	0.33	0.31	0.05
AC1	0.44	0.19	0.03
AC2	0.14	0.16	0.18
AC3	0.00	0.65	0.02

Notes: The table provides statistics (averaged over the sample period) on the fraction of expenditures spent at the own neighborhood and at the CD1 and CD2 commercial districts. NAP and AC correspond to Non-Affluent Peripheral, and Affluent Central neighborhoods (see text).

Table 6: Estimates of utility function parameters

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln (price at destination)	8.768	9.283	1.691	1.725	5.065	4.727	1.630	4.730
	(5.788)	(5.491)	(.763)	(.749)	(1.421)	(1.304)	(.774)	(1.302)
ln (price at destination) X housing prices					-0.253	-0.232		-0.232
					(.083)	(.078)		(.078)
Distance to destination	0.272	0.365	0.197	0.334	0.393	0.423	0.423	0.411
	(.049)	(.072)	(.036)	(.045)	(.13)	(.12)	(.12)	(.119)
Distance to destination Xsenior citizen					0.002	0.004	0.004	0.004
					(.004)	(.007)	(.007)	(.006)
Distance to destination Xdriving to work					-0.002	-0.003	-0.003	
					(.002)	(.002)	(.002)	
Distance to destination X car ownership								-0.002
								(.001)
Shopping at home	2.489	1.723	3.035	2.089	1.977	1.890	1.889	1.910
	(.426)	(.526)	(.397)	(.41)	(.435)	(.426)	(.426)	(.424)
Fixed origin effects	NO	YES	NO	YES	YES	YES	YES	YES
Fixed destination effects	NO	NO	YES	YES	YES	YES	YES	YES
Fixed period effects	YES	YES	YES	YES	YES	YES	YES	YES
Fixed destination effects X housing prices	NO	NO	NO	NO	NO	YES	YES	YES
# observations	1819	1819	1819	1819	1819	1819	1819	1819
R^2	0.243	0.382	0.657	0.775	0.776	0.784	0.783	0.783

Notes: The table reports estimates of equation (3). The price and distance variables were entered with a negative sign in the regression so that the estimates in the table are estimates of α and β. The second row reports the coefficient for the interaction of the natural log of prices at the destination with housing prices at the origin. The distance variable is interacted with the origin neighborhood's share of senior citizens, its share of residents who drive to work, and its car ownership share. The term "Destination$\times$ housing p." refers to the interaction of destination neighborhood fixed effects with housing prices at the origin neighborhood. Standard errors in parentheses are 2-way clustered at the origin and destination levels.

Table 7: Distribution of estimated elasticities across neighborhoods (absolute value)

A. Own price elasticity

mean	sd	min	p10	p25	p50	p75	p90	max
				$\sigma = 0.7$				
4.82	0.92	3.00	3.86	3.99	4.95	5.87	5.95	6.13
				$\sigma = 0.8$				
6.43	1.37	3.78	5.01	5.31	6.54	7.94	8.32	8.47

B. Distance semi-elasticity

mean	sd	min	p10	p25	p50	p75	p90	max
0.35	0.06	0.06	0.28	0.31	0.35	0.39	0.42	0.45

Notes: All elasticities computed given the baseline demand estimates (column 6 of Table 6) for November 2008. Own price elasticities are presented for alternative values of σ, while distance semi-elasticities are at the neighborhood level and do not depend on σ. The table shows statistics for elasticities computed at each of the 15 destinations where the composite good price could be computed using at least 21 items. Price elasticities (panel A) were computed over for each of these 15 destinations and so the reported statistics were computed over 15 values. Distance semi-elasticities (Panel B) were computed for each of the 690 (46x15) origin-destination pairs, and so the reported statistics were computed over 690 values.

