Retail Prices in a City*

Alon Eizenberg
The Hebrew University and CEPR

Saul Lach
The Hebrew University and CEPR

Merav Yiftach
Israel Central Bureau of Statistics

February 2017

Abstract

We study grocery price differentials across neighborhoods in a large metropolitan area (the city of Jerusalem, Israel). Prices in commercial areas are persistently lower than in residential neighborhoods. We also observe substantial price variation within residential neighborhoods: retailers that operate in peripheral, non-affluent neighborhoods charge some of the highest prices in the city. Using CPI data on prices and neighborhood-level credit card data on expenditure patterns, we estimate a model in which households choose where to shop and how many units of a composite good to purchase. The data and the estimates are consistent with very strong spatial segmentation. Combined with a pricing equation, the demand estimates are used to simulate interventions aimed at reducing the cost of grocery shopping. We calculate the impact on the prices charged in each neighborhood and on the expected price paid by its residents - a weighted average of the prices paid at each destination, with the weights being the probabilities of shopping at each destination. Focusing on prices alone provides an incomplete picture and may even be misleading. Specifically, we find that interventions that make the commercial areas more attractive and accessible yield only minor price reductions, yet expected prices decrease in a pronounced fashion. The benefits are particularly strong for residents of the peripheral, non-affluent neighborhoods.

*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 Ruthie Harari-Kremer for her help with maps, and to Elka Gotfryd for mapping zipcodes into statistical subquarters. We thank Steve Berry, Pierre Dubois, Phil Haile, JF Houde, Volker Nocke, Mark Rysman, Katja Seim, Avi Simhon, Yuya Takashi, Ali Yurukoglu and Christine Zulehner for helpful comments, as well as seminar participants at Carlos III, CEMFI, Frankfurt, Harvard, Johns Hopkins, Yale and Wharton, and participants at the Israeli IO day (Tel Aviv, 2014), EARIE (Milan 2014) and the Economic Workshop at IDC (2015).

Correspondence: alon.eizenberg@mail.huji.ac.il (Eizenberg), Saul.Lach@mail.huji.ac.il (Lach), meravo@cbs.gov.il (Yiftach).

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

Applied economists have long been interested in price variation across retail locations. Much of this work documented variation in prices across neighborhoods within a city. Urban scholars, starting with Caplovitz (1963), have focused on relating the observed price differentials to variation in socioeconomic and demographic factors ("do the poor pay more?")[1] In this paper, we explore the variation in the cost of grocery shopping across neighborhoods in the city of Jerusalem, Israel. Our goal is not to determine whether the poor pay more, but rather to explore the determinants of the cost of grocery shopping. In our analysis, this cost is determined as an equilibrium outcome of a structural model of demand and supply. In equilibrium, the cost incurred by residents of a given neighborhood is affected in a nontrivial fashion by the neighborhood's socioeconomic standing, its spatial location relative to the city's large commercial centers, and the degree of intra-neighborhood retail competition.

While shopping at the neighborhood of residence is prevalent, it is by no means exclusive. As we report below, on average, only 22% of expenditures are spent in the home neighborhood. Thus, observed price variation across neighborhoods is not sufficient for inferring the variation in the cost of grocery shopping across neighborhoods. Motivated by this observation, our approach will analyze both prices *and* shopping patterns across neighborhoods.[2] To this end, we compute the *expected price* paid by a random resident of the neighborhood. This expected price is a weighted average of the prices charged at each retail destination in the city, with the weights being the probabilities with which residents of the relevant neighborhood shop at these various destinations. It therefore combines information on prices and on shopping patterns. We study the variation across neighborhoods in both the prices charged by retailers operating in the neighborhood, and in the expected price paid by its residents. We also explore the manner by which these prices are affected by policy interventions.

Jerusalem is composed of very distinct residential neighborhoods, and also has several popular commercial areas. Hard discount supermarkets are located in the commercial areas, whereas residential neighborhoods feature more expensive supermarkets. Our first step is to characterize the prices charged by retailers in each of these (residential and commercial) neighborhoods using

[1] Price differentials, especially when products are homogeneous, hint at violations of the "law of one price" and have therefore also been of interest to Industrial Organization researchers. See Baye et al. (2006) for a review of theoretical models rationalizing "price dispersion" in equilibrium and the empirical work documenting its existence and characteristics in various markets.

[2] See Frankel and Gould (2001) for the general point that neighborhood of residence and location of shopping need not be perfectly correlated. This point has also been recently emphasized by Houde (2012). See, among others, Aguiar and Hurst (2007), Griffith et al. (2009), Kurtzon and McClelland (2010) for analyses of survey data where recorded prices correspond to prices actually paid by households.

price data from the Israeli Central Bureau of Statistics (ICBS). These data cover 27 everyday grocery items sold at about 60 retailers in Jerusalem (in 2007 and 2008). We aggregate these individual-item prices into a neighborhood-level "composite good" price. This price exhibits substantial variation across neighborhoods, holding products' qualities fixed. Prices in residential areas are, in general, higher than in commercial areas. For example, in November 2008, the price of the composite good in an affluent residential neighborhood (Rehavya) was 24 percent higher than the price in a popular commercial area located 3.6 km away (Talpiot). The average difference between the prices charged in residential and commercial areas is about 8 percent.

Examining variation in the prices charged by retailers across *residential neighborhoods* also reveals some interesting patterns. Very high prices are charged not only in the centrally-located, affluent neighborhood of Rehavya, but also in three of the least affluent neighborhoods: Neve Yaaqov, Givat Shapira and Qiryat HaYovel. The common feature of these three neighborhoods is their peripheral location, at some distance from the city's center and from the main commercial areas. In fact, retailers in those neighborhoods charge higher prices than retailers in more affluent residential neighborhoods that are located closer to the main commercial areas. This suggests that cross-neighborhood shopping plays an important role in determining equilibrium prices. Simply put, the intensity of competition from the commercial areas' hard discount chains affects the pricing decisions of retailers located in residential neighborhoods, and this intensity is higher, the closer is the residential neighborhood to the commercial center.

This mechanism is nicely illustrated by anecdotal evidence. Residents from Qiryat HaYovel, one of the three disadvantaged neighborhoods mentioned above, initiated a consumer boycott in January 2014 against a supermarket located in their neighborhood. They claimed that prices in this supermarket were much higher than those charged in other branches of the same chain that operate in the city's commercial areas. The boycott organizers cited travel cost as the main impediment to their shopping in the commercial areas: "Young families will not travel to Talpiot or Givat Shaul (the two main commercial areas) to shop and, instead, shop in the neighborhood for lack of time."[3] The boycott organizers arranged transportation services and encouraged residents to shop outside the neighborhood. The boycott ended after the chain agreed to lower the cost of a basket of goods by 14%, according to the organizers. This figure approaches the price differentials between Qiryat HaYovel and the commercial areas measured in our sample period, which pre-dates the boycott.

To document shopping patterns, we use data on grocery expenditures from a credit card company. These are neighborhood-level aggregate data that report expenditures by residents of each "origin" neighborhood (identified by the buyer's zipcode) spent in each "destination"

[3] "Ynet" (an Israeli news outlet), January 13th 2014.

neighborhood (identified by the seller's zipcode). To the best of our knowledge, this is a new source of data on shopping patterns. The expenditure data reveal considerable variation in the fraction of expenditures spent within the home neighborhood. Residents of the affluent Rehavya neighborhood made 44 percent of their grocery spending "at home", while those in the Geulim neighborhood did not shop at home at all.[4] The most popular commercial area is Talpiot where households made, on average, 27 percent of their grocery purchases. Here also there is variation across residential neighborhoods. Residents of the Geulim neighborhood which borders with the Talpiot commercial area performed 65 percent of their purchases there, while residents of Rehavya, located 3.6 km away, performed only 19 percent of their shopping at Talpiot.

The next step in our analysis is the formulation of a structural model of demand, following the literature on the estimation of differentiated-product demand systems using aggregate data (Berry 1994, Berry, Levinsohn and Pakes 1995, Nevo 2001). In the model, households make the discrete choice of where to shop for the composite good by maximizing preferences that depend on price, distance and unobserved characteristics of the shopping experience. We allow demographics to affect price and distance sensitivities. A nested logit structure allows us to consider retailers located within a neighborhood as closer substitutes than retailers located in different neighborhoods. We follow Björnerstedt and Verboven's (2016) adaptation of the discrete choice framework to allow consumers to also choose the quantity of purchased units.

We assume that consumers are perfectly informed regarding all shopping locations and the prices and amenities offered there. This stands in contrast to a familiar "search cost" literature in which price differentials are explained as a consequence of consumers being imperfectly informed about prices (Stigler, 1961). In Jerusalem, prices in residential neighborhoods are *persistently* higher than those in the commercial areas. The exact location of the low price stores is common knowledge. This is likely to be true in many urban settings, and we thus choose to ignore potential information frictions and emphasize spatial frictions instead.[5]

The demand model is helpful in three different ways. First, the model clarifies the conditions under which observed credit-card expenditure shares can be used to measure the probabilities with which residents of each origin neighborhood choose to shop at each destination neighborhood. This is not trivial due to two reasons: the measurement error brought about by, among other issues, observing credit card expenditures rather than total expenditures, and the fact that aggregate neighborhood-level expenditures mask individual-level heterogeneity in the quantity

[4] The Geulim subquarter includes three affluent areas: Geulim (Baqa), Givat Hananya (Abu Tor), and Yemin Moshe.

[5] But see Dubois and Perrone (2015) for a different view. Other empirical studies based on the imperfect information paradigm are, for example, Sorensen (2000), Lach (2002), Brown and Goolsbee (2002), and Chandra and Tapatta (2011).

of purchased groceries. We then use the estimated probabilities to compute the *expected price* paid by a random resident of each neighborhood. This expected price is typically lower than the price charged by the retailers operating in the neighborhood, since the neighborhood's residents take advantage of the opportunity to shop at cheaper locations. Nonetheless, the expected prices paint the same picture regarding the three peripheral, non-affluent neighborhoods discussed above: residents of these neighborhoods face some of the highest expected prices in the city, in addition to being charged very high prices at their local neighborhood's supermarkets.

Second, the estimated demand model delivers reasonable price and distance elasticities and sheds light on the role played by spatial frictions in household preferences. Our model departs from standard applications by deriving the econometric error term from non-random measurement error in the expenditure data. We show how to use the panel structure of the data to obtain consistent estimation. Third, combined with a pricing equation, the estimated demand model allows us to back out retailers' marginal costs, and to compute counterfactual price equilibria under various policy interventions.

We consider three types of interventions that aim at reducing the cost of grocery shopping. First, we reduce the disutility from travel, with the interpretation of improvements in the city's transportation infrastructure. A second intervention improves the unobserved amenities of shopping at the major commercial areas which we interpret as providing better parking and general organization of the commercial areas. Finally, the third intervention increases within-neighborhood competition via the entry of additional retailers into residential neighborhoods.[6]

In the first two interventions (reduced disutility from travel, and improved amenities at the shopping areas), equilibrium prices are only mildly reduced.[7] In contrast, the expected price decreases considerably in those interventions. The benefits to the peripheral, less-affluent neighborhoods are particularly pronounced. For instance, when amenities at the major shopping area of Talpiot are improved, the expected price paid by residents of Qiryat HaYovel drops by 7%, while the price charged by retailers in Qiryat HaYovel itself is reduced only by 0.6%. The expected price falls by much more than the price charged in the neighborhood since residents shop much more intensely at the lower price supermarkets of Talpiot: specifically, the probability that Qiryat HaYovel's residents shop at Talpiot rises from 0.28 in the observed equilibrium to 0.76 under this intervention. Note that considering only the effect on equilibrium prices would miss

[6]Given tractability considerations, we treat the entry decisions of supermarkets as fixed. The IO literature has developed tools for studying endogenous retail entry and location choices (see Seim 2006, Beresteanu, Ellickson and Misra 2010, Aguirregabiria, Mira, Roman 2007, and Ellickson, Houghton and Timmins 2013). We view this restriction as reasonable given the stability of supermarket locations over long periods of time stemming from strict zoning restrictions and space constraints.

[7]In certain scenarios, prices in some neighborhoods are even slightly increased. We provide an explanation for this counter-inutitive result in Section 4.2.

the substantial benefits implied by this policy intervention. We therefore stress the importance of the joint analysis of prices and shopping patterns as summarized by the expected prices.

Another insight is provided by comparing the three interventions. The greatest reduction in expected prices is brought about by the second scenario, in which amenities at the commercial areas are improved. On average across neighborhoods, expected prices drop by 5.8% (noting that the prices actually charged, averaged across retail locations, drop by less than 0.5%, again emphasizing the importance of accounting for shopping patterns). Moreover, as we discuss in Section 4, this second intervention is also associated with lower social costs than improving the transportation infrastructure, or facilitating additional supermarket entry into residential neighborhoods. Thus, our findings suggest that the cost of grocery shopping can be reduced by making shopping at the commercial centers more attractive.[8] Notably, the benefits to residents of peripheral, non-affluent neighborhoods are particularly strong.

Following a literature review, the paper proceeds as follows. In Section 2, we present our price and expenditure data. Section 3 presents the model of consumer demand and its estimation. Section 4 describes our pricing model and its implied margins, as well as counterfactual experiments. Section 5 offers concluding remarks.

Literature. A vast urban economics literature compares prices across residential locations. 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 systematic price variation across store types, neighborhoods and cities. Prices in suburban locations were about 4 percent lower than in central city stores where poorer population lived. Chung and Myers (1999) analyze survey data for the Twin Cities metropolitan area, and also find that the price of a weekly home food plan was higher in poorer neighborhoods. Recent work challenges these findings and reports that prices in richer zip codes (Hayes, 2000) or prices paid by high income households (Aguiar and Hurst, 2007) are significantly higher. Kurtzon and McClelland (2010) study a BLS telephone 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." Frankel and Gould (2001) document price differences *across cities* and find that higher prices are associated with the absence of lower middle-class consumers. These authors use city-level price variation, citing the difficulty of conducting neighborhood-level analysis stemming from cross-neighborhood shopping. Our paper differs from the above literature in that our focus is not on whether "the poor pay more" per se, and in that our structural approach allows us to evaluate counterfactual policies.

[8] The city of Jerusalem in fact plans to improve both access to the main shopping area of Talpiot, via the extension of the light rail system, and its internal organization ("The plan: the Talpiot industrial zone to undergo a revolution in the next decade," Kol Hair, a local Jerusalem newspaper, April 2016).

The literature on spatial frictions in economics is vast with classic theoretical contributions including Hotelling (1929) and Salop (1979). Smith and Hay (2005) offer a theoretical model to study competition across shopping centers, focusing on agglomeration effects stemming from consumers' preference for one-stop shopping (See also Dluhosch and Burda 2007). Several recent empirical papers have taken a structural approach to study spatial competition in various industries, including Adams and Williams (2014), Miller and Osborne (2012), Thomadsen (2005), Davis (2006), McManus (2007), and Houde (2012), whose demand model considers the "home" neighborhood for gasoline consumers as their entire commuting path between home and work. Davis, Dingel, Monras and Morales (2015) examine the role of spatial frictions in determining restaurant choices in New York using data from Yelp.com. They find that travel time is a first order determinant of restaurant choice. Interest in shopping patterns is not limited to the choice of location. Griffith, Leibtag, Leicester, and Nevo (2009) examine how purchasing on sale, buying in bulk (at a lower per unit price), buying generic brands and choosing outlets impacts household grocery expenditures.

Finally, a substantial body of empirical work considers spatial competition among supermarkets. For example, Chintagunta, Dubé, and Singh (2003) study pricing policies by multi-store supermarket chains. Smith (2004) estimates a discrete-continuous model of consumer demand in which both the choice of the retailer and total expenditures are endogenously determined. Dubois and Jódar-Rosell (2010) study price and brand competition across supermarkets. They estimate a discrete-continuous demand model and use a supply-side model to identify heterogeneous marginal costs. They consider a counterfactual analysis in which travel costs are reduced and explore the impact on retailers' prices and brand offerings. Figurelli (2013) estimates transportation costs within a model in which consumers choose where to shop, employing a control function approach to address the endogenous choice of the bundle of goods purchased at the store. Our paper addresses a different economic question relative to this extant literature and, as a consequence, our framework differs from that employed in those papers.

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 across the city, and the data on consumer expenditures.

2.1 Jerusalem's urban structure: neighborhoods

Jerusalem's urban structure provides a convenient arena for the study of price differentials across neighborhoods. The city's population resides in clearly distinct neighborhoods that differ substantially in their socioeconomic makeup and are, for the most part, spatially-separated.[9] Neighborhoods are geographically spread out and moving between them typically requires some mode of transportation. While distinct neighborhoods with established identities are a key feature of Jerusalem, there is no formal statistical definition that precisely matches the notion of a "neighborhood." We therefore use the Israel Central Bureau of Statistics's (ICBS) closely-related concept of a *subquarter,* and use the terms *neighborhood* and *subquarter* interchangeably. A subquarter includes several *statistical areas* with territorial continuity between them.[10] Our analysis covers 46 neighborhoods: 40 residential subquarters and 6 "commercial areas."

The two major commercial areas are Talpiot and Givat Shaul. Additional commercial areas are Romema, the Central Bus Station, the market at Mahane Yehuda, and the large Malcha shopping mall (see Table A1 in Appendix A for additional details). We defined these six commercial areas as collections of statistical areas that are predominantly commercial with minimal residential presence. These areas were typically carved out of a larger subquarter. For instance, the original Talpiot subquarter was partitioned into two parts: a collection of primarily-residential statistical areas, and a collection of primarily-commercial statistical areas. Figure 1 displays the city's neighborhoods, highlighting the 46 neighborhoods covered by our study.

The 46 neighborhoods are in the western part of the city and include all major predominantly Jewish neighborhoods, but exclude the historical "old city" and the predominantly Arab neighborhoods in the eastern part of Jerusalem. Several considerations motivate this exclusion. First, despite a strong integration of many economic activities across the Arabic and Jewish communities in the city of Jerusalem, these populations tend to reside in distinct neighborhoods that are near-exclusive Arabic or Jewish. Second, residents of the two communities consume quite distinct baskets of goods. For example, cottage cheese is an everyday staple among the Jewish population but it is not consumed by the Arab population. Third, while residents of the western neighborhoods do perform some shopping in eastern neighborhoods, and vice-versa, this is not the norm when it comes to the weekly grocery shopping trip. Lastly, our credit card expenditures data will be less representative of expenditures by Arab households because of their low usage of credit cards.[11]

---

[9] The population of Jerusalem in 2008 was 763,600 (495,000 Jews and 268,600 Arabs) and its area was 126 km^2 (http://www.jiis.org/.upload/publications/facts-2008-eng.pdf).

[10] A statistical area is a small geographic unit as homogeneous as possible, generally including 3,000 – 4,000 persons in residential areas. http://www.cbs.gov.il/mifkad/mifkad_2008/hagdarot_e.pdf.

[11] According to the ICBS, in 2013 the percentage of Arab households in Israel hav-

Figure 1: Neighborhoods included in the study

The neighborhoods are matched to demographic information from the 2008 Israel Census of Population. We focus on demographics that are likely to shift price and travel sensitivities: "car ownership" (percentage of the subquarter's households having at least one car), "driving to work" (percentage of those aged 15 and over who used a private car or a commercial vehicle as a driver to get to work), and "senior citizen" (percentage above the age of 65). We also observe housing prices, obtained from the country's Tax Authority's records of real estate transactions. These prices are a proxy for the subquarter's wealth and are measured by the 2007-2008 average price per square meter.[12] Table 1 shows the distribution of these variables across neighborhoods.

Table A2 in Appendix A shows the neighborhood-specific values of these variables, and reveals

ing a credit card was 53 percent, while that of Jewish households was 88 percent (http://www.cbs.gov.il/reader/newhodaot/hodaa_template.html?hodaa=201515045). Credit cards may also be less accepted by retailers in the Arab neighborhoods.

[12]We thank Daniel Felsenstein for providing the housing price data.

Table 1: Distribution of demographics across 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. Driving to work = percentage of those aged 15 and over 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 above age 65.

considerable variation across neighborhoods. The population fraction owning at least one car, for instance, is only 7 percent in Mea Shearim, an ultra-orthodox neighborhood, but reaches 89 percent in Har Homa, a new neighborhood located in the outskirts of Jerusalem. Similarly, housing is relatively cheap in Pisgat Zeev North, while being 2.5 times as expensive in the affluent neighborhood of Rehavya. This variation will help identify heterogeneity in price and distance sensitivities across neighborhoods.

Distance across neighborhoods plays an important role in our analysis. The ICBS prepared a matrix of the shortest road distance between the centroids of each pair of statistical areas. We used this information to generate a matrix of distances between each pair of the 46 neighborhoods. The distance d_{jn} between neighborhoods j and n is an average of the distances between each pair of statistical areas that belong in neighborhoods j and n, for $1 \leq j, n \leq 46$. Certain neighborhoods are themselves quite large, and so we define neighborhood j's "own distance" d_{jj} as the mean distance between the centroids of each pair of the statistical areas included in it. Table 2 shows the distance between each neighborhood and the City center, the two main shopping areas (Talpiot and Givat Shaul) and the average distance to all the other neighborhoods. The latter provides a rough idea of how "isolated" each neighborhood is. Neve Yaaqov, one of the three peripheral neighborhoods mentioned in the introduction, is the most isolated neighborhood in this sense.

2.2 Price data

Price data for 27 products were collected during September and November 2007, and November 2008, in 60 distinct stores across Jerusalem. About 55 percent of the stores were supermarkets,

Table 2: Distance across neighborhoods (Kilometers)

Subquarter	City center	Talpiot	Givat Shaul	All (mean)	Subquarter	City center	Talpiot	Givat Shaul	All (mean)
Neve Yaaqov	9.2	13.2	12.0	10.8	Bayit va-Gan	5.7	5.7	4.7	6.0
Pisgat Zeev North	7.5	11.6	10.6	9.3	Ramat Sharet, Ramat Denya	6.5	4.8	5.9	6.5
Pisgat Zeev East	7.0	11.0	10.2	8.9	Qiryat Ha-Yovel north	6.1	5.4	5.0	6.1
Pisgat Ze'ev (northwest)	6.1	10.2	9.4	8.1	Qiryat Ha-Yovel south	6.6	5.0	5.9	6.5
Ramat Shlomo	5.1	9.4	6.9	7.0	Qiryat Menahem, Ir Gannim	8.5	7.0	7.6	8.3
Ramot Allon north	6.5	10.6	7.0	7.7	Manahat slopes	5.6	3.6	6.5	6.0
Ramot Allon	6.0	10.0	6.1	7.3	Gonen (Qatamon)	4.0	1.9	6.1	5.2
Ramot Allon South	6.1	10.2	6.6	7.3	Rassco, Giv'at Mordekhay	3.0	2.8	5.0	4.8
Har Hozvim, Sanhedriya	2.4	6.7	4.6	4.9	German Colony, Gonen	2.5	2.3	5.6	4.7
Ramat Eshkol, G. Mivtar	3.0	7.2	5.7	5.5	Qomemiyyut, YMCA	1.3	3.4	5.2	4.5
Ma'alot Dafna, S. Hanavi	2.0	6.1	5.1	4.9	Ge'ulim, G. Hananya, Y. Moshe	2.8	2.1	6.5	5.2
Giv'at Shapira	3.7	7.8	7.1	6.4	Talpiot, Arnona, M. Hayyim	4.0	1.2	7.5	5.7
Mamila, Morasha	0.9	4.3	5.1	4.6	East Talpiyyot	5.0	3.0	8.8	6.9
Ge'ula, Me'a She'arim	1.2	5.5	4.5	4.5	East Talpiyyot (east)	4.9	3.3	8.8	6.9
Makor Baruch, Z. Moshe	1.3	5.4	3.7	4.4	Homat Shmuel (Har Homa)	7.2	3.4	10.4	8.3
City Center	0.6	4.4	4.4	4.4	Gilo east	7.2	3.6	9.0	7.6
Nahlaot, Zichronot	1.1	4.5	3.7	4.3	Gilo west	8.4	4.9	10.2	8.8
Rehavya	1.5	3.6	4.5	4.4	Talpyiot shopping area	4.4	0.0	7.5	5.7
Romema	3.0	6.6	3.4	5.0	Givat Shaul shopping area	4.4	7.5	0.0	6.0
Giv'at Sha'ul	4.1	7.5	2.8	5.8	Malcha shopping center	5.2	3.1	6.2	5.7
Har Nof	5.1	8.1	2.8	6.6	Romema shopping area	2.0	5.6	3.1	4.5
Q. Moshe, Bet HaKerem	3.5	5.5	2.6	4.8	Central Bus Station	2.0	5.6	3.1	4.5
Nayot	2.9	4.6	3.8	4.8	Mahane Yehuda	1.1	5.0	3.5	4.2

20 percent were open market stalls and 15 percent were grocery stores. The data were collected by ICBS 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. The selected products have the same universal product code (UPC) and are therefore identical across stores (e.g., the same brand, size, packaging, etc.), implying that they have the same quality.[13] The 27 products were chosen among the hundreds of products in the CPI because of their popularity. This guarantees that they are sampled in a relatively large number of stores and that they are actually bought by most households.

The list of products, their mean price and coefficient of variation are displayed in Tables B1 and B2 (Appendix B). The products consist of 13 popular and frequently purchased foodstuffs, 11 fruits and vegetables and 3 miscellaneous products. Many products, notably vegetables, exhibit substantial price dispersion, while others, such as cottage cheese or coffee, have more concentrated price distributions.

We do not observe prices in all 46 neighborhoods, only in 26 of them: 21 residential neighborhoods and five of the six commercial areas (we do not observe prices for the Central Bus Station). Furthermore, not all 60 stores are surveyed in each period. Because in some neighborhoods there are periods when no stores were sampled, we have a total of 73 neighborhood-period observations on prices (instead of $78 = 26 \times 3$). Finally, not all 27 products are surveyed in each store-period combination.

Table 3 lists the neighborhoods where we observe prices, the number of stores in our sample, and the total number of observed products across the neighborhood's various stores, noting again that these products are not necessarily observed in each of the neighborhood's stores. As the table shows, most neighborhoods have a single store in the sample. Mahane Yehuda, an attractive fresh produce open market, has the largest number of stores (all but one are market stalls), followed by the Talpiot shopping area where the hard discount supermarkets are located. The rightmost column of Table 3 reports, in addition, the total number of supermarkets in each neighborhood, regardless of whether prices were sampled in them. This measure, obtained from the ICBS, plays an important role in modeling the extent of within-neighborhood competition. In Table A2 (Appendix A), we report this number of supermarkets for each of the 46 neighborhoods, including those where no prices are observed.

The composite good and its neighborhood-level price. Typically, households perform a main shopping trip once a week, buying a variety of goods. We will therefore focus our analysis on two household choices: where to shop, and how many units of a "composite good" to buy. We

[13] We emphasize that even among fruits and vegetables there are no noticeable quality differences across stores at the same point in time because the ICBS collects prices on produce of a specific type.

Table 3: Number of sampled stores and observed products

	# sampled stores			# observed products			# supermarkets
Neigborhood	Sep07	Nov07	Nov08	Sep07	Nov07	Nov08	
Neve Yaaqov	1	1	1	27	27	27	1
Pisgat Zeev North	1	1	1	26	26	27	1
Ramot Allon north	2	2	2	24	25	25	1
Ramat Eshkol, Giv'at-Mivtar	1	1	1	11	10	9	0
Ma'alot Dafna, S. Hanavi	1	0	0	10	0	0	0
Givat Shapira	2	2	2	27	27	27	2
Ge'ula, Me'a She'arim	3	4	3	12	12	13	0
City Center	1	2	2	6	7	6	2
Rehavya	2	2	2	24	25	24	1
Romema	2	2	2	24	23	22	1
Giv'at Sha'ul	1	1	1	3	4	3	0
Har Nof	1	1	1	25	21	22	1
Qiryat Moshe, Bet Hakerem	3	3	3	27	27	27	2
Nayot	1	1	1	11	11	11	1
Ramat Sharet, Ramat Denya	1	1	0	1	1	0	0
Qiryat Ha-Yovel south	3	2	2	27	26	26	1
Rassco, Giv'at Mordekhay	2	2	2	26	27	27	1
Ge'ulim, G. Hananya, Y. Moshe	1	1	1	26	25	23	1
Talpyot, Arnona, M.Hayim	1	1	1	4	4	2	0
Gilo east	0	1	0	0	1	0	0
Gilo west	2	2	2	12	13	12	0
Talpiot shopping	7	7	7	27	27	27	5
Givat Shaul shopping	3	3	3	27	27	26	3
Malcha shopping	1	1	1	3	4	4	1
Romema shopping	1	1	1	27	27	23	3
Mahane Yehuda	10	10	9	25	24	24	1
Total	54	55	51				

define the price of the composite good charged in a given neighborhood as a weighted average of the prices of its individual products using CPI weights.

Let ω_i be the weight of product i used in the CPI, $i = 1, \ldots, 27$, and let Ω_{nt} be the set of products observed in neighborhood n at time t.[14] Then the price of the composite good is

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

where p_{nit} is the *average* price of product i in neighborhood n in period t across all stores selling the product in the neighborhood and Ω_{nt} is the set of products for which we observe prices in

[14]Note that by using the same weights across all neighborhoods we ensure that differences in the price of the composite good reflect price differences and nothing else. In addition, there are no data on neighborhood-specific CPI weights.

neighborhood n in period t.

We can think of a unit of the composite good underlying the price p_{nt} as composed of a fraction $\omega_i/\sum_{i\in\Omega_{nt}}\omega_i$ of the unit in which product $i's$ price is measured. For example, the composite good in Neve Yaaqov includes 95 gr. of potatoes, 6 percent of a packet of Turkish coffee, etc. The price p_{nt} corresponds to the price of a *single unit* of the composite good. In the model, households are allowed to purchase multiple units of the composite good.

Table 3 shows that the set of products Ω_{nt} varies across neighborhoods. For example, the composite good includes only one product in Ramat Sharet and in Gilo east, but it includes 27 products in Neve Yaaqov. At first glance, this is puzzling because the selected 27 products are every-day popular products that should be available at any reasonable grocery store and supermarket. This stems from the definition of a product as corresponding to a single UPC, and to the fact that some stores may be missing the product because they carry a different version of what is essentially the same product differing, perhaps, in size, packaging or brand.

This variation presents a challenge: we want to define a composite good that would be as homogeneous as possible without reducing the sample size too much. Our leading specification therefore computes the price p_{nt} only for neighborhoods where Ω_{nt} includes at least 21 products. We compute the price p_{nt} in the 15 neighborhoods (including four commercial areas) in Table 3 in which at least 21 items have observed prices, treating prices at the remaining neighborhoods as unobserved.[15] Several robustness checks are performed: we use a threshold lower than 21 products, impute the missing prices by projecting product-specific prices on demographics, construct the composite good from fruits and vegetables only, and use prices from supermarkets only. As reported below, these alternative definitions yield qualitatively similar demand patterns.

The aggregation of prices to the neighborhood level, as opposed to the store level, is motivated by several factors. First, our expenditure data, described below, are at the neighborhood level. Second, since not all items are observed in all stores, the aggregation to the neighborhood level mitigates the incidence of missing prices. Third, residential neighborhoods tend to be served by smaller, more expensive store formats, whereas commercial neighborhoods exhibit larger, hard-discount stores. This suggests that the bulk of the price variation should be observed across, but not within, neighborhoods. This observation is consistent with quantitative analysis: when we regress the prices of each individual good on a set of neighborhood and period dummy variables, we find that these dummies explain at least 50 percent of the price variation in 22 out of the 27 regressions (with 25 out of 27 delivering an R-squared measure of at least 0.43, while the median

[15] Notice that there are no neighborhoods with an observed number of products between 14 and 20. 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).

R-squared is 0.59). These quantitative and qualitative aspects of the variation in prices motivate our focus on neighborhood-level price indices, and will be consistent with our model in which within-neighborhood symmetry in mean-utility levels across stores will be assumed (we return to this issue in Sections 3.1 and 4.1 below).

To gain a sense of the quantitative importance of cross-neighborhood price variation, we examine the savings for residents of a neighborhood j from shopping at the cheapest place in the city instead of at their own home neighborhood j. These gross savings are defined by $100 \times (p_{jt} - Min_n p_{nt})/p_{jt}$ and are computed for each of the 15 neighborhoods with valid prices in each period. The histogram is displayed in Figure 2. The mean gross gain is 13 percent and the maximum gross gain is 22 percent.

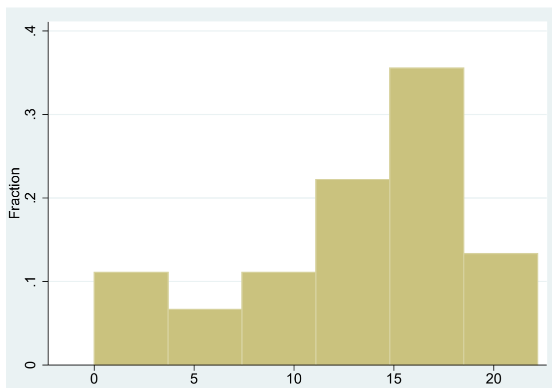

Figure 2: A histogram of percentgae savings from shopping at the cheapest destination across neighborhoods

Table 4 displays the price of the composite good at the 15 neighborhoods, ranked from cheapest to most expensive. There is substantial variation across neighborhoods, with the maximum price being about 24-29 percent above the minimum price.[16] Prices in commercial areas are consistently lower than in most residential neighborhoods (except for the Romema shopping in November 2008). The Talpiot commercial area and the Mahane Yehuda market are always among the cheapest locations. Prices are high not only in the affluent residential neighborhood of Rehavya, but also in less affluent neighborhoods such as Qiryat HaYovel south, Givat Shapira, and Neve Yaaqov. Rankings are persistent: the rank correlation of p_{nt} between September and November 2007 is 0.68, while that between November 2007 and November 2008 at 0.57 is still quite high even though 12 months elapsed between the two measurements. This persistence supports the notion that the location of the cheap stores is well known among Jerusalem residents.

[16] The composite good's price increased by 10% between November 2007 and November 2008. To provide a benchmark, the CPI inflation for food between December 2007 and December 2008 was 8.3%.