Table 8: Estimated costs and margins

		σ=0.7		σ=0.8	
	p	c	(p-c)/p	c	(p-c)/p
Average (all)	7.80	6.11	0.22	6.52	0.16
Median (all)	7.85	6.07	0.20	6.44	0.15
Median CD1-CD2	6.98	5.69	0.18	6.04	0.13
Median residential	7.87	6.10	0.21	6.44	0.16
Median NAP1-NAP3	8.14	6.66	0.17	7.00	0.13
Median AC1-AC3	7.85	5.88	0.25	6.37	0.19

Notes: The table reports the composite good price (p), marginal cost (c), and price-cost margin (p-c)/p in each destination neighborhood where the composite good price could be computed using at least 21 items in November 2008. Costs and margins are reported under two alternative values for the correlation parameter σ. CD, NAP and AC correspond to Commercial Districts, Non-Affluent Peripheral, and Affluent Central neighborhoods (see text).

Table 9: Counterfactual changes to posted prices

Retail location	Observed posted price	Reduced travel disutility		Improved amenities		Additional entry
	(NIS)	Distance	Distance & κ	CD1	CD1-CD2	
Average (all)	7.80	-0.9%	-1.4%	-0.3%	-0.3%	-2.6%
Median (all)	7.85	-0.7%	-1.0%	0.0%	-0.3%	-3.0%
Median CD1-CD2	6.98	-0.7%	-0.6%	0.4%	0.3%	0.0%
Median residential	7.87	-0.5%	-0.5%	-0.1%	-0.8%	-3.4%
NAP1	8.01	3.3%	4.7%	-0.1%	-0.3%	-2.8%
NAP2	8.14	0.4%	0.8%	0.0%	-0.1%	-1.3%
NAP3	8.19	-0.5%	-0.5%	-0.6%	-0.9%	-3.5%
AC1	8.52	-8.2%	-12.0%	-3.6%	-1.1%	-6.8%
AC2	7.85	-1.3%	-3.7%	0.2%	-0.8%	-1.9%
AC3	7.76	-0.2%	-0.3%	-0.1%	-0.2%	-3.0%

Notes: The table reports statistics on the percentage changes to the posted prices charged at locations where prices are observed (11 residential neighborhoods and 4 commercial districts) under the various counterfactual exercises, computed in the third time period (November 2008). CD, NAP and AC correspond to Commercial Districts, Non-Affluent Peripheral, and Affluent Central neighborhoods (see text). The first and second row report average and median values computed over all 15 neighborhoods with valid prices, respectively. The third and fourth rows report median values computed over the two main commercial districts CD1 and CD2, and over all residential neighborhoods, respectively. The first column pertains to the posted prices (NIS) observed in the data. The second column reports price changes under the counterfactual that reduces the distance disutility parameter β by 50%, and the third column pertains to reducing, in addition, the "shopping at home" utility parameter κ by 50%. The fourth (fifth) column pertains to improving the amenities at the commercial districts CD1 (CD1 and CD2), while the last column refers to the counterfactual that adds a supermarket to each residential neighborhood.

Table 10: Counterfactual changes to the Average Price Paid (APP)

Retail location	Observed APP	Reduced travel disutility		Improved amenities		Additional entry
Median residential	7.72	-1.8%	-3.2%	-4.7%	-5.6%	-1.6%
NAP1	7.86	0.4%	0.0%	-2.2%	-3.2%	-2.6%
NAP2	7.85	-3.5%	-5.5%	-6.6%	-7.2%	-0.7%
NAP3	7.72	-3.3%	-4.7%	-7.0%	-7.3%	-1.2%
AC1	7.98	-5.7%	-7.3%	-8.6%	-8.7%	-3.2%
AC2	7.67	-2.9%	-3.4%	-4.7%	-6.2%	-0.5%
AC3	7.28	-1.1%	-1.2%	-4.1%	-4.3%	-0.3%

Notes: The table reports the percentage changes in the Average Price Paid (i.e., in the weighted average of prices that takes into account the probabilities with which residents shop in different destinations) faced by residents of the 11 residential neighborhoods where prices are observed. All computations performed for the third time period (November 2008). The first column shows the APP (NIS) corresponding to the observed equilibrium, whereas additional columns show the counterfactual change to the APP given alternative parameter values: the second column pertains to the counterfactual that reduces the distance disutility parameter β by 50%, and the third column pertains to reducing, in addition, the "shopping at home" utility parameter κ by 50%. The fourth (fifth) column pertains to improving the amenities at the commercial districts CD1 (CD1 and CD2), while the last column refers to the counterfactual that adds a supermarket to each residential neighborhood. CD, NAP and AC correspond to Commercial Districts, Non-Affluent Peripheral, and Affluent Central neighborhoods (see text).