Table 4: Price of composite good across neighbothoods and time

Sep-07		Nov-07		Nov-08	
Ramot Allon north	6.23	**Talpyiot shopping area**	6.15	**Talpyiot shopping area**	6.89
Talpyiot shopping area	6.33	Ramot Allon north	6.56	**Givat Shaul shopping**	7.07
Mahane Yehuda	6.84	**Mahane Yehuda**	6.81	**Mahane Yehuda**	7.20
Romema shopping area	7.03	Pisgat Zeev North	6.89	Pisgat Zeev North	7.36
Har Nof	7.13	Har Nof	6.93	Ramot Allon north	7.61
Neve Yaaqov	7.15	**Romema shopping area**	6.99	Har Nof	7.62
Rassco, Giv'at Mordekhay	7.32	Baq'a, Abu Tor, Yemin Moshe	7.06	Baq'a, Abu Tor, Yemin Moshe	7.76
Pisgat Zeev North	7.34	Rehavya	7.27	Qiryat Moshe, Bet Ha-kerem	7.85
Givat Shaul shopping	7.45	**Givat Shaul shopping**	7.30	Rassco, Giv'at Mordekhay	7.87
Giv'at Shapira	7.54	Neve Yaaqov	7.31	Neve Yaaqov	8.01
Qiryat Moshe, Bet Ha-kerem	7.55	Rassco, Giv'at Mordekhay	7.34	Giv'at Shapira	8.14
Romema	7.61	Qiryat Ha-Yovel south	7.36	Romema	8.17
Baq'a, Abu Tor, Yemin Moshe	7.68	Romema	7.38	Qiryat Ha-Yovel south	8.19
Qiryat Ha-Yovel south	7.80	Giv'at Shapira	7.39	Rehavya	8.52
Rehavya	8.01	Qiryat Moshe, Bet Ha-kerem	7.61	**Romema shopping area**	8.69
Mean	7.27		7.09		7.80
Standard deviation	0.50		0.38		0.52

Notes: the table lists the price of the composite good in each location and time period where it could be computed using at least 21 observed products (see text). Commercial areas appear in bold.

Insights into our research question are provided by exploring the distribution of prices across neighborhoods in Figures 3 and 4, describing the 15 prices in our third sample period, November 2008. Figure 3 shows that some of the highest prices in the city are charged by retailers located in the peripheral neighborhoods of Neve Yaaqov, Givat Shapira and Qiryat HaYovel.

Figure 4 plots composite good prices against housing prices, along with a linear predicted line (note that commercial areas also have a small residential population, explaining why we observe residential housing prices there). This figure clarifies that retailers in the three peripheral neighborhoods mentioned above charge some of the highest prices, *despite the fact that these are some of the least affluent residential neighborhoods.* Neighborhoods such as Geulim (Baqa) or Bet Hakerem, in contrast, are much more affluent, yet pay lower prices. From Figure 3, we see that the latter two neighborhoods are located in the vicinity of the cheaper supermarkets in the major commercial areas, Talpiot and Givat Shaul. In our model, this spatial feature would imply that prices in these two neighborhoods are disciplined by the lower prices at the commercial areas, whereas no such effect operates in the peripheral neighborhoods.[17]

[17] The linear predicted line in Figure 4 suggests a positive relationship between composite good and housing prices. While one may be tempted to conclude that "the rich pay more," we note that the small number of data points (15 observations) is not enough to draw such a conclusion.

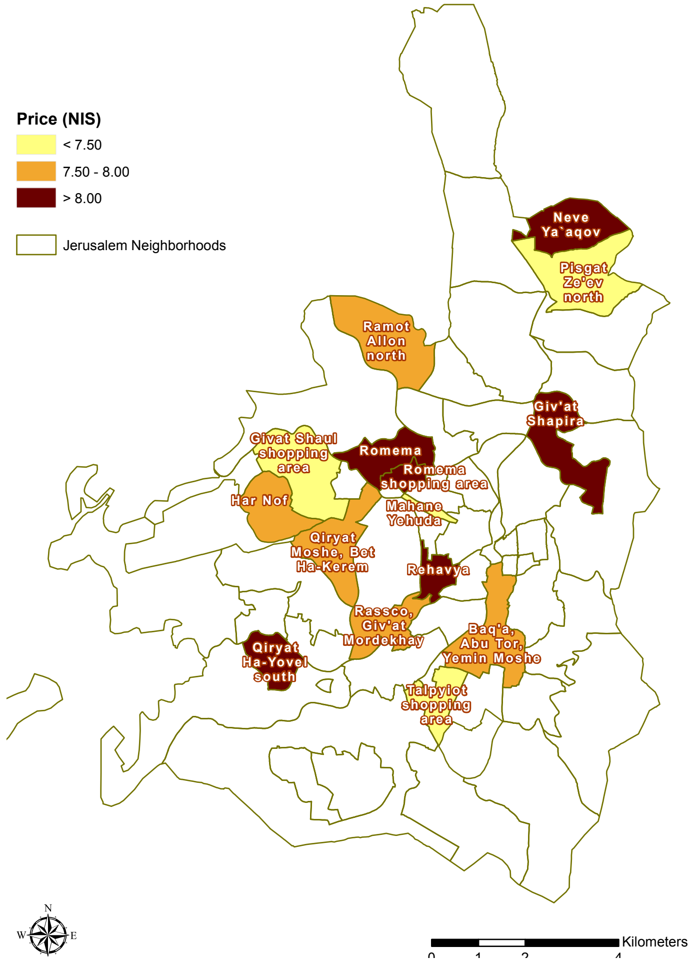

Figure 3: Composite good prices across the city, November 2008

2.3 Expenditure data

We obtained data on consumers' expenditures from a credit card company that operates in Israel. Institutional details suggest that customers of this company should not be different from customers of other companies. The use of debit cards is minimal in Israel. Our data should therefore be representative of transactions performed via payment cards. Nonetheless, grocery shopping is also performed using cash and checks, and our framework shows how to exploit the panel structure of the data to address the measurement error that results from this omission.

The data capture expenditures by Jerusalem's residents in supermarkets, grocery stores, bakeries, delicatessen, butcher stores, wine stores, fruits and vegetables stores and health stores, covering all store types where our 27 products are likely to be sold. We observe total neighborhood-level expenditures during the same three periods for which we have price data. The data list total expenditures by residents of each origin neighborhood j performed at each destination

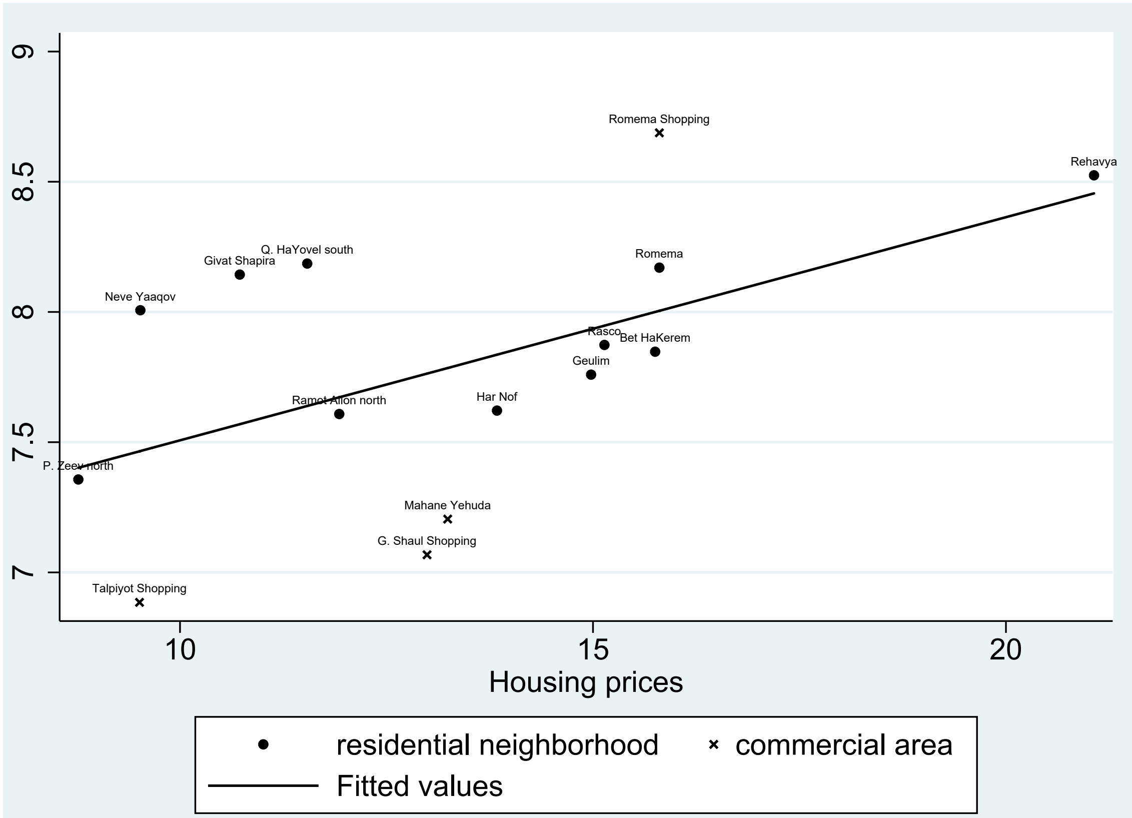

Figure 4: Composite good prices plotted against housing prices, November 2008

neighborhood n where $j, n \in \{1, ..., 46\}$.Simply put, the expenditure data are provided in a 46 by 46 matrix providing the expenditure flow between each pair of neighborhoods. The data were constructed as follows: first, the neighborhood of residence for individual card holders was identified using their zip codes. Similarly, the destination neighborhoods for particular transactions by card holders were identified using the stores' zip codes.[18] Finally, the expenditure data were aggregated to the neighborhood level matrix described above, and were provided to us at that level of aggregation (that is, we do not observe data at the individual household or store level).

We also observe the total expenditures of residents of each origin neighborhood at destinations outside the city. Jerusalem does not have substantial satellite cities surrounding it which provide attractive shopping opportunities. We therefore conjecture that a substantial portion of the shopping outside the city may represent cases where individuals actually reside outside Jerusalem,

[18]This required a mapping between zipcodes and neighborhoods (subquarters). Such a mapping is not trivial since zipcodes can map into multiple neighborhoods. We created a unique mapping of zip codes into subquarters via a "majority rule": the zip code was mapped to the subquarter with which it has the largest geographical overlap. We thank Elka Gotfryd from the Department of Geography at The Hebrew University for her invaluable help.

yet their mailing address erroneously identifies them as Jerusalem residents (e.g., students who study in universities outside the city but have not updated their mailing address). For this reason, in our structural model, we will define the "outside option" as shopping in the 31 destinations in Jerusalem where we do not have valid price data, ignoring expenditures outside Jerusalem.[19]

Table 5 provides descriptive statistics regarding the expenditure data. The most popular commercial area is Talpiot where, on average, 27 percent of expenditures are incurred. The top destination accounts for 42 percent of the expenditures on average. In many cases (16 to 20 out of the 46 neighborhoods depending on the period), the top destination is the Talpiot commercial area. Givat Shaul is at a distant second place, although it is quite popular among nearby neighborhoods (e.g., Har Nof, Bet Hakerem, Nayot). Most expenditures are not incurred within the home neighborhood, yet home-neighborhood shopping is substantial capturing, on average, 22% of total expenditures. The home destination is the top destination in 12 to 17 cases depending on the period.

We would like to use these expenditure data to infer the probability of shopping at the various destinations in the city for the purpose of computing the expected price incurred by residents of each origin neighborhood. This, however, requires us to overcome two main issues. First, consumers purchase different quantities of the composite good, implying that the level of expenditures is not a direct indicator of the incidence of purchase. Second, the credit card data raises measurement issues: they do not cover expenditures made in cash or checks, and do cover more than the 27 products included in our composite good. The model presented in the next section will allow us to clarify the assumptions needed to overcome these challenges.

3 A structural model of demand in the city

The model of households' preferences is presented in Section 3.1. Section 3.2 derives the estimating equation for this model, while Section 3.3 presents the estimated parameters and the implied demand elasticities. In Section 3.4 we use the shopping probabilities implied by the model to compute expected prices for residents of each neighborhood.

[19] Robustness checks in which we added the expenditures incurred outside Jerusalem to the outside option yield remarkably close results to the ones reported in Section 3.3, reassuring us that this measurement issue does not drive our findings.

Table 5: Credit card expenditure fractions

Neighborhood	Fraction spent at			
	Own neighborhood	Top neighborhood	Talpiot	Givat Shaul
Neve Yaaqov	0.25	0.36	0.03	0.02
Pisgat Zeev North	0.68	0.68	0.10	0.03
Pisgat Zeev East	0.22	0.23	0.23	0.06
Pisgat Ze'ev (northwest & west)	0.01	0.35	0.24	0.08
Ramat Shlomo	0.18	0.28	0.01	0.02
Ramot Allon north	0.25	0.25	0.12	0.06
Ramot Allon	0.15	0.15	0.15	0.08
Ramot Allon South	0.31	0.31	0.18	0.11
Har Hozvim, Sanhedriya	0.08	0.32	0.01	0.02
Ramat Eshkol, Givat Mivtar	0.56	0.56	0.05	0.02
Ma'alot Dafna, Shmuel Hanavi	0.18	0.28	0.08	0.02
Giv'at Shapira	0.42	0.42	0.18	0.04
Mamila, Morasha	0.05	0.29	0.29	0.06
Ge'ula, Me'a She'arim	0.24	0.32	0.06	0.02
Makor Baruch, Zichron Moshe	0.03	0.35	0.04	0.02
City Center	0.10	0.18	0.16	0.05
Nahlaot, Zichronot	0.03	0.33	0.17	0.04
Rehavya	0.44	0.44	0.19	0.03
Romema	0.54	0.54	0.03	0.02
Giv'at Sha'ul	0.60	0.60	0.03	0.16
Har Nof	0.30	0.31	0.01	0.31
Qiryat Moshe, Bet HaKerem	0.14	0.38	0.16	0.18
Nayot	0.08	0.22	0.14	0.20
Bayit va-Gan	0.05	0.21	0.17	0.10
Ramat Sharet, Ramat Denya	0.12	0.31	0.31	0.07
Qiryat Ha-Yovel north	0.21	0.21	0.21	0.07
Qiryat Ha-Yovel south	0.33	0.33	0.31	0.05
Qiryat Menahem, Ir Gannim	0.52	0.52	0.21	0.03
Manahat slopes	0.07	0.55	0.55	0.06
Gonen (Qatamon) A - I	0.07	0.55	0.55	0.03
Rassco, Giv'at Mordekhay	0.31	0.47	0.47	0.03
German Colony, Gonen (Old Qatamon)	0.07	0.61	0.61	0.03
Qomemiyut (Talbiya), YMCA Compound	0.01	0.31	0.29	0.05
Geulim (Baqa), Givat Hananya, Yemin Moshe	0.00	0.65	0.65	0.02
Talpiyyot, Arnona, Mekor Hayyim	0.15	0.71	0.71	0.02
East Talpiyyot	0.01	0.71	0.71	0.03
East Talpiyyot (east)	0.01	0.66	0.66	0.02
Har Homa	0.00	0.72	0.72	0.03
Gilo east	0.21	0.46	0.46	0.02
Gilo west	0.26	0.46	0.46	0.03
Talpiot shopping area	0.76	0.76	0.76	0.03
Givat Shaul shopping area	0.41	0.41	0.06	0.41
Malcha shopping center	0.01	0.60	0.60	0.05
Romema shopping area	0.60	0.60	0.04	0.03
Central Bus Station	0.14	0.27	0.16	0.01
Mahane Yehuda	0.06	0.26	0.26	0.08
Average	0.22	0.42	0.27	0.06

Notes: The table shows, for each neighborhood, the fractions (averaged over the sample period) of its residents' expenditures spent at the neighborhood itself, at the top destination, and at the Talpiot and Givat Shaul shopping centers.

3.1 A model of household preferences

Define the set $\mathcal{J}$ of origin neighborhoods – the "origins" – such that $J = |\mathcal{J}| = 46$ in our application.[20] A household residing in any one of the origin neighborhoods $j = 1, ..., J$ makes a discrete choice of where to shop for the composite good, and a continuous choice: how many units of this good to purchase.[21] We note that consumers certainly purchase more than the 27 items that are included in the composite good. As we show below, we formally take into account the discrepancy between the observed expenditure on all items and the expenditure on the more restricted composite good. Furthermore, our model of consumer preferences controls for differences in availability and variety of additional items across destinations using fixed utility effects.

We model a total of sixteen possible shopping destinations – the "destinations" – for each household. Let $\mathbb{N}$ denote the set of fifteen Jerusalem neighborhoods in which we observe the price of the composite good, indexed by $n = 1, ..., N$ where $N = |\mathbb{N}| = 15$. We let $n = 0$ denote the outside option, defined as shopping in one of Jerusalem's neighborhoods for which we do not observe prices. Put differently, each of the 46 origins can serve as a destination, but we distinguish between two types of destinations: destinations that belong in $\mathbb{N}$ (i.e., where prices are observed) represent 15 "inside options," while all remaining 31 destinations (i.e., neighborhoods that belong in $\mathcal{J}$ but not in $\mathbb{N}$) are lumped together as the outside option. In particular, note that $\mathbb{N} \subset \mathcal{J}$. We maintain that this limitation is not crucial since, as required for the computation of the monthly CPI, our observed prices cover the main commercial areas and important residential neighborhoods . Neighborhoods where we do not observe prices typically do not feature attractive retailers such as supermarkets or important minimarkets.

The continuous choice — how many units of the composite good to purchase — is modeled by adapting Björnerstedt and Verboven's (2016, hereafter BV) nested logit model of demand to our setup. Their chosen functional form implies that households spend a constant fraction of their income on the composite good. While the literature offers more sophisticated strategies for introducing this continuous dimension (e.g., Smith 2004, Figurelli 2013) into supermarket demand, those papers have relied on different data (namely, scanner, micro-level data) and addressed different questions relative to our work. In the context of our aggregate (neighborhood-level) demand data, we view the simple strategy adopted as an attractive choice.

[20] Recall that the six shopping areas also contain some small residential population. To maintain internal consistency, we therefore consider each commercial area as a residential origin.

[21] Our model assumes that households purchase the composite good on a single shopping trip. Smith (2004) uses household level survey data to show that households concentrate their grocery shopping in a single shopping trip, but also engage in "top-up" trips. Our observed aggregate expenditures are uninformative about such distinctions, and we therefore do not model them.