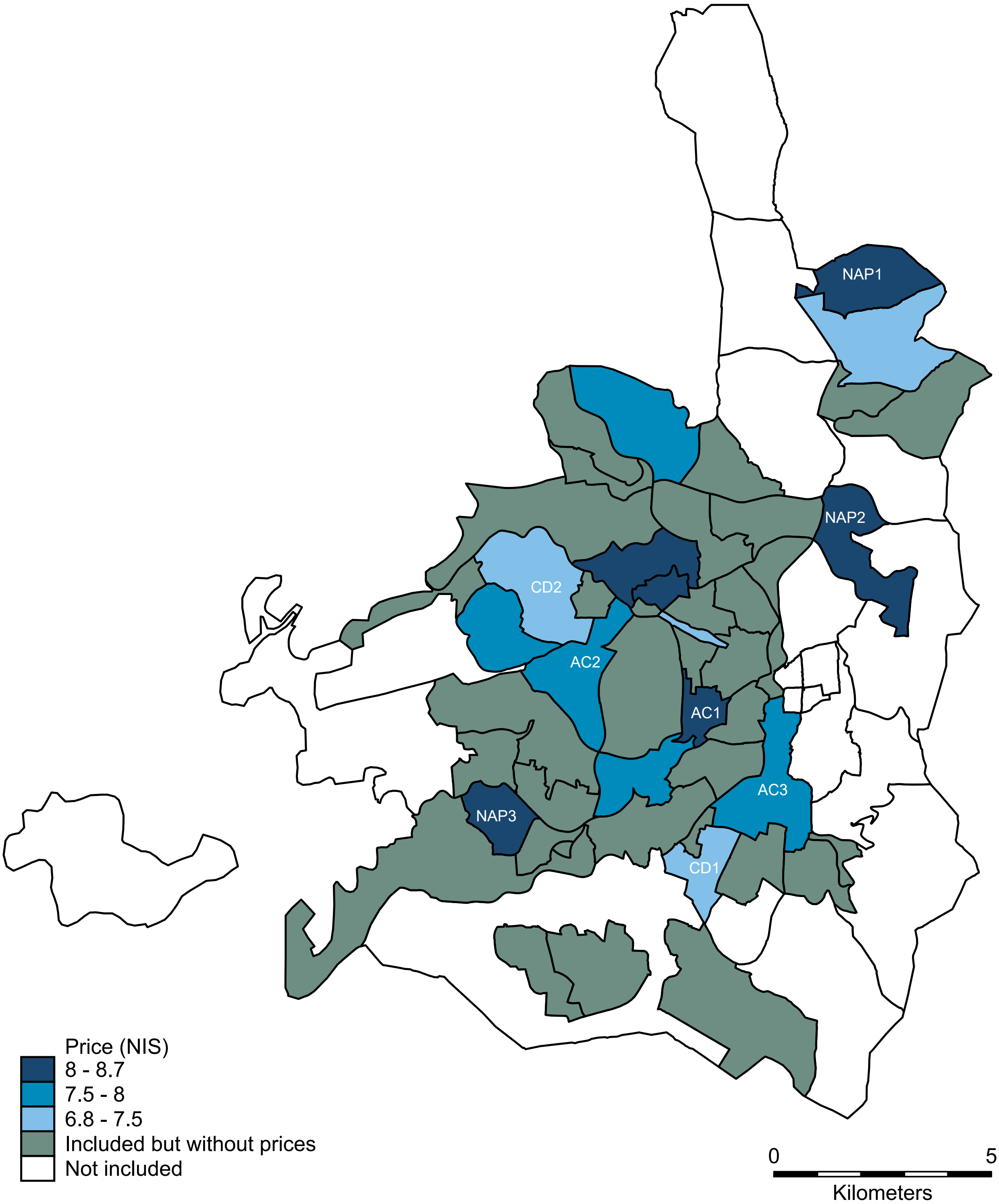

Figure 1: Neighborhoods included in the analysis and price levels

Figure 2: Composite good prices plotted against housing prices, November 2008

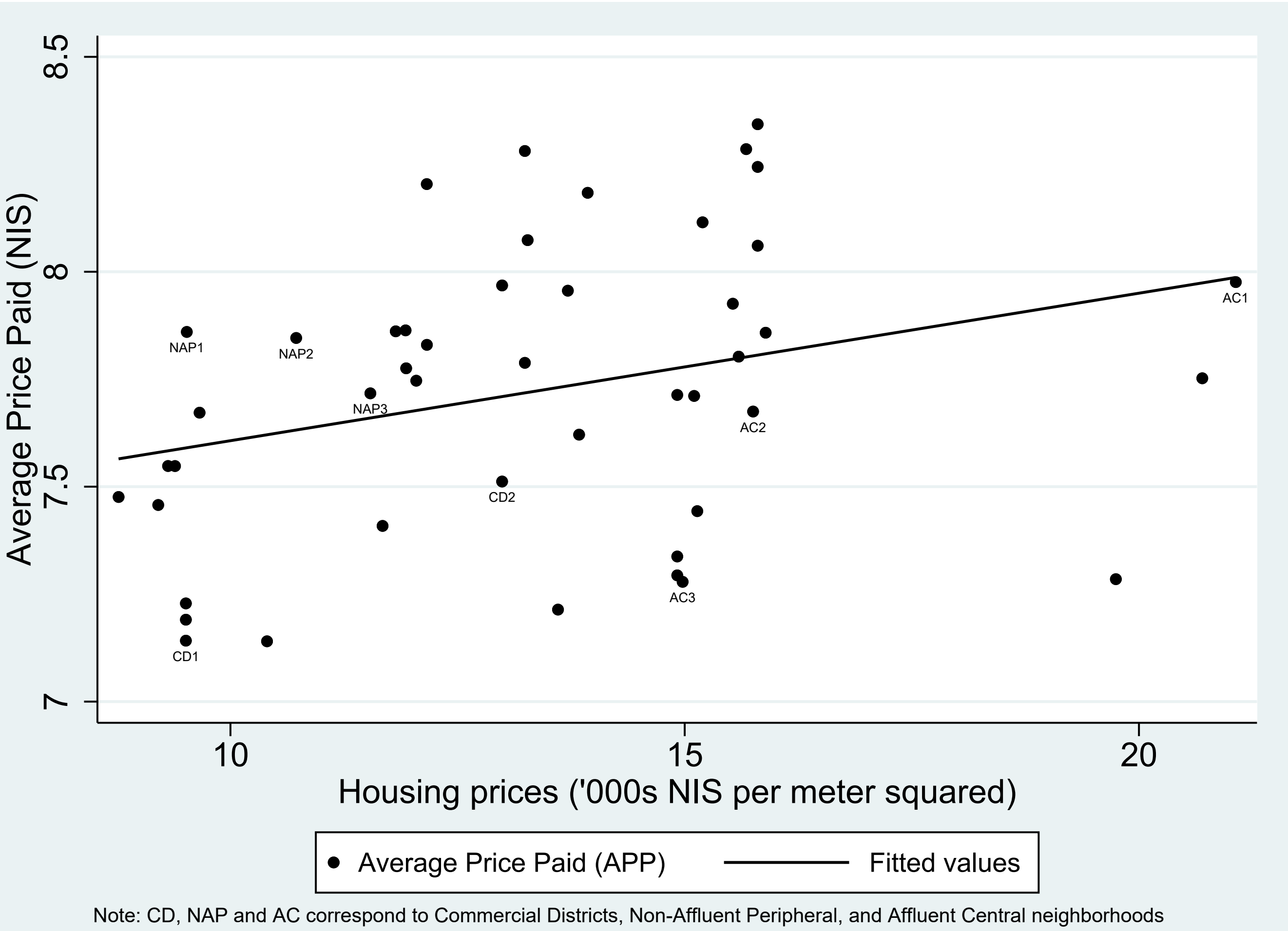

Figure 3: Average Prices Paid (APP) plotted against housing prices, November 2008