The nests are destination neighborhoods, allowing stores within a neighborhood to be closer substitutes than stores located in different neighborhoods. Stores within a neighborhood are symmetrically differentiated: they offer identical mean utility levels, but are allowed to offer distinct benefits to individual households via idiosyncratic error terms.[22] This symmetry assumption is motivated by the fact that our expenditure data are at the aggregate destination neighborhood level, rather than at the individual store level. Note that we do observe prices at the store level. However, the symmetry assumption allows us to construct a neighborhood-level price index (see section 2) that utilizies price information from several stores within the neighborhood. Given that not all items are observed in all stores, the neighborhood-level price index is advantageous relative to a store-level price index.

While motivated by a practical data issue, the symmetry assumption is, in fact, reasonable and not particularly restrictive. As reviewed in the data section above, most of the price variation is explained by neighborhood and time dummy variables. This quantitative finding is consistent with institutional details: supermarkets in a residential neighborhood would typically all be of a certain format (smaller, expensive supermarkets) whereas supermarkets in a commercial area are hard-discount, larger supermarkets. Finally, note that the economic content of this assumption is not that stores within a neighborhood are homogenous. Rather, it is assumed that they have an identical mean utility level. Individual consumers do not view them as perfect substitutes because of the non-symmetric idiosyncratic terms. For example, a certain household may strictly prefer one of the neighborhood's supermarkets because of its greater proximity to the household's residence.

Omitting the time index from the notation, the (indirect) utility of household h residing in neighborhood $j \in \mathcal{J}$ from buying the composite good at store s located in neighborhood $n \in \mathbb{N}$ is given by

$$U_{hjsn} = \nu_c + \nu_j + \nu_n + hp_j \cdot \nu_n + \left(\gamma^{-1} \ln y_j - \ln p_{sn}\right) \cdot x_j \alpha - d_{jn} \cdot x_j \beta + \kappa \cdot h_{jn} + \zeta_{hn}(\sigma) + (1-\sigma)\epsilon_{hjsn} \quad (2)$$

The constant ν_c shifts the utility from all "inside options" relative to the utility from the outside option. The origin fixed effects ν_j capture utility differences across origin neighborhoods (essentially their different valuations of the outside option, as will become clear below). The destination fixed effects ν_n capture quality differences across destinations. These capture various amenities at location n (parking space, opening hours, etc.), and, in addition, may also capture differences in grocery product variety (i.e., the availability of products other than our basic 27 items) and the availability of additional attractions (e.g., other businesses). The term $hp_j \cdot \nu_n$

[22] Of course, stores located in different neighborhoods are characterized by different mean utility levels. See Berry and Waldfogel (1999) for a model with symmetric differentiation of products within local markets.

interacts the origin neighborhood's housing prices with the destination neighborhood's fixed effect. This allows us to control for the possibility that residents of more affluent neighborhoods systematically prefer certain destinations that offer amenities that are attractive for an affluent population (e.g., a health store or a spa). The mean income in neighborhood j is y_j, and is unobserved. The vector x_j contains the demographic features of neighborhood j displayed in Table 1 (and a constant term). The shortest road distance between each origin neighborhood j and each destination neighborhood n is denoted by d_{jn} for any $(j, n) \in \mathcal{J} \times \mathbb{N}$.

The price charged by store s in neighborhood n is denoted by p_{sn}. Given the symmetric differentiation of stores within a neighborhood, we will focus on equilibria where prices also satisfy within-neighborhood price symmetry, i.e., $p_{sn} = p_n$ for every store s in neighborhood n. The price p_n is computed from the observed data using (1). This assumption will be consistent with the pricing model introduced in Section 4.1.[23] The price vector is denoted by $\mathbf{p} = (p_1, ..., p_N)$.

The parameter vectors α and β capture price and distance sensitivities, respectively. These sensitivities vary with the origin neighborhood demographic characteristics (e.g., the percentage of individuals owning a car, percentage of senior citizens, etc.).[24] Price and distance sensitivities do not have to be affected by the same demographics because some of the elements of α and β can be set to zero. Note that the marginal utility from money is decreasing because of the logarithm specification. Additional flexibility is allowed by interacting the price regressor with origin-neighborhood housing prices, serving as a proxy for income.

The "shopping at home" dummy variable h_{jn} takes the value 1 if $j = n$, and zero otherwise. As we already account for the effect of distance via d_{jn}, κ reflects the benefits of shopping in the home neighborhood on top of the implied savings of travel time (and direct travel costs). Put differently, κ introduces nonlinearity in the household's travel costs: it captures a "fixed cost" associated with shopping outside the home neighborhood, possibly related to the need to drive, or give up a convenient parking space near home.

The idiosyncratic term $\zeta_{hn}(\sigma) + (1-\sigma)\epsilon_{hjsn}$ follows the typical assumptions for the nested logit model (Berry 1994). The shock ϵ_{hjsn} is drawn from a Type-I Extreme Value distribution that is I.I.D. across all households, origins, destination stores and time (the latter's index is omitted here). It captures idiosyncratic variation in the utility of shopping at store s in destination n for a particular household living in neighborhood j. For example, this household may particularly

[23] We must introduce the notation in a way that allows stores within a neighborhood to charge different prices since characterizing the pricing equilibrium involves writing down each store's first-order condition with respect to its own price.

[24] Note that we do not use household level data on demographic characteristics but rather neighborhood-level means. An alternative would be to estimate a random coefficient model by drawing from the observed empirical distribution of these demographic variables in each neighborhood (Berry, Levinsohn and Pakes (1995), Nevo (2001)). The homogeneity assumption, however, substantially simplifies the estimation.

value shopping at this store s if it is on the way home from work, or close to the kids' school. The random variable ζ_{hn} has a unique distribution that depends on the parameter σ and guarantees that the entire term $\zeta_{hn}(\sigma)+(1-\sigma)\epsilon_{hjsn}$ follows the Type-I Extreme Value distribution (Cardell 1997). The parameter σ takes values in the interval $[0,1)$. As this parameter approaches zero, the term ζ_{hn} approaches zero as well, corresponding to the familiar conditional logit model (McFadden 1974). In contrast, as this parameter approaches 1, the unobserved tastes of household h towards stores located in destination neighborhood n become perfectly correlated. This parameter, therefore, governs the intensity of within-neighborhood competition: higher values of it imply that stores located in the same neighborhood become closer substitutes to one another.[25]

It is convenient to decompose the utility function as follows:

$$U_{hjsn} = \gamma^{-1} \ln y_j \cdot x_j\alpha + \delta_{jsn} + \zeta_{hn}(\sigma) + (1-\sigma)\epsilon_{hjsn},$$

where

$$\delta_{jsn} = \nu_c + \nu_j + \nu_n + hp_j \cdot \nu_n - \ln p_{sn} \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn}$$

is the mean utility level, common to all origin-j residents who shop at s in destination n. Notice that, as long as stores within a given neighborhood n charge symmetric prices, i.e., $p_{sn} = p_n$, the mean utility is symmetric across these stores as well.

The model is completed by specifying the utility of a resident of neighborhood j from shopping at the outside option $n = 0$, defined as the only member of its nest:

$$U_{hjs0} = \gamma^{-1} \ln y_j \cdot x_j\alpha + \zeta_{h0}(\sigma) + (1-\sigma)\epsilon_{hjs0} \tag{3}$$

This definition normalizes, without loss of generality, j-residents' mean utility from the outside option at $\delta_{j0} = 0$. The terms v_j in the mean utility δ_{jsn} associated with "inside options" allow for heterogeneity in the utility from the outside option across origin neighborhoods. This is particularly important given that, for residents of neighborhoods in which the price is not observed, the choice to shop in their home neighborhood is considered part of the outside option.

Choice probabilities. Integrating over the density of the idiosyncratic terms delivers the familiar nested logit formula for the probability that a resident from origin neighborhood j shops at store s located in neighborhood n, conditional on shopping at n,

$$\pi_{js/n}(\mathbf{p};\theta) = e^{\left(\gamma^{-1}\ln y_j \cdot x_j\alpha + \delta_{jsn}\right)/(1-\sigma)} / D_{jn} \tag{4}$$

[25] If consumers prefer traveling to a commercial area because it allows them to visit several supermarkets and buy different items in each, the nested logit structure would be misspecified, as it does not allow supermarkets to serve as complements. At the same time, most consumers are not likely to split their grocery shopping across two stores within a single shopping trip. Moreover, greater product variety in shopping areas is controlled for via the ν_n destination fixed effects.

where $\theta = (\alpha, \beta, \kappa, \sigma)$ are the model's parameters, and the term D_{jn} is defined by

$$D_{jn} = \sum_{s=1}^{L_n} e^{\left(\gamma^{-1}\ln y_j \cdot x_j\alpha + \delta_{jsn}\right)/(1-\sigma)} \quad \text{for } n = 1, ..., 15, \text{ and } D_{j0} = e^{\gamma^{-1}\ln y_j x_j \alpha/(1-\sigma)}$$

where L_n denotes the number of retailers located in neighborhood n. In the empirical application, we take this to be the number of supermarkets, as reported in Table 3, with certain adjustments that account for the role of additional store formats such as grocery stores and market stands (noting that one of our robustness checks, reported in Appendix Table F1, estimates the demand model using prices sampled in supermarkets only). We return to this when discussing the supply-side model in Section (4.1) below.

The probability that a resident from origin j shops in neighborhood n (the "nest share") is,

$$\pi_{jn}(\mathbf{p}; \theta) = D_{jn}^{1-\sigma} / \sum_{m=0}^{N} D_{jm}^{1-\sigma} \tag{5}$$

The probability of shopping at store s located in neighborhood n is given by multiplying the terms in (4) and (5). Imposing within-neighborhood price symmetry, $p_{sn} = p_n$, we have,

$$\begin{aligned} &D_{jn} = L_n \cdot e^{\left(\gamma^{-1}\ln y_j \cdot x_j\alpha + \delta_{jn}\right)/(1-\sigma)} \\ &\delta_{jsn} = \delta_{jn} = \gamma^{-1}\ln y_j \cdot x_j\alpha + \nu_c + \nu_j + \nu_n + hp_j \cdot v_n - \ln p_n \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn} \\ &\pi_{js/n}(\mathbf{p}; \theta) = 1/L_n \\ &\pi_{jsn}(\mathbf{p}; \theta) = \pi_{jn}(\mathbf{p}; \theta)/L_n \end{aligned} \tag{6}$$

Quantity choice. Conditional on buying at store s in destination n, the quantity demanded by household h residing in neighborhood j of the composite good is, using Roy's identity, $q_{hjsn} = \gamma \frac{y_j}{p_{sn}}$, so that expenditure on the composite good is a constant fraction γ of the (representative) household's income.[26] Since our estimation procedure (see equation (9) below) relies on normalized expenditures, we can allow the fraction γ to vary across origin neighborhoods, and we do not need to estimate it. For notational simplicity, therefore, we keep it constant.

In an equilibrium with $p_{sn} = p_n$ each store in the neighborhood is visited with equal probability and demand per household residing in neighborhood j for the composite good sold at destination n is

$$q_{hjn} = \gamma \frac{y_j}{p_n} \tag{7}$$

[26] Defining $f(y_j, p_{sn}) = (\gamma^{-1}\ln y_j - \ln p_{sn})$, Roy's identity implies that $q_{hjsn} = -\frac{\partial f/\partial p_{sn}}{\partial f/\partial y_j}$.

Finally, we note that the expected monetary expenditure of household h residing in neighborhood j in destination neighborhood n at time t can be written as $e_{hjnt} = \pi_{jnt} q_{hjnt} p_{nt} = \pi_{jnt} \gamma y_j$, using (7) and taking income to be time-invariant. Because income is assumed identical across households within the neighborhood, q_{hjnt} and e_{hjnt} do not vary within the neighborhood, and aggregate expenditures by neighborhood j residents in neighborhood n are,

$$E_{jnt} = H_j e_{hjnt} = H_j \pi_{jnt} \gamma y_j \tag{8}$$

where H_j is the number of households in neighborhood j. As we show below, observing H_j will not be necessary for our analysis.

3.2 Estimating the demand model

Motivated by the within-neighborhood store symmetry, we pursue a variant of Berry's (1994) inversion strategy: rather than inverting a product (in our case, store) level market share equation, we invert a nest-level expenditure share equation that equates the *nest* expenditure shares predicted by the model to those observed in the data. This enables us to solve for the mean utility level. Using (5), (8) and the definition of the mean utility δ_{jn} from (6), we obtain:[27]

$$\begin{aligned}
\ln\left(\frac{E_{jnt}}{E_{j0t}}\right) &= \ln\left(\frac{H_j \pi_{jnt} \gamma y_j}{H_j \pi_{j0t} \gamma y_j}\right) = \ln\left(\frac{\pi_{jnt}}{\pi_{j0t}}\right) = \ln(L_n^{1-\sigma} \cdot e^{\delta_{jn}}) \\
&= (1-\sigma)\ln L_n + \delta_{jnt} \qquad (9) \\
&= \nu_c + \nu_j + (\nu_n + (1-\sigma)\ln L_n) + hp_j \cdot \nu_n + \nu_t - \ln p_{nt} \cdot x_j \alpha - d_{jn} \cdot x_j \beta + \kappa \cdot h_{jn}
\end{aligned}$$

Importantly, the time-invariant number of symmetric retailers at destination n, L_n, cannot be separated from the destination fixed effect v_n, implying that identification of the parameter σ will not be possible without variation over time in the number of competitors. We discuss below our approach for tackling this issue.

Equation (9) cannot be estimated just yet, as it has no error term. Moreover, the left-hand side contains expenditure shares that are implied by the model but are measured with error in the data. There are two sources for this measurement error. First, observed prices pertain to (at most) 27 products, whereas observed credit-card expenditures correspond to purchases of many additional products. Second, we observe credit-card expenditures instead of total expenditures.

Let $\tilde{E}^{cc}_{jnt}$ denote the observed credit-card expenditures by neighborhood j residents in neighborhood n at time t. These are expenditures at all relevant establishments (supermarkets, grocery

[27] Note that the time fixed effect v_t is part of the definition of δ_{jnt}. Again, the model in Section 3.1 omited all time indices for expositional clarity.

stores, bakeries, etc.), as described in Section 2.3, i.e., they contain expenditures on products other than the 27 in our composite good. Let E_{jnt} be the *unobserved* expenditures on our composite good made of (at most) 27 products using any payment means (cash, credit cards and checks). The model yields predictions for the unobserved E_{jnt} rather than for the observed $\tilde{E}^{cc}_{jnt}$.

To link both types of expenditures we let $\tilde{E}_{jnt}$ denote expenditures using any payment means on *all* products sold at the relevant establishments (i.e., not just on our 27 products). Without loss of generality, we can always express expenditures on the 27 products, E_{jnt}, as a proportion of $\tilde{E}_{jnt}$,

$$E_{jnt} = \lambda_{jnt}\tilde{E}_{jnt} \tag{10}$$

where $0 \leq \lambda_{jnt} \leq 1$. Similarly, observed *credit-card* expenditures on *all* products, $\tilde{E}^{cc}_{jnt}$, can also always be expressed as a proportion of expenditures by any payment means on *all* products $\tilde{E}_{jnt}$,

$$\tilde{E}^{cc}_{jnt} = \tau_{jnt}\tilde{E}_{jnt} \tag{11}$$

where $0 \leq \tau_{jnt} \leq 1$. Combining the above definitions, we get that *observed* expenditures $\tilde{E}^{cc}_{jnt}$ are related to total expenditures on the composite good for which we observe prices, E_{jnt}, by

$$\tilde{E}^{cc}_{jnt} = \frac{\tau_{jnt}}{\lambda_{jnt}}E_{jnt} \tag{12}$$

Substituting into (9), we obtain an equation in terms of observed expenditures,

$$\ln\left(\frac{\tilde{E}^{cc}_{jnt}}{\tilde{E}^{cc}_{j0t}}\right) = \nu_c + \nu_j + (\nu_n + (1-\sigma)\ln L_n) + hp_j \cdot \nu_n + \nu_t - \ln p_{nt} \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn} + w_{jnt} \tag{13}$$

where $w_{jnt} = \ln\left(\frac{\tau_{jnt}}{\lambda_{jnt}}\frac{\lambda_{j0t}}{\tau_{j0t}}\right)$.

The measurement error w_{jnt} therefore plays the role of the econometric error term. The other terms unobserved by the econometrician are $(\nu_c, \nu_j, \nu_n, \nu_t)$. What matters for consistent estimation of $\theta = (\alpha, \beta, \kappa, \sigma)$ is the correlation between the unobservables and the regressors (prices and distances). The structure of our data – multiple destinations for each origin and vice-versa, as well as three periods of data on prices and expenditures – enables us to control for the unobserved $(\nu_c, \nu_j, \nu_n, \nu_t)$ via dummy variables (where ν_c is simply a constant). We therefore allow $(\nu_c, \nu_j, \nu_n, \nu_t)$ to be correlated with prices and distances. With respect to w_{jnt}, we assume the following:

Assumption 1. *Conditional on origin, destination and time fixed effects, w_{jnt} is uncorrelated with prices and distances.*

This assumption implies that the proportionality factors λ_{jnt} and τ_{jnt} may depend on fixed neighborhood characteristics but not on prices and distance, given these characteristics. For example, the fraction of expenditures made through credit card purchases, τ_{jnt}, may differ across destinations (e.g., there are less credit card purchases in the open market of Mahane Yehuda) but these differences are not related to prices nor to distances to these destinations, conditional on the fixed effects. We also allow τ_{jnt} to vary across origin neighborhoods because differences in income, age composition, etc., may be correlated with the extent of credit card use. The fraction λ_{jnt} of total expenditures accounted for by our composite good may also vary across origins and destinations. Because the composite good includes very popular and basic everyday products, per-capita expenditures are not likely to vary much across households. The variation in λ_{jnt} would then be a result of the variation in total expenditures on all goods, $\tilde{E}_{jnt}$, which is likely to be correlated with households' income, composition and other demographics and less with the price of our composite good at destination n. Assumption 1 relates this variation to neighborhood characteristics, but not to prices at destination nor to distance to it, given these characteristics. We believe these are reasonable assumptions in the present context.