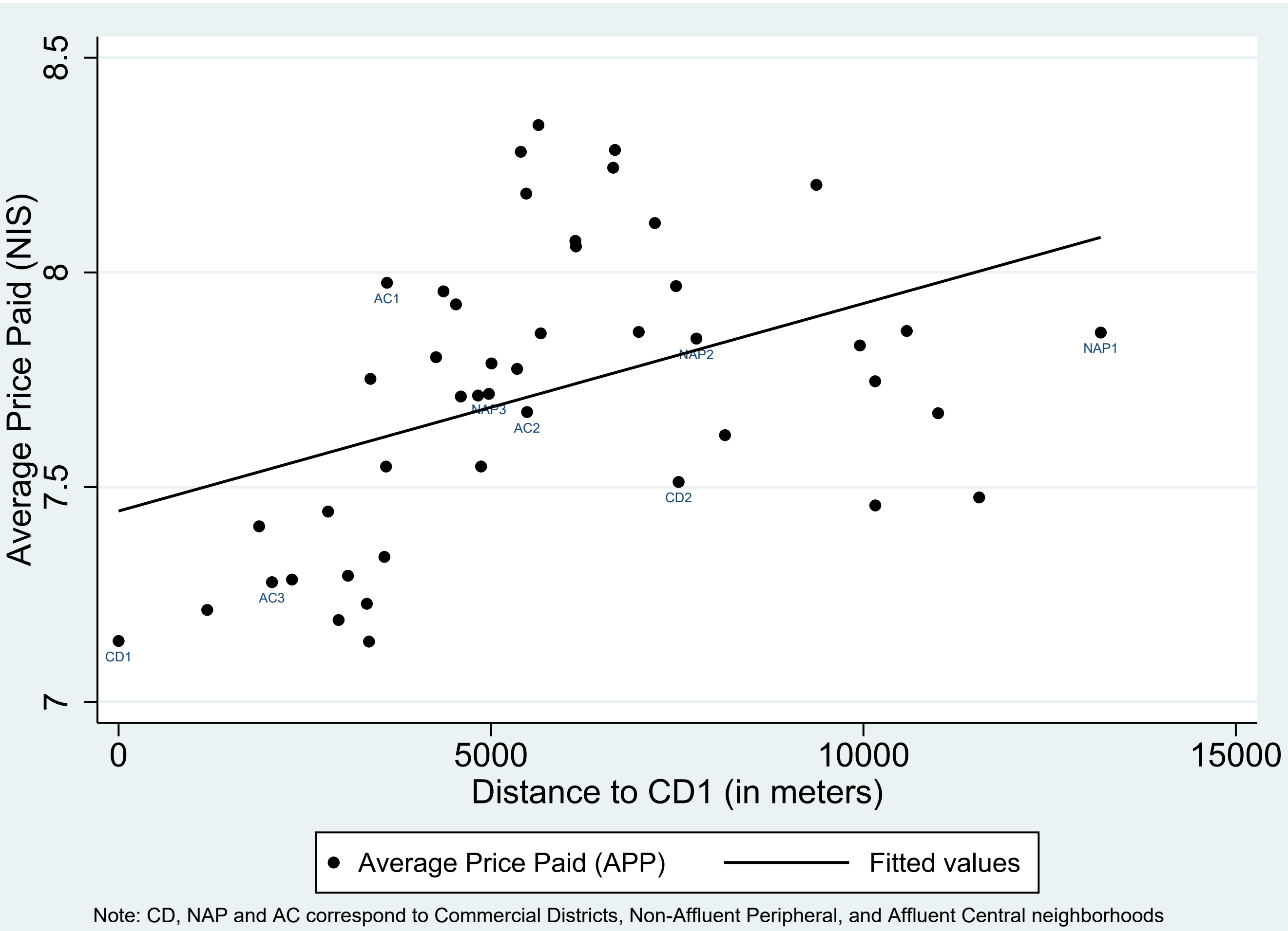

Figure 4: Average Prices Paid (APP) plotted against distance to CD1, November 2008