Using Assumption 1, we can linearly project w_{jnt} on origin, destination and time dummies and write it as a linear combination of these dummies and a projection error u_{jnt} uncorrelated with the dummies (and therefore demographics), by construction, and with the distances, d_{jn}, and prices at destination, p_{nt}, by assumption. The estimating equation therefore becomes

$$\ln\left(\frac{\tilde{E}^{cc}_{jnt}}{\tilde{E}^{cc}_{j0t}}\right) = \phi_c + \phi_j + \phi_n + \phi_t + hp_j \cdot v_n - \ln p_{nt} \cdot x_j\alpha - d_{jn} \cdot x_j\beta + \kappa \cdot h_{jn} + u_{jnt} \tag{14}$$

Purging w_{jnt} of its correlation with the various fixed effects implies that estimating the origin, destination and time dummies (ϕ_j, ϕ_n, ϕ_t) would not identify the origin, destination and time fixed *utility effects* (v_j, v_n, v_t). This issue will require some attention when analyzing the quantitative implications of the estimated model such as choice probabilities, elasticities, and margins. The consistent estimation of the parameters (α, β, κ), however, only requires Assumption 1.

We estimate equation (14) by OLS. In constructing the regressors entering (14) we note that, since x contains a constant term, the first term in $-d_{jn} \cdot x_j\beta$ equals $-d_{jn}\beta_d$, while the first term in $-\ln p_{nt} \cdot x_j\alpha$ equals $-\ln p_{nt}\alpha_p$. Both price and distance sensitivities, therefore, have a base parameter (β_d and α_p, respectively) and interactions with demographic effects. Observations used to estimate the parameters in (14) consist of all triplets (j, n, t) pertaining to origin neighborhood j, destination neighborhood n and time period t.

A practical issue with this regression is "zero" expenditure shares. While the nested logit specification predicts a positive expenditure share by residents of any origin j at any destination n, observed credit card expenditures $\tilde{E}^{cc}_{jnt}$ are sometimes zero. When this occurs (about 12 percent of all potential observations), we cannot compute the LHS variable in (14) for that observation. Our practical solution is to drop such observations from the sample implying that our actual sample size is reduced from a potential $46 \times 15 \times 3 = 2070$ observations to 1819 observations. The results are qualitatively robust to substituting a very small number for $\tilde{E}^{cc}_{jnt}$.[28]

Identification: an informal discussion. Identification of β_d, the first term in $d_{jn} \cdot x_j\beta$, is obtained by relating the variation in expenditures (net of origin, destination, time and distance effects) in location n to the variation in the distance to n from neighborhoods having the same demographics. Identification of the other elements of β is obtained by relating this net variation in expenditures to the variation in demographics across neighborhoods having the same distance to n. Identification of α_p, the first term in $\ln p_{nt} \cdot x_j\alpha$, is obtained by relating the net variation in expenditures to the variation in price over time in the same destination neighborhood. Identification of the other elements of α is obtained by relating the net variation in expenditures in location n to the variation in demographics across neighborhoods. Note that since we have multiple observations on expenditures in destination n and from origin j, we could estimate destination and origin fixed effects (ϕ_n and ϕ_j) even with a single data period.

The parameter σ, however, is fundamentally unidentified, posing a difficult problem. Absent variation over time in L_n, the number of competitors in destination n, we cannot separately identify the components of ϕ_n: note that in the estimation equation (14), the fixed effect ϕ_n captures the sum of the utility terms $v_n + (1-\sigma)\ln L_n$ from equation (13) and the linear projection of w_{jnt} on the n-destination dummy variable.[29] One possible solution would be to combine supply-side moments (e.g., requiring that marginal costs would be independent of certain neighborhood-level characteristics) along with the demand-side moments to pin down σ. Instead, the solution we employ in practice is to calibrate σ so that it generates reasonable markups. While this simpler approach has limitations, it alleviates the need to rely on our pricing model in generating the demand estimates. We further discuss this approach in Section 3.3.

Our model and estimation follow familiar strategies in the IO literature based on the nested logit model (McFadden 1978) and on Berry's (1994) inversion strategy for the estimation of demand functions using aggregate data. In our setup, each origin neighborhood constitutes a "market," and retailers, nested into destination neighborhoods, play the role of "products" over which households make a discrete choice. Three aspects distinguish our strategy from the standard approach. First, we adopt BV's (2016) version of the nested logit model which allows

[28]See Appendix F. This is not a formally valid correction but one often used in practice. Gandhi, Lu and Shi (2013) propose a partial-identification strategy to address this type of challenge.

[29]The problem does not arise from our choice to invert the nest shares. Were we to invert the individual product (store) shares, as it is typical, the estimation equation would include the term $\sigma \ln(1/L_n)$ (where $1/L_n$ is the within-nest share, see Berry 1994). Once again, this term would be absorbed by the fixed effect ϕ_n.

for non-constant purchased quantities across households. As in their framework, log price, rather than price, appears on the right-hand side of the estimating equation. Second, we invert nest shares rather than "product" shares. Third, we explicitly model measurement error in the context of our data, and use it to construct the econometric error term.

The standard approach, in contrast, typically ignores measurement error and derives the econometric error term by specifying an unobserved random shifter at the product level. In our context, this would imply adding an unobserved utility shifter v_{jnt} to equation (9), which would be known to firms and therefore correlated with prices, generating an endogeneity problem. We do not specify such an error term because we view the measurement error as a more serious threat to identification in our setup than the potential presence of v_{jnt}.

If, however, systematic demand unobservables v_{jnt} are present our model will be misspecified and this would jeopardize our estimation strategy. The presence of v_{jnt} would imply that residents of certain origin neighborhoods j have a systematic preference for traveling to certain destination neighborhoods n, over and above the overall tendency to travel to n (which is controlled for by the v_n fixed effect), and for reasons not related to the distance d_{jn} or to the price at the destination p_n. We do not expect such systematic tendencies to be important. One scenario that could generate such tendencies is that residents of affluent origin neighborhoods may prefer traveling to specific destinations since these destinations offer unobserved amenities that are particularly attractive to wealthy individuals.[30] We included the term $hp_j \cdot \nu_n$ (origin's housing prices interacted with destination fixed effects) to control for such possibilities. As shown in the next section, this inclusion has little bearing on the estimated coefficients, reinforcing our prior beliefs that such systematic effects, to the extent that they are present, are not likely to be quantitatively important in the current context.

Another scenario that would violate our assumptions is that households may use credit cards in their major shopping trip, and cash in small "top-up" trips, and that the latter shopping is performed close to home. This would mean that our measurement error would be correlated with distance, even after controlling for fixed effects, violating Assumption 1.[31] However, as long as the "top-up" trips primarily take place in the home neighborhood, this issue can be overcome by altering Assumption 1 to condition not only on origin, destination and time fixed effects, but also on the "shopping at home" dummy variable h_{jn}. This will not change our estimated coefficients but would change the interpretation of the "shopping at home" coefficient. Specifically, as with the ν terms, this coefficient would confound the utility effect κ with measurement error.

[30] This argument is related to Figurelli's (2013) point that there is an interaction between the choice of which goods to buy and the choice of store location.

[31] We are grateful to Pierre Dubois for pointing out this possibility.

3.3 Estimation results

Table 6 shows OLS estimates of equation (14) for various specifications. We compute standard errors by 2-way clustering at the origin and destination level, i.e., allowing for arbitrary correlation between observations sharing an origin and/or a destination. We entered the regressors $\ln p_{nt} \cdot x_j\alpha$ and $d_{jn} \cdot x_j\beta$ with a negative sign, as specified in (14), so that the estimates in the table are direct estimates of α and β.

Table 6: Estimates of utility function parameters

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln (price)	8.768	9.283	1.691	1.725	5.065	4.727	1.630	4.730	5.865	4.646
	(5.788)	(5.491)	(.763)	(.749)	(1.421)	(1.304)	(.774)	(1.302)	(1.537)	(1.333)
ln (price) X housing prices					-0.253	-0.232		-0.232	-0.315	-0.228
					(.083)	(.078)		(.078)	(.091)	(.079)
Distance	0.272	0.365	0.197	0.334	0.393	0.423	0.423	0.411	0.471	0.487
	(.049)	(.072)	(.036)	(.045)	(.13)	(.12)	(.12)	(.119)	(.116)	(.107)
Distance X seniors					0.002	0.004	0.004	0.004	0.004	0.006
					(.004)	(.007)	(.007)	(.006)	(.005)	(.007)
Distance X driving to work					-0.002	-0.003	-0.003		-0.003	
					(.002)	(.002)	(.002)		(.001)	
Distance X car ownership								-0.002		-0.003
								(.001)		(.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	YES	YES
Fixed destination effects	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES
Fixed period effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Destination X housing prices	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES
# observations	1819	1819	1819	1819	1819	1819	1819	1819	1819	1819
R2	0.243	0.382	0.657	0.775	0.776	0.784	0.783	0.783	0.762	0.770

Notes: 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. Standard errors in parentheses are (2-way) clustered at the origin and destination levels.

The various specifications in Table 6 capture the effect of three main variables of interest: price, distance, and the "shopping at home" indicator, while controlling for different combinations of fixed effects and allowing for interactions with various socioeconomic characteristics. Across all specifications, "shopping at home" has a positive and highly significant coefficient which is consistent with the relatively high frequency of home neighborhood shopping observed in the data (Table 5). As expected, the coefficients of log price and distance are always positive, consistent with consumer utility declining with higher prices and longer distances. Column (4) corresponds to our model (14) with the three set of dummies (origin, destination and period) but without interacting the main regressors with demographics. All three effects (price, distance and "shopping at home") are strongly significant, even after controlling for the complete set of fixed effects implied by the theory. Interestingly, 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)-(10), and yields higher estimation precision. This is consistent with the unobserved destination characteristics v_n (e.g., availability of parking, opening hours, product variety etc.) being important determinants of consumer utility.

Interaction terms allow price and distance sensitivities to vary with characteristics of the neighborhood of origin, and display intuitive coefficients signs. 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. Retired 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. These effects, however, are only significant when omitting the "shopping at home" dummy variable in columns (9) and (10).

Following discussion in the previous subsection, we control for an interaction term between origin housing prices and destination dummies in column (6). 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. Columns (7)–(10) present variations of the baseline specification. Omitting the interaction of price with housing prices in column (7) generates the same marginal effect of log price as that from column (6) evaluated at the mean housing price. Columns (8) – (10) present additional results using "car ownership" instead of "driving to work" and omitting the "shopping at home" indicator. Overall, estimates in columns (7) – (10) are very close to the baseline specification in column (6).[32]

Demand elasticities. We next examine the quantitative economic implications of the para-

[32] Additional specifications, not reported, show that an interaction of price with family size is not significant and does not alter the other coefficients.

meter estimates via the computation of demand elasticities. Price elasticities are calculated at the store level even though we do not observe store-level demand, since it is these elasticities that are crucial for pricing decisions (see Section 4.1). We therefore calculate the elasticity of demand at store s located in destination n with respect to the price charged at the store, p_{sn}.

Demand for the composite good at store s located in neighborhood n from households residing in neighborhood j is $Q_{jsnt} = \frac{E_{jsnt}}{p_{snt}} = H_j \pi_{jsnt} \frac{\gamma y_j}{p_{snt}}$, where E_{jsnt} is the total expenditure of origin neighborhood j's residents at store s located in neighborhood n and π_{jsnt} is the probability that a resident from origin j shops at the store. Aggregate demand at the store from all origin neighborhoods is $Q_{snt} = \sum_{j=1}^{J} Q_{jsnt}$. The retailer's own price elasticity is therefore

$$\eta_{snt,p} = \frac{p_{snt}}{Q_{snt}} \frac{\partial Q_{snt}}{\partial p_{snt}} = -\sum_{j=1}^{J} \frac{Q_{jsnt}}{Q_{snt}} \left[1 + x_j \alpha \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} \pi_{js|nt} - \pi_{jsnt} \right) \right] \tag{15}$$

where $\pi_{js|n}$ was defined in (4). This elasticity measures the percentage change in demand at store s located in destination n in response to a one percent increase in the composite good's price charged at that store. This is a quantity-weighted average of origin-specific price elasticities.

In what follows we assume within-neighborhood symmetry in the observed equilibrium. Under this assumption, we have that the neighborhood's retailers split the demand equally so that $\pi_{js|nt} = 1/L_n$, and $\pi_{jsnt} = \pi_{jnt}/L_n$ (see (6)). It also follows that $Q_{jsnt}/Q_{snt} = Q_{jnt}/Q_{nt}$, i.e., the fraction of sales at store s that are made to customers arriving from neighborhood j is equal to the fraction of total neighborhood n's sales to origin j's residents.

The semi-elasticity of the neighborhood-level demand Q_{jnt} with respect to the distance between j and n is (imposing within-neighborhood symmetry),

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

This measures the percentage change in demand from residents of neighborhood j at destination n in response to a $1km$ increase in the distance between neighborhoods j and n (for $j \neq n$).

Examining the elasticity terms, the actual computation of these elasticities requires an estimate of σ, and data on the number of stores L_n, in addition to the choice probabilities π_{jnt} (noting that $\pi_{js|nt} = 1/L_n$, and that $Q_{jsnt}/Q_{snt} = Q_{jnt}/Q_{nt}$ are known). As discussed above, σ is not identified with the data at hand. Since this parameter determines the extent of within-neighborhood competition and, therefore, the equilibrium markups, one approach is to pick a value of σ that yields reasonable markups.[33] Based on conversations with people familiar with

[33] Such an approach has some precedence in the literature. BV (2016), for instance, calibrate a parameter τ that governs the degree to which firms consider rival profits in their own profit function (where a value of 1 corresponds to perfect collusion) to generate reasonable markups.

the industry, retail markups of 15-25 percent are reasonable for the type of products studied in this paper.[34] As shown in Section 4.1 where we describe the supply-side model and derive the markups, setting $\sigma = 0.7$ yields an average (median) markup of 22 (20) percent and therefore this is the value chosen for σ. As a sensitivity check, we also estimate elasticities and markups, and conduct counterfactual experiments given an alternative value of $\sigma = 0.8$. As we report below, this robustness check has no impact on the paper's findings.[35]

We next turn to the measurement of the number of competitors in destination n, L_n. We define it as the number of supermarkets operating in neighborhood n in 2008. We therefore do not count grocery stores and other non-supermarket retail establishments. This definition is driven by our view of the retail competition studied in this paper. Small grocery stores are not a close substitute to supermarkets in the context of a households' main weekly shopping trip (e.g., because of limited availability of items). To have a well-defined measure of within-neighborhood competition, we therefore count supermarkets only. In order to partially take into account the role played by additional retail formats, we modify the definition of L_n to equal the number of supermarkets *plus 1* within residential destinations, while keeping it equal to the number of supermarkets in the commercial areas. This modification results in estimated margins that we view as more reasonable, and has a negligible effect on the qualitative findings of the counterfactual analyses reported in Section 4.2. The number of supermarkets is shown in the last column of Table 3.[36]

The last element required for the computation of the elasticities are the choice probabilities π_{jnt}. In typical applications, these probabilities are simply equated to the observed market (or, in our case, expenditure) shares via the market share equation given the observed equilibrium, and can be easily computed from the formulae above given any counterfactual scenario. In our application, this need not be the case due to the measurement error and the fact that the

[34] 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.

[35] Given the structure of the destination fixed effect ϕ_n, namely that it equals sum of the utility terms $v_n + (1-\sigma)\ln L_n$ from equation (13) and the linear projection of w_{jnt} on v_n, we could regress the estimated fixed effects $\hat{\phi}_n$ on $\ln L_n$ to estimate $1-\sigma$. When we do this we get an estimate of $\hat{\sigma} = 0.81$, imprecisely estimated because of the small number of observations. This has some similarities to the minimum distance procedure in Nevo (2001). This estimate, however, 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.

[36] The number of supermarkets per neighborhood was provided by the ICBS and includes supermarkets which were not included in the price sample (e.g., in Pisgat Zeev north). A specific issue arises with respect to the open market of Mahane Yehuda where many small sellers – open stands – are present. Absent clear theoretical guidance on how to make this number comparable to the numbers of supermarkets in other locations, we choose to set $L_n = 2$ in that location (because there is a small supermarket in the neighborhood). Using different values has, of course, an immediate impact on the margins implied for this specific neighborhood. However, it makes no difference for the qualitative findings of the paper.

estimated fixed effects (ϕ) confound the utility fixed effects (ν) with measurement error effects. As a consequence, even though the parameters α, β, κ are consistently estimated given Assumption 1, the mean utility levels δ are not identified, and hence, neither are the choice probabilities, absent additional assumptions.