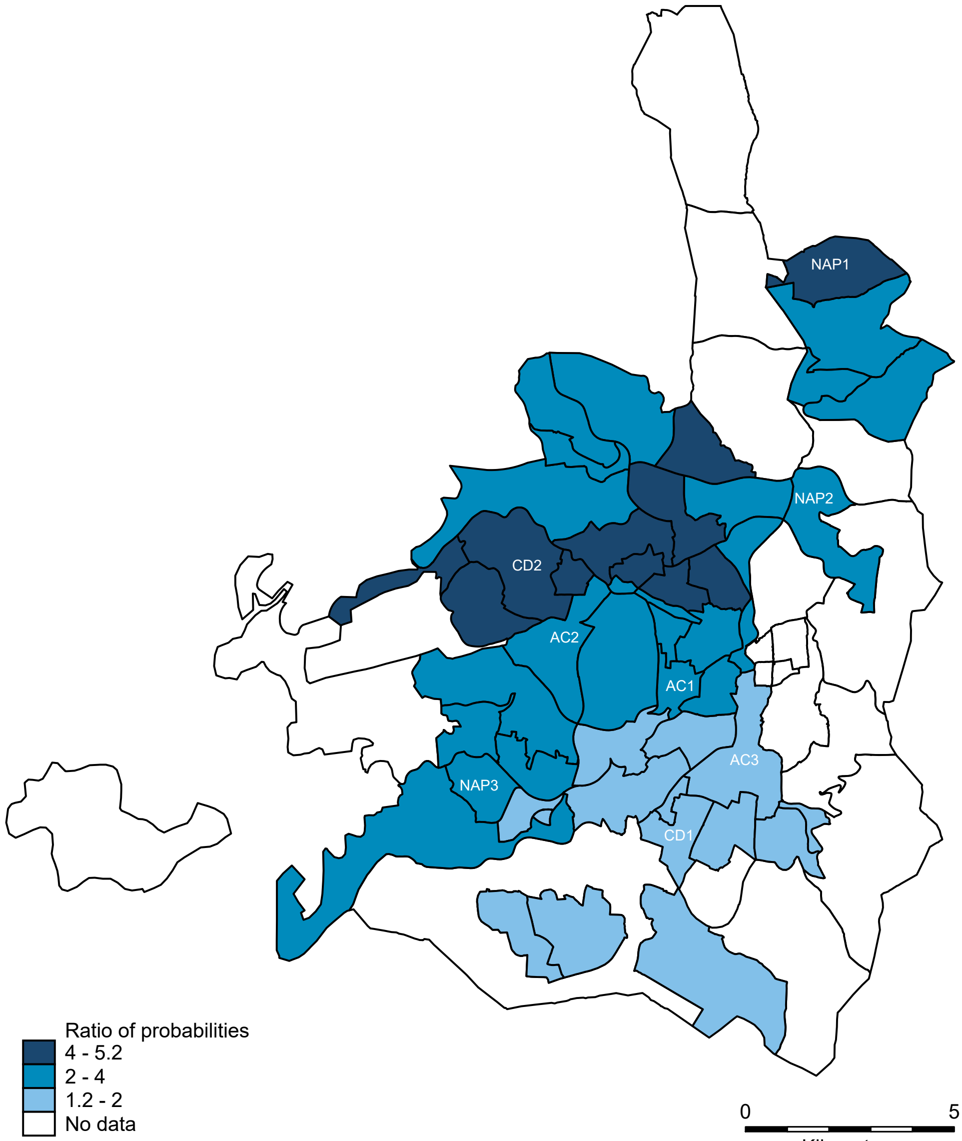

Figure 5: Ratio of counterfactual (improved amenities at CD1) to observed probability of shopping at the main commercial district CD1, November 2008