To tackle this issue, we first note that, using (8) and (12), observed expenditure shares s_{jnt}^{CC} satisfy:

$$s_{jnt}^{CC} = \frac{\tilde{E}_{jnt}^{cc}}{\sum_{m=0}^{N} \tilde{E}_{jmt}^{cc}} = \frac{\left(\frac{\tau_{jnt}}{\lambda_{jnt}}\right) E_{jnt}}{\sum_{m=0}^{N} (\frac{\tau_{jmt}}{\lambda_{jmt}}) E_{jmt}} = \frac{(\frac{\tau_{jnt}}{\lambda_{jnt}}) H_j \pi_{jnt} \gamma y_j}{\sum_{m=0}^{N} (\frac{\tau_{jmt}}{\lambda_{jmt}}) H_j \pi_{jmt} \gamma y_j} = \frac{(\frac{\tau_{jnt}}{\lambda_{jnt}}) \pi_{jnt}}{\sum_{m=0}^{N} (\frac{\tau_{jmt}}{\lambda_{jmt}}) \pi_{jmt}}$$

Examining the above expression we note that if, for any fixed origin neighborhood j, the ratio $(\tau_{jnt}/\lambda_{jnt})$ is constant across destinations n, then these ratios cancel out and we get that the observed credit-card expenditure share s_{jnt}^{CC} is equal to the choice probability π_{jnt},

$$s_{jnt}^{CC} = \frac{\pi_{jnt}}{\sum_{m=0}^{N} \pi_{jmt}} = \pi_{jnt} \tag{16}$$

We therefore proceed by imposing the following assumption:

Assumption 2. *The ratio τ/λ is fixed over origin-destination pairs, i.e., $(\frac{\tau_{jnt}}{\lambda_{jnt}}) = (\frac{\tau_{\ell mt}}{\lambda_{\ell mt}})$ for all $j, \ell \in \mathcal{J}$ and $m, n \in \mathbb{N}$.*

Note that this assumption allows the parameters τ and λ to vary across locations, and only requires their ratio to be equal. Assumption 2 serves two purposes: first, it allows us to estimate the choice probabilities in the observed equilibrium from the observed expenditure shares. This allows us to compute elasticities, and, as we show in the next subsection, the *expected prices* paid by residents of each origin neighborhood. Second, Appendix C shows that Assumption 2 allows us to compute mean utility levels, choice probabilities and markups under counterfactual scenarios, enabling our policy analyses. We stress that Assumption 2 is not required for the consistent estimation of the parameters α, β, κ — Assumption 1 was sufficient for that purpose. In this sense, our framework clarifies the different sets of assumptions that can be used to accomplish different goals in the presence of the measurement error in expenditure data. While this assumption enables the usage of s_{jnt}^{CC} to measure π_{jnt}, it is clearly much weaker than simply assuming that the two are identical (i.e., ignoring the measurement error altogether).

Turning to the estimated elasticities, we report their distribution in Table 7. We employ the leading specification (column 6 of Table 6) and compute price elasticities for each destination,

Table 7: Distribution of estimated elasticities (absolute value)

Own price elasticity										
σ	mean	sd	min	p10	p25	p50	p75	p90	max	N
0.7	4.82	0.92	3.00	3.86	3.99	4.95	5.87	5.95	6.13	15
0.8	6.43	1.37	3.78	5.01	5.31	6.54	7.94	8.32	8.47	15
Distance semi-elasticity										
	mean	sd	min	p10	p25	p50	p75	p90	max	N
	0.35	0.06	0.06	0.28	0.31	0.35	0.39	0.42	0.45	690

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. Price elasticities are computed for each destination. Distance semi-elasticities computed for each origin-destination pair.

and distance semi-elasticities for each origin-destination pair. We present estimates for the last period, November 2008, and those are nearly identical to the average over the three periods.

The average (median) store-level own *price elasticity* $\eta_{snt,p}$ 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 neighborhood, this relatively-elastic demand seems reasonable. Moreover, expenditures on the composite good represent a non-trivial fraction of the household budget and this tends to make households relatively more responsive to prices. Note also that the elasticity is sensitive to the choice of σ: increasing σ to 0.8 generates a higher mean price elasticity of 6.43. At the same time, this modification makes no difference in terms of the qualitative findings of the paper. The average *distance semi-elasticity* $\eta_{jnt,d}$ is 0.35 in absolute value (the median is also 0.35) 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, on average. This suggests a substantial scope for spatial competition and is consistent with anecdotal evidence surveyed in the Introduction.

Our estimated model provides another way of assessing the price-distance trade-off. One may formulate this trade-off as follows: consider residents of location j who shop at location n. What is the highest price increase these consumers are willing to accept for destination n to become closer (in travel time) to their location by (the equivalent of) $1km$? Examining the utility function, one easily observes that the percentage change in prices that keeps their utility unchanged (after the decrease in distance) is $100\left(\exp\left(\frac{x_j\beta}{x_j\alpha}\right)-1\right)$ whose median value over the

46 origin neighborhoods is 24.5 and is indicative of a substantial spatial dimension in households' preferences.

Finally, we note that in Appendix F we present robustness checks for both the demand estimates and the implied elasticities under alternative computations for the price of the composite good, as motivated in Section 2.2.

The estimated demand model, combined with a supply-side assumption stated later, will allow us to explore the manner by which policy interventions affect prices and consumer behavior in equilibrium. Before we do this, the next subsection shows how the model is used to compute the expected prices paid by residents of each residential neighborhood.

3.4 Combining prices and shopping patterns: expected prices

The model above defined the probability that a resident from neighborhood j buys the composite good in neighborhood n (at any of its stores), π_{jn}. As discussed above, the model's assumptions imply that these probabilities can be estimated directly from the observed expenditure shares using (16). These probabilities are identical for all households residing in neighborhood j and describe their shopping patterns across Jerusalem's neighborhoods. In particular, π_{jj} is the probability of shopping in the home neighborhood and Table 5 indicates that $\pi_{jj} < 1$ for every neighborhood j. It follows that the observed price p_j is not the only price faced by households in neighborhood j. Different households (from the same neighborhood) end up buying in different locations because of the idiosyncratic terms in their utility function. We combine these probabilities and observed prices into an *expected price* for residents of neighborhood j:

$$p_j^E \equiv \sum_{m=0}^{N} \pi_{jm} p_m \tag{17}$$

This expected price weights the price in each of the destination neighborhoods by the probability that a resident from neighborhood j shops there. It is therefore interpreted as the cost of grocery shopping incurred by a random resident of origin neighborhood j. In order to compute p_j^E, and compare it to p_j, the price charged by retailers operating at neighborhood j, we use the observed expenditure shares to estimate the π_{jm} terms. As for prices, we observe p_m at each of the 15 neighborhoods where valid price data was collected, but we also need to know the price charged to households who shop at the outside option — the 31 neighborhoods without valid observed prices — labeled $m = 0$. The price at the outside option p_0 is, of course, not observed. Since the "outside option neighborhoods" are residential neighborhoods where we believe most shopping opportunities are at expensive grocery stores rather than at low price supermarket chains, we set p_0 as the price charged in Qiryat HaYovel south (e.g., $p_0 = 8.19$ in November

2008). This is one of the three peripheral, non-affluent neighborhoods discussed above, and it is also the neighborhood that launched the consumer boycott in January 2014.

Figure 5 plots the expected price against housing prices in each of the 46 neighborhoods in November 2008 (along with a linear predicted line). Only selected neighborhoods are labeled. In 8 out of the 11 residential neighborhoods with valid prices, the expected price is substantially lower than the observed price.[37] This reflects the fact that households engage in cross-neighborhood shopping, in part for the purpose of reducing their costs. The price dispersion of the expected price is lower: the standard deviation of the observed price is 0.52 while that of the expected price is 0.34, though of course the latter is computed with more price observations (15 and 46 observations, respectively).

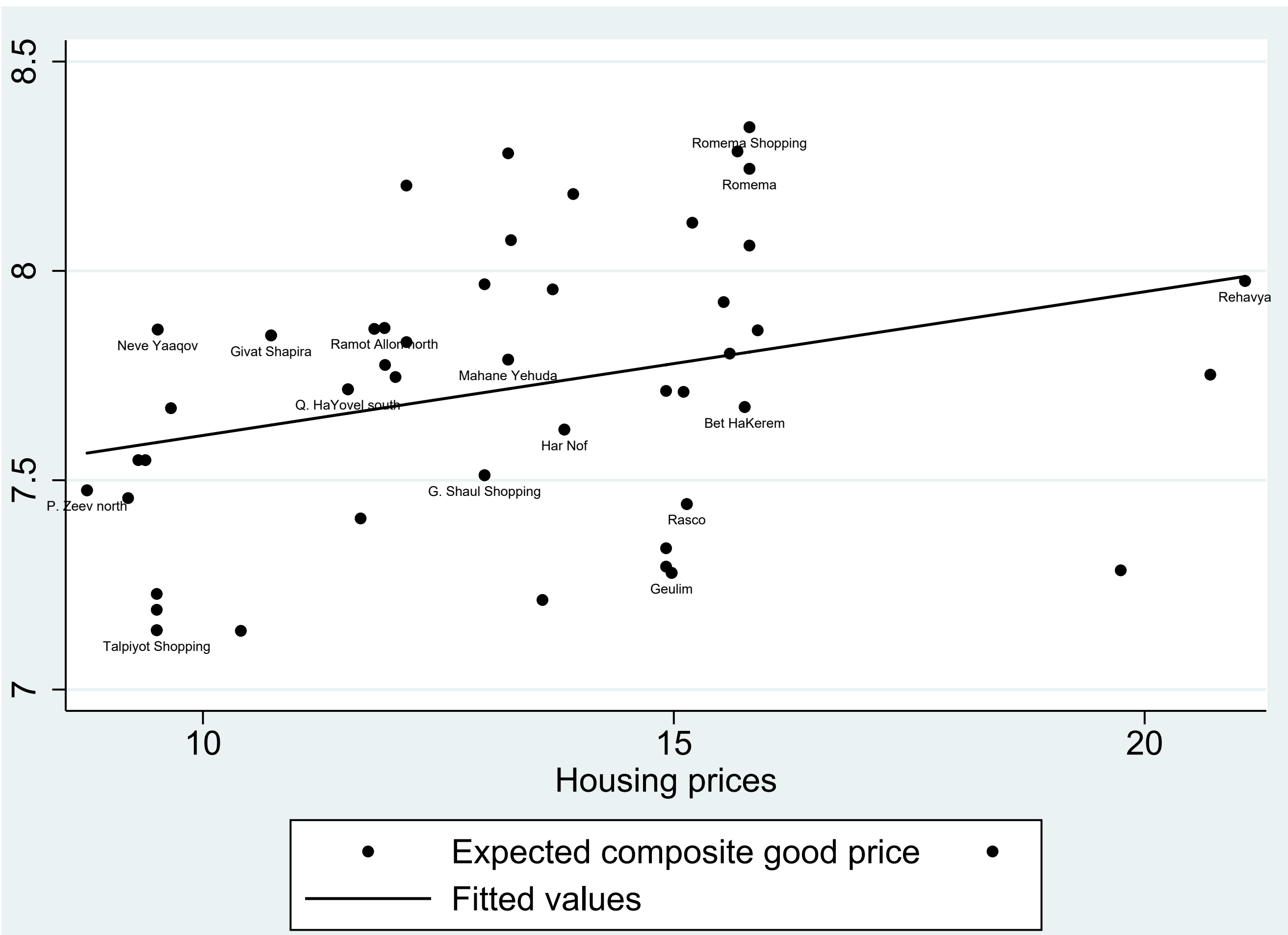

Figure 5: Observed prices and expected prices plotted against housing prices, November 2008

Furthermore, it is of interest to compare p_j^E with the minimum price across all 15 neighborhood

[37] When p_j^E is higher than p_j (in Pisgat Zeev north, Ramot Allon north and Romema) the difference is small.

$p_{\min} = Min_n(p_n)$. This minimum price would have been the price actually paid if households were to determine their shopping destination based on price only (ignoring equilibrium effects). The expected price is, on average, 12.2 percent higher than $p_{\min}$ (the range being between 3.7 and 21.2 percent). This reflects the monetary value of spatial frictions faced by households (i.e., $\beta \neq 0$ and $\kappa \neq 0$) as well as their preferences for specific shopping destinations (v_n and the idiosyncratic terms), and it gives a rough indicator of the extent to which prices can be expected to decline were these frictions to be removed – an analysis to which we return below.

Finally, and importantly, the expected prices at the peripheral, non-affluent neighborhoods (Qiryat HaYovel south, Givat Shapira and Neve Yaaqov) are higher than those faced by residents of more affluent neighborhoods that are located closer to the commercial areas such as Geulim (Baqa) or Bet Hakerem. This suggests examining the relationship between the expected price p_j^E and distance to the main commercial area Talpiot, $d_{jTalpiot}$. This is plotted in Figure 6 (along with a linear predicted line) which shows a strong positive relationship between households' distance to this main commercial center, and the expected price they pay. This is yet another manifestation of the role played by spatial frictions in determining the cost of grocery shopping incurred by residents of various neighborhoods.

4 Spatial differentiation and pricing

In this section we introduce a model of equilibrium pricing decisions in which retailers located across the city's neighborhoods simultaneously choose prices in a differentiated Nash-Bertrand fashion. The first subsection presents this pricing model and derives the price-cost margins implied by this model and the estimated demand system. The second subsection performs counterfactual exercises.

4.1 A pricing model and implied margins

In line with the assumptions of the demand model (Section 3.1), L_n symmetrically-differentiated retailers are present in each destination n, where $n = 1, ..., N$. The L_n retailers within neighborhood n have the same marginal cost c_n. Retailers in the entire city engage in a simultaneous pricing game resulting in a Nash-Bertrand equilibrium. In equilibrium, each retailer's price maximizes her profits given the prices charged by all other retailers in the city: those located within the same neighborhood, and those located in other neighborhoods. Given rival prices p_{-sn}, the price p_{sn} charged by retailer s in destination neighborhood n maximizes the profit function,

$$\Pi_{sn} = (p_{sn} - c_n)Q_{sn}(p_{sn}; p_{-sn})$$

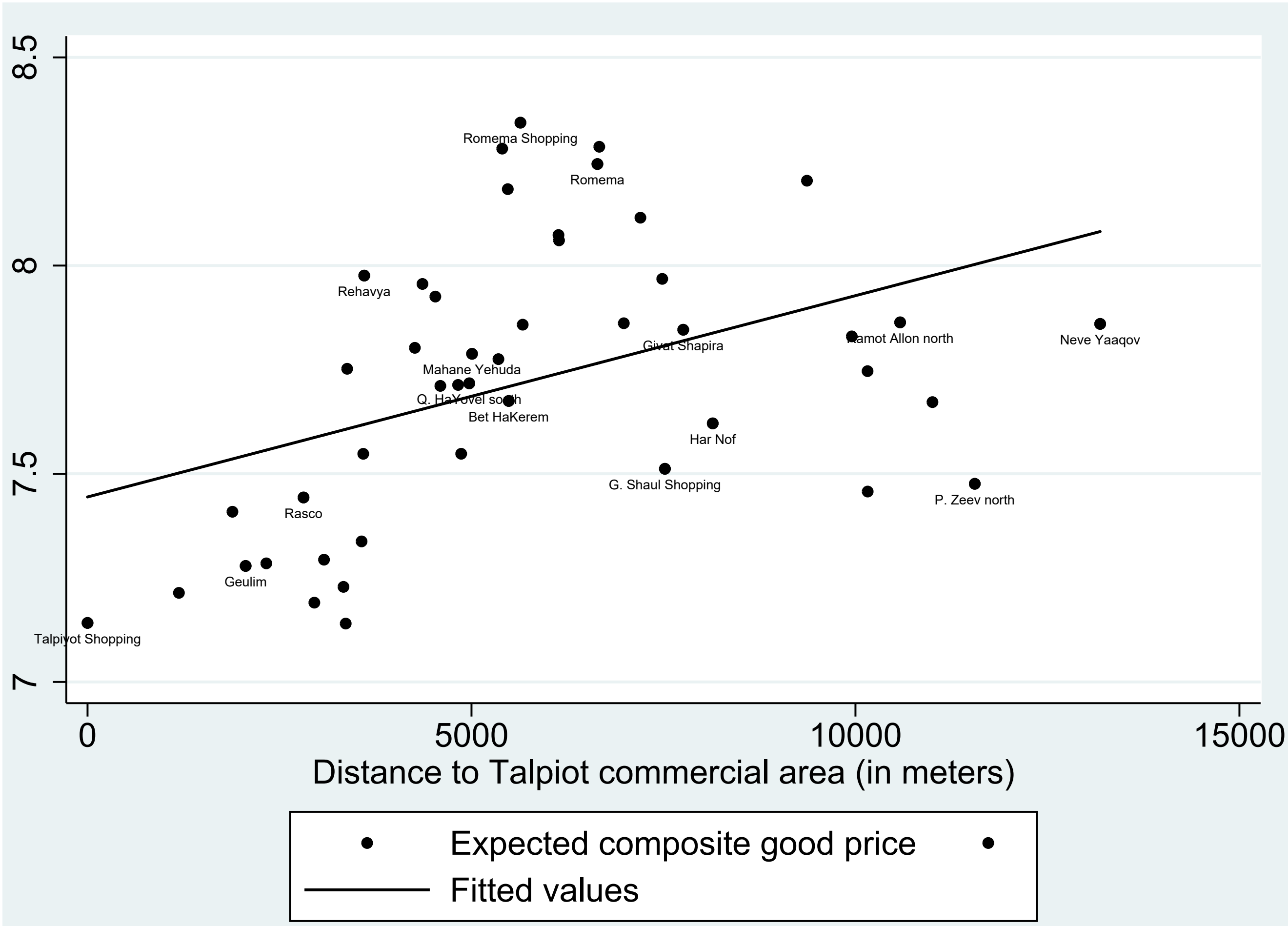

Figure 6: Expected prices plotted against distance to Talpiot, November 2008

where $Q_{sn} = \sum_{j=1}^{J} Q_{jsn}$ is the total quantity sold by retailer s in neighborhood n, obtained by summing over all the origin neighborhoods j from which customers arrive at this retailer. Rearranging yields the familiar inverse elasticity formula for the equilibrium margins,

$$\frac{p_{sn} - c_n}{p_{sn}} = -\frac{1}{\eta_{sn,p}} = \frac{1}{\sum_{j=1}^{N} \frac{Q_{jsn}}{Q_{sn}} \left[1 + x_j \alpha \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} \pi_{js|n} - \pi_{jsn}\right)\right]} \tag{18}$$

where the last equality follows from (15).

We follow the literature by assuming the existence of a unique interior Nash equilibrium in prices.[38] We further assume that the unique pricing equilibrium satisfies within-neighborhood symmetry, a natural assumption given the assumed symmetry of the non-price components of

[38] Caplin and Nalebuff (1991) demonstrate such uniqueness under stronger conditions than those imposed here. See also Nocke and Schutz (2015).

mean-utility levels.[39] In the observed equilibrium, therefore, stores within the neighborhood charge an identical price (equal to the measured neighborhood price p_n in (1)), provide identical mean utility levels, and garner identical market shares. It follows that when exploring the observed equilibrium, we use (6) to replace $\pi_{js|n}$ by $1/L_n$, and π_{jsn} by π_{jn}/L_n. It also follows that $\frac{Q_{jsn}}{Q_{sn}} = \frac{Q_{jn}}{Q_n}$. The role of within-neighborhood competition is clear: higher values of L_n are associated with lower markups, and the magnitude of this effect depends on the parameter σ: the derivative of the margin with respect to σ is negative (as long as $L_n > 1$). This is intuitive given that higher values of σ imply greater substitutability of stores within a neighborhood.

Margins are affected by spatial differentiation as reflected in the interactions between shopping probabilities and demographics, and by within-neighborhood competition captured by the number of retailers. The margin at destination n increases in π_{jn} because a larger π_{jn} reflects higher preferences for shopping at n by neighborhood j residents. This effect is mediated via demographics: the effect of a high π_{jn} is stronger, the higher is the sensitivity of residents of j to price, reflected in a high value of $x_j\alpha$. The share of sales by retailers located at n to households from neighborhood j, Q_{jn}/Q_n, also matters for these retailers' margin. In residential neighborhood n, the term Q_{nn}/Q_n – the fraction of the sales by retailers located at n made to residents of the same neighborhood – is usually large and its associated expression $\left[1 + x_n\alpha\left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma}(1/L_n) - (\pi_{nn}/L_n)\right)\right]$ will be dominant in determining the margin at n. If n is an affluent residential neighborhood with high housing prices, the price sensitivity $x_n\alpha$ will be small, operating in the direction of increasing the margin.

Table 8 displays the estimated costs and margins by neighborhood in November 2008, noting that very similar quantitative and qualitative patterns obtain when averaging over the three time periods. We present results for the third time period mainly because this is the time period in which we conduct the counterfactual analyses reported below, and it is instructive to report costs and margins that correspond exactly to the counterfactual experiment.

Using the baseline value $\sigma = 0.7$, the average and median estimated margins are 22 and 20 percent, respectively. 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 discussion in Section 3.2). We also compute margins assuming $\sigma = 0.8$, generating somewhat lower margins given the higher substitutability among stores within neighborhoods. As expected, margins in residential neighborhoods are generally higher than those in the large commercial areas of Talpiot and Givat

[39] When generating counterfactuals we will compute such an equilibrium at the estimated parameter values. The role of this assumption is to rule out other equilibria, i.e., equilibria that do not respect the within-neighborhood symmetry property.

Table 8: Estimated costs and margins

Retail location		sigma=0.7		sigma=0.8	
	p	c	(p-c)/p	c	(p-c)/p
Neve Yaaqov	8.01	6.66	0.17	7.00	0.13
Pisgat Zeev North	7.36	6.10	0.17	6.44	0.12
Ramot Allon north	7.61	6.07	0.20	6.44	0.15
Giv'at Shapira	8.14	6.81	0.16	7.18	0.12
Rehavya	8.52	5.69	0.33	6.27	0.26
Romema	8.17	6.12	0.25	6.59	0.19
Har Nof	7.62	5.74	0.25	6.19	0.19
Qiryat Moshe, Bet Ha-Kerem	7.85	5.88	0.25	6.37	0.19
Qiryat Ha-Yovel south	8.19	6.47	0.21	6.88	0.16
Rassco, Giv'at Mordekhay	7.87	5.84	0.26	6.30	0.20
Geulim (Baqa)	7.76	6.23	0.20	6.59	0.15
Talpiot shopping area	6.89	5.73	0.17	6.06	0.12
Givat Shaul shopping area	7.07	5.65	0.20	6.02	0.15
Romema shopping area	8.69	7.00	0.19	7.44	0.14
Mahane Yehuda	7.20	5.63	0.22	5.99	0.17
Average	7.80	6.11	0.22	6.52	0.16
Median	7.85	6.07	0.20	6.44	0.15

Notes: The table reports the composite good price (p), marginal cost (c), and price-cost margin in each destination neighborhood in which prices are observed in November 2008, our third sample period. Costs and margins are reported under two alternative values for the correlation parameter sigma. Shopping areas appear in bold type.

Shaul. Our model attributes this to both spatial differentiation across neighborhoods and to low within-neighborhood competition in residential areas.

A limitation of our supply model is that we ignore multi-store pricing. This stems from the fact that our expenditure data are at the neighborhood level, rather than at the store level. This motivated our assumption of symmetric differentiation within the neighborhood. Simply put, we cannot treat supermarkets within a neighborhood as being systematically different (in terms of size, chain ownership or otherwise). Nonetheless, many of the supermarkets in residential neighborhoods are part of a chain that also operates supermarkets in the commercial areas. This could, in principle, result in a motivation to raise prices in the residential neighborhoods that is absent from our model.[40]

During the sample period, neighborhood supermarkets were mainly operated by two chains:

[40] For example, the supermarket located at the residential neighborhood of Qiryat HaYovel is owned by a chain operating a supermarket in the popular commercial area of Talpiot. Price setting at the chain level would consider the cross-elasticity between stores: raising prices in Qiryat HaYovel would drive some consumers to shop in Talpiot, and some of those sales in Talpiot would be garnered by the same chain.

"Shufersal" and "Mega." Those chains also operate some supermarkets in the commercial centers. However, the prominent supermarkets in the comercial areas are operated by hard-dicount chains (notably, "Rami Levy"). These hard discounters are absent from the residential neighborhoods. Moreover, the commercial area supermarkets of "Shufersal" and "Mega" operate under hard-discount formats and their pricing is strongly constrained by the low prices charged by "Rami Levy."

As the commercial areas are largely dominated by hard discount chains that dictate low prices there, while being virtually absent from the residential neighborhoods, we view the competitive arena as largely reflecting spatial competition between the hard discount stores located in the commercial areas, and the more standard stores located in the residential neighborhoods. In light of this market structure, our model should still provide reasonable predictions notwithstanding its limitations.

4.2 Policy interventions

We use our pricing model, along with the estimated demand system, to conduct three counterfactual exercises. In these exercises, we examine the role played by various aspects of the competitive environment in generating the city's price equilibrium. Our first scenario involves an improvement in the transportation system that reduces the utility cost of travel within the city. A second scenario improves the unobserved aspects (ν_n in the terminology of our demand model) of shopping at the major commercial areas. Finally, we consider an increase in within-neighborhood competition via the entry of additional supermarkets into residential neighborhoods. Intuitively, all three scenarios operate in the direction of enhancing competition and lowering prices. We shall examine the impact on the prices charged in equilibrium, and on the shopping patterns, i.e., the probabilities with which residents of each origin neighborhood shop at each destination. Combining the two will inform us regarding the impact of the interventions on the *expected prices*.[41] Following a succinct explanation of some technical aspects, we discuss each scenario in turn, and then summarize the combined takeaway from all three.

Computation. We solve for counterfactual price equilibria, focusing on equilibria that satisfy within-neighborhood price symmetry. It follows that the pricing equilibrium is characterized by a system of first-order conditions, containing one "representative" first-order condition per destination neighborhood. This is the FOC that characterizes the optimal pricing decision of a representative retailer in the neighborhood, as defined in (18). It is convenient to organize the FOCs in vector form:

[41] The structural model allows us to compute the impact on welfare but since our interventions directly affect utility parameters we find this less appealing.

$$(p-c)\bullet d(p)=p \tag{19}$$

where $\bullet$ represents element-by-element multiplication and d is a vector defined by

$$d(p)=\begin{bmatrix} \sum_{j=1}^{J}\frac{Q_{j1}}{Q_1}\left[1+x_j\alpha\left(\frac{1}{1-\sigma}-\frac{\sigma}{1-\sigma}(1/L_1)-\pi_{j1}/L_1\right)\right] \\ \sum_{j=1}^{J}\frac{Q_{j2}}{Q_2}\left[1+x_j\alpha\left(\frac{1}{1-\sigma}-\frac{\sigma}{1-\sigma}(1/L_2)-\pi_{j2}/L_2\right)\right] \\ \vdots \\ \sum_{j=1}^{J}\frac{Q_{jN}}{Q_N}\left[1+x_j\alpha\left(\frac{1}{1-\sigma}-\frac{\sigma}{1-\sigma}(1/L_N)-\pi_{jN}/L_N\right)\right] \end{bmatrix}$$

The system of equations in (19) is solved by the price equilibrium vector p (assumed to be unique per discussion above). We used the baseline estimates from column 6 in Table 6 and $\sigma = 0.7$. Appendix E reports counterfactual results using the value $\sigma = 0.8$, delivering very similar results. In each counterfactual experiment, we vary the relevant primitives and then compute the vector p that solves (19), i.e., the counterfactual price equilibrium. Additional technical details on computation of the left hand side of (19) are available in Appendix C.[42]

Scenario 1: reducing the disutility from travel. We conduct two experiments. In the first experiment we reduce by 50 percent the utility cost associated with traveling d_{jn} kilometers for all origins and destinations (j,n). This scenario can be best thought of as a somewhat radical improvement in the transportation infrastructure. Examples may include improving public transportation or the roads. In practice, we add $0.5d_{jn}x_j\beta$ to the utility garnered by residents of each origin j from traveling to each destination n. One interpretation is that travel *time* is halved, but other interpretations are also possible. For example, it could be that travel becomes more pleasant in addition to a reduction in travel time.

In the second experiment, we again reduce the disutility from travel by 50 percent as above and, in addition, reduce the utility boost of shopping at home κ by half: that is, we subtract 0.5κ from the utility of shopping at one's home neighborhood. An example of a policy that may reduce κ is the deployment of mass-transportation systems that connect residential neighborhoods with the rest of the city, significantly reducing the need to use a private automobile to shop outside the neighborhood.[43] This may reduce the "fixed cost" of shopping outside the neighborhood associated, say, with giving up a convenient parking space close to home.

[42] The counterfactual analyses must deal with the same issue discussed above in the context of estimating elasticities and markups in the observed equilibrium: the fact that we estimate the fixed effects ϕ that confound measurement error components with the utility fixed effects v. Assumption 2 once again allows us to overcome this issue. Appendix C provides complete details.

[43] A very conveneient light rail system which stops at the Mahane Yehuda open market started operating in Jerusalem in August 2011.

Scenario 2: improving the shopping experience at the commercial centers. This scenario involves improving the destination fixed effects ν_n associated with the city's main commercial centers. These fixed effects capture many aspects of the shopping experience that are unobserved to the econometrician, but here we wish to consider an increase in ν_n resulting from a policy change. For instance, the city may improve the physical infrastructure at destination n by making it cleaner and more pleasant, by working together with local businesses to improve parking availability and convenience and so forth. Boosting the utility of shopping at n would induce more consumers to shop there, potentially making the market more competitive. Our goal is to investigate the implications of such improvements for the city's price equilibrium.

Some institutional details that motivate this exercise stem from the casual observation that the utility cost of traveling to the commercial areas in Jerusalem extends beyond the fixed cost of leaving the home neighborhood (captured by κ) and the cost of traveling d_{jn} kilometers: it also involves the experience that shoppers face upon arrival at the major commercial areas. These areas (Talpiot, Givat Shaul and the open market at Mahane Yehuda) are non-residential neighborhoods characterized by highly-congested traffic and limited parking. Consumers arriving at those commercial areas incur substantial time loss and inconvenience navigating through these neighborhoods, whether by public transportation or by private automobiles. The entry points into these commercial areas are also highly congested, so that shoppers experience substantial time loss before they can actually access the supermarkets. The improvement in ν_n, the destination fixed effect, could result from setting up large parking spaces at the entry points to the commercial area with a convenient shuttle service into the heart of the area. Interestingly, the city of Jerusalem recently announced plans to improve the Talpiot shopping area exactly along these lines.[44] We consider two experiments: one where only the utility of shopping at Talpiot is increased, and another one where the utility of shopping at all three major commercial areas – Talpiot, Givat Shaul and the open market at Mahane Yehuda – increases.

Since ν_n is measured in utiles that have no cardinal meaning we consider scenarios in which this term is increased by one standard deviation. Moreover, given that ν_n is unidentified due to measurement error, we compute the standard deviation of ϕ_n, the fixed effect that confounds the effect of ν_n with the measurement error effect, and add it to the mean utility of shopping in destination n. Naturally, 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, will have little bearing on the qualitative findings, as we discuss below.

Scenario 3: Intensified intra-neighborhood competition. Here we consider the effect

[44] "The plan: the Talpiot industrial zone expected to undergo a revolution over the next decade," Kol Ha'ir (a local Jerusalem newspaper, April 15th 2016).

of increasing L_n, the number of supermarkets in neighborhood n, by 1 for each residential neighborhood. Such additional entry should, of course, be ideally modeled as endogenous. Here, in contrast, we do not wish to formally study the incentives for such additional entry but rather to quantify the impact of additional entry on the price equilibrium in a way that allows comparison to the effects studied in the previous two scenarios (moreover, see the discussion above regarding the stability of supermarket locations overt time).

Results. Tables 9 and 10 provide our main results. Table 9 reports the impact of the policy counterfactuals on prices, while Table 10 does the same for expected prices. All counterfactuals are performed in our third sample period (November 2008).

Table 9: Percentage change in prices under counterfactual scenarios

Retail location	Observed price	Disutility from travel: Reduced 50%	Disutility from travel: + Reduced κ	Improved amenities: Talpiot only	Improved amenities: Three areas	Additional entry
Neve Yaaqov	8.01	3.3%	4.7%	-0.1%	-0.3%	-2.8%
Pisgat Zeev North	7.36	0.3%	0.7%	-0.9%	-1.1%	-3.4%
Ramot Allon north	7.61	0.2%	0.3%	-0.4%	-0.5%	-3.3%
Giv'at Shapira	8.14	0.4%	0.8%	-0.1%	-0.1%	-1.3%
Rehavya	8.52	-8.2%	-12.0%	-3.1%	-1.5%	-6.8%
Romema	8.17	-1.3%	-2.7%	1.1%	1.6%	-4.4%
Har Nof	7.62	-0.7%	-1.7%	0.0%	-1.2%	-4.3%
Q. Moshe, Bet HaKerem	7.85	-1.3%	-3.7%	0.2%	-0.7%	-1.9%
Qiryat HaYovel south	8.19	-0.5%	-0.5%	-0.6%	-0.8%	-3.5%
Rasco, Givat Mordekhay	7.87	-1.7%	-3.4%	-0.6%	-0.9%	-4.6%
Geulim (Baqa)	7.76	-0.2%	-0.3%	-0.1%	-0.2%	-3.0%
Talpiot shopping	6.89	-0.3%	-0.2%	0.4%	0.1%	0.0%
Givat Shaul shopping	7.07	-1.1%	-1.0%	0.3%	0.3%	0.0%
Romema shopping	8.69	-1.0%	-1.0%	0.4%	0.2%	0.1%
Mahane Yehuda	7.20	-1.5%	-1.4%	0.1%	-0.1%	0.1%
Mean (residential)		-0.9%	-1.6%	-0.4%	-0.5%	-3.6%
Median (residential)		-0.5%	-0.5%	-0.1%	-0.7%	-3.4%

Notes: The table reports the percentage changes in prices charged at locations where prices are observed (11 residential neighborhoods and four commercial areas appearing in bold type) under the various policy interventions, computed at the third time period (November 2008). See text for expanations of each scenario. The last two rows report statistics that are computed over the 11 residential neighborhoods only.

Table 9 presents the percentage change in the prices charged by retailers operating in each of the 15 neighborhoods where prices were observed. Eleven of those destinations are residential neighborhoods, and four are commercial areas that appear in bold type. Under all scenarios, the impact on pricing at the commercial areas is minimal. The bottom two rows provide statistics for the eleven residential neighborhoods, revealing modest price changes. On average across the eleven residential neighborhoods, prices decline by 0.4%-1.6% under the first two scenarios, and by 3.6% under the third scenario that admits additional supermarket entry into the residential neighborhoods. We stress that such additional entry may not be feasible, and may be associated with substantial social opportunity costs due to zoning restrictions and lack of space. A price reduction of about 3.5% does not appear as a sufficient incentive to incur such costs.

Several additional aspects of the results merit discussion. First, the averages mask a lot of variation in the price response across neighborhoods. For example, prices in the affluent neighborhood of Rehavya decline the most in all three scenarios whereas the price declines in our three "disadvantaged" neighborhoods – Qiryat HaYovel south, Neve Yaaqov and Givat Shapira – are much smaller (recalling that what we mean by this term is that these are peripherally located, non-affluent neighborhoods that pay some of the highest prices in the observed equilibrium).

The latter two neighborhoods actually experience price *increases* under the first scenario, in which the utility costs of travel are reduced. This may seem counterintuitive, as this intervention should exert downward pressure on prices. This pattern, nonetheless, may be explained by changes in the composition of demand faced by the retailers in these neighborhoods. When traveling from the peripheral neighborhoods becomes less costly, the households that continue shopping at those peripheral, expensive destinations are those with very large idiosyncratic shocks favoring shopping there, making the demand faced by retailers located there *less elastic*. The retailer may therefore profitably raise prices rather than reduce them. Less surprising is the result that prices in the commercial areas increase when they are made more attractive via improved amenities.

In sum, the proposed interventions do not appear to reduce prices in a substantial fashion. Table 10, in contrast, presents the impact of the same interventions on the *expected prices* paid by residents of the same eleven residential neighborhoods that appear in Table 9. For completeness, Table D1 in Appendix D shows the impact on expected prices 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 prices displayed in Table 9.

The first column of Table 10 reports the expected price paid by residential neighborhoods at the observed equilibrium (corresponding to the data points in Figure 6). The other columns report the impact of the interventions on the expected price paid by residents of each neighborhood. Several clear patterns emerge. First, the percentage reduction in *expected prices* displayed here is much bigger than the percentage reduction in *prices* displayed in Table 9. Across the first two scenarios (reduced disutility from travel, and improved amenities at the commercial areas), expected prices fall by 2.4%-5.6% (averaged across residential neighborhoods), compared to the average drop in prices of 0.4%-1.6% reported above. Expected prices, therefore, respond much

Table 10: Percentage change in expected prices under counterfactual scenarios

Retail location	Observed expected price	Disutility from travel Reduced 50%	+ Reduced κ	Improved amenities Talpiot only	Three areas	Additional entry
Neve Yaaqov	7.86	0.4%	0.0%	-2.2%	-3.4%	-2.6%
Pisgat Zeev North	7.48	-1.5%	-1.5%	-3.2%	-3.7%	-2.7%
Ramot Allon north	7.86	-3.3%	-3.5%	-5.2%	-6.7%	-1.6%
Giv'at Shapira	7.85	-3.5%	-5.5%	-6.6%	-7.3%	-0.7%
Rehavya	7.98	-5.7%	-7.3%	-8.6%	-8.9%	-3.2%
Romema	8.24	-1.8%	-2.2%	-0.8%	-3.1%	-2.7%
Har Nof	7.62	-1.5%	-1.9%	-0.6%	-5.1%	-1.8%
Q. Moshe, Bet HaKerem	7.67	-2.9%	-3.4%	-4.7%	-6.2%	-0.5%
Qiryat HaYovel south	7.72	-3.3%	-4.7%	-7.0%	-7.4%	-1.2%
Rasco, Givat Mordekhay	7.44	-1.8%	-3.2%	-5.4%	-5.6%	-1.6%
Geulim (Baqa)	7.28	-1.1%	-1.2%	-4.1%	-4.2%	-0.3%
Mean		-2.4%	-3.1%	-4.4%	-5.6%	-1.7%
Median		-1.8%	-3.2%	-4.7%	-5.6%	-1.6%

Notes: The table reports the percentage changes in expected prices charged at the same 11 residential neighborhoods displayed in Table 9. See text for detailed explanations of each scenario. All analyses performed for the third time period (November 2008).

more intensely than the equilibrium prices themselves.

Notice that an average reduction of 5.6 percent is quite substantial because, as remarked in Section 3.4, the average difference between the expected price and the minimum price in the observed equilibrium is about 12.2 percent and this can be interpreted as an upper bound to the price effect of the interventions.

The difference between the two analyses stems from the fact that Table 9 considers only the impact on the equilibrium prices charged at the different locations, while the expected prices of Table 10 take into account, in addition, the changes in shopping patterns. This is evident in Figure 7 that compares the probability of shopping at the Talpiot commercial area in the observed equilibrium, to the same probability under the intervention that improves amenities at Talpiot. The probability of shopping at Talpiot increases for residents of all neighborhoods, and substantially more for those located in the periphery. While the price charged at Talpiot increases slightly, it is still low, and, as a consequence, expected prices decline considerably.

Second, benefits to the three disadvantaged neighborhoods are substantial. In the scenario that improves amenities at the Talpiot commercial area, the expected price paid by residents of Qiryat HaYovel — the neighborhood where the boycott took place — drops by a substantial 7%. The price charged by the retailers in that neighborhood dropped by 0.6% only, as shown in Table 9. Expected prices at the other two neighborhoods, Neve Yaaqov and Givat Shapira, drop by 2.2% and 6.6%, respectively, whereas prices charged by the retailers in both of these neighborhoods only drop by 0.1%. Simply put, evaluating this policy intervention in terms of its effect on residents of these neighborhoods would be highly misleading if it considers changes in prices alone. Such an analysis would suggest very mild benefits, if at all. In contrast, the analysis that considers, in addition, the impact on shopping probabilities, embedded into the computation of expected prices, suggests substantial reductions in the cost of grocery shopping incurred by residents of these neighborhoods. As shown above, this point applies to residential neighborhoods in general, and not only to the three disadvantaged ones.

Third, among the various scenarios, scenario 2 is the one that brings the most substantial benefits in terms of reducing the expected prices: they drop, on average, by 4.4% and 5.6% given improvement in amenities at Talpiot only, and at the three major commercial areas, respectively. These are substantial average gains, and, as we saw, the gains to the disadvantaged neighborhoods are particularly high. This is interesting because, among the three scenarios, this second scenario is the one that seems to be the least costly. Unlike the third scenario, it does not require admitting additional supermarkets into the residential neighborhoods (to the extent that this is even possible). Unlike the first scenario, it does not require a major improvement of the city's transportation infrastructure, an endeavour that may be extremely expensive.

Finally, as noted above, we added one standard deviation of the distribution of the estimated fixed effects ϕ_n to the utility of shopping at the commercial areas which may be greater than a one standard deviation of the distribution of the utility fixed effects ν_n. This issue, however, does not drive our findings. For example, if we repeat the experiment that improves amenities at Talpiot by adding one half of a standard deviation of ϕ_n to the utility of shopping at this commercial center, we obtain that the average drop in expected prices across the 46 origins is 2.3%, whereas the average drop in prices across the 15 destinations is only 0.1%. In other words, the notion that expected prices are affected much more than prices themselves still obtains, regardless of this issue.

These results exemplify the usefulness of a quantitative analysis of equilibrium relationships for policy recommendations as opposed to a partial-equilibrium ("holding all other things constant"), and usually qualitative, analysis.[45] First, although prices were expected to decline they did not decline by much (and even increased in some instances). Second, shopping patterns changed substantially suggesting that the cost of living can be reduced by interventions that facilitate consumers' ability to access low-price stores, *even if prices across the city do not change by much.*

[45] 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. We have ignored this feedback effect in our analysis as it requires a model of retailers' choice of amenities which is beyond the scope of this paper.

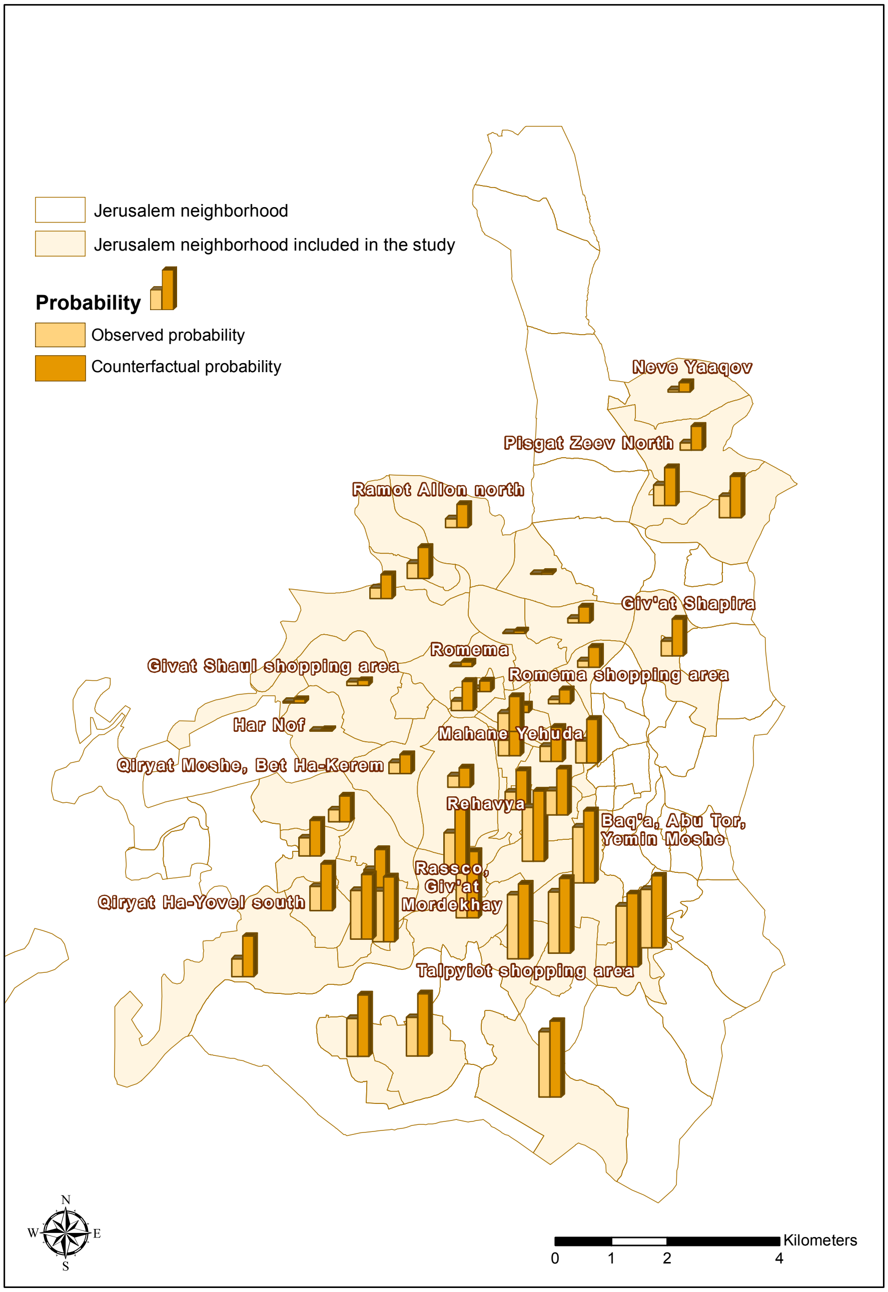

Figure 7: Observed vs. Counterfactual (given improved amenities at Talpiot) probability of shopping at Talpiot, November 2008

Our results therefore show that assessing a policy intervention by its effect on prices alone would be incomplete if its effect on shopping mobility is ignored.

5 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, to explore the determinants of price differentials and shopping patterns within the city. We document several important patterns: prices in residential neighborhoods are persistently higher than prices in commercial areas. When comparing among residential neighborhoods we find, in general, that retailers at several peripheral, non-affluent neighborhoods charge some of the highest prices in the city. Retailers operating in more affluent neighborhoods display interesting variation: some of them charge very high prices, while others, that are in close proximity to the cheap shopping areas, charge low prices. We establish that spatial frictions play an important role in generating these patterns.

Our framework allows us to examine another measure of the cost of grocery shopping faced by neighborhoods' residents: the expected price paid by a random neighborhood resident. This measure takes into account the probabilities with which residents visit the various shopping destinations across the city. In the observed equilibrium, the expected prices also display the patterns discussed above, i.e., they are higher for neighborhoods that are located at a greater distance from the main shopping areas.

Our policy interventions demonstrate the value of considering both price measures. Interventions that facilitate households' access to the main shopping areas, or make shopping there more attractive, have a rather small effect on the prices charged in equilibrium. The effect on the expected prices, in contrast, is substantial, and is particularly enjoyed by residents of the peripheral, less-affluent neighborhoods. The greatest reduction in the cost of grocery shopping is afforded by the intervention that improves amenities at the commercial areas, which is also the one that is likely to involve the least social costs. We stress that these conclusions would be completely missed by an analysis that considers the impact on prices alone. Our structural model allows for the joint analysis of prices and shopping patterns and their responses to interventions.

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 how to lower the cost of living by facilitating household mobility.

References

[1] Adams, B. and K.R. Williams (2013): "Zone Pricing and Spatial Menu Costs: Evidence from Drywall," *Mimeo*

[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., Mira, P., and H. Roman (2007): "An Estimable Dynamic Model of Entry, Exit, and Growth in Oligopoly Retail Markets," *American Economic Review, Papers and Proceedings* 97(2): 449–454

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

[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] Beresteanu A., Ellickson P.B., and S. Misra (2010): "The Dynamics of Retail Oligopoly," *mimeo*

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

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

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

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

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

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

[13] 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.

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

[15] Davis, R.D., Dingel, J.I., Monras, J., and E. Morales (2016), "Spatial and Social Frictions in the City: Evidence from Yelp," *Mimeo*

[16] Dluhosch, B., and M. C. Burda (2007): "Malls and the City," *Mimeo*

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

[18] Ellickson, P.B., Houghton, S., and C. Timmins (2013): "Estimating network economies in retail chains: a revealed preference approach," *RAND Journal of Economics,* 44(2): 169–193

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

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

[21] Griffith, R., Leibtag, E., Leicester, A., & Nevo, A. (2009). "Consumer shopping behavior: how much do consumers save?" *The Journal of Economic Perspectives*, 23(2), 99-120.

[22] Hay, D. and H. Smith (2005): "Streets, Malls, and Supermarkets" *Journal of Economics & Management Strategy*, 14(1): 29–59

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

[24] Houde, J.-F. (2012): "Spatial differentiation and vertical contracts in retail markets for gasoline," *American Economic Review* (forthcoming)

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

[26] Lach, S. (2002): "Existence And Persistence Of Price Dispersion: An Empirical Analysis," *Review of Economics and Statistics*, 84(3): 433-444

[27] 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.

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

[29] Miller, N.H. and M. Osborne (2014): "Spatial differentiation and price discrimination in the cement industry: evidence fro

[30] Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2), 307-342

[31] 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

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

[33] 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).

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

[35] Seim, K. (2006), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37(3), 619-640

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

[37] Sorensen, A.T. (2000): "Equilibrium price dispersion in retail markets for prescription drugs," *Journal of Political Economy* 108: 833– 850

[38] Stahl, D.O. (1989): "Oligopolistic pricing with sequential consumer search," *American Economic Review* 79: 700-712

[39] Stigler, G.J. (1961): "The Economics of Information," *Journal of Political Economy*, 69:213-225

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

[41] Varian, H.R. (1980): "A model of sales," *American Economic Review* 70: 651–659

A Subquarters and demographics

Neighborhoods are identified with the subquarters defined by the ICBS with some exceptions. ICBS-defined subquarters are distinct sets of statistical areas. The exceptions are 1) the commercial areas (appearing in bold in Table A1 below) that were carved out from existing subquarters as discussed in Section 2.1, and 2) four subquarters that were added to accommodate the expenditure data received from the credit card company. These additional subquarters share some of the statistical areas with other subquarters and are denoted in Table A1 with a star *. Although these four subquarters share the same statistical areas (and therefore the same demographics) they do have different zipcodes and therefore different expenditure data.

Table A1 presents our 46 subquarters (neighborhoods) and provides the statistical areas that are included in each neighborhood. Table A2 provides neighborhood-level statistics referred to in the main text.

Table A1: Composition of residential and commercial neighborhoods

Subquarter (neighborhood)	statistical areas					
Neve Yaaqov	111	112	113	114	115	116	
Pisgat Zeev North	121	122	123	124	125		
Pisgat Zeev East	131	132	133	134	135	136	
Pisgat Ze'ev (north - west & west) *	135	136					
Ramat Shlomo	411	412	413				
Ramot Allon north	421	422	423	424	425	426	
Ramot Allon	431	432	433	434	435	436	
Ramot Allon South *	435						
Har H-hozvim, Sanhedriyya	511	512	513	514	515		
Ramat Eshkol, Giv'at-Mivtar	521	522	523				
Ma'a lot Dafna, Shmuel Ha-navi	531	532	533				
Giv'at Shapira	541	542	543				
Mamila, Morasha	811	812					
Ge'ula, Me'a She'arim	821	822	823	824	825	826	
Makor Baruch, Zichron Moshe	831	832	833	834	835	836	
City Center	841	842	843	844	845	846	847
Nahlaot, Zichronot	851	852	854	855	856	857	858
Rehavya	861	862	863	864			
Romema	911	912	913	915	916		
Giv'at Sha'ul	921	922	923	925			
Har Nof	931	932	933	934			
Qiryat Moshe, Bet Ha-Kerem	1011	1012	1013	1014	1015	1016	
Nayot	1021	1022	1023	1024			
Bayit va-Gan	1031	1032	1033	1034	1035		
Ramat Sharet, Ramat Denya	1041	1042	1043	1044			
Qiryat Ha-Yovel north	1121	1122	1123	1124			
Qiryat Ha-Yovel south	1131	1132	1133	1134			
Qiryat Menahem, Ir Gannim	1141	1142	1143	1144	1145	1146	1147
Manahat slopes *	1147						
Gonen (Qatamon)	1211	1212	1213	1214	1215	1216	1217
Rassco, Giv'at Mordekhay	1221	1222	1223				
German Colony, Gonen (Old Qatamon)	1311	1312	1313	1314			
Qomemiyyut (Talbiya), YMCA Compound	1321	1322					
Ge'ulim, Giv'at Hananya, Yemin Moshe	1331	1332	1333	1334	1335	1336	
Talpiyyot, Arnona, Mekor Hayyim	1341	1342	1343	1344	1346		
East Talpiyyot	1351	1352	1353	1354	1355		
East Talpiyyot (east) *	1355						
Homat Shmuel (Har Homa)	1621	1622	1623				
Gilo east	1631	1632	1633	1634			
Gilo west	1641	1642	1643	1644			
Talpyiot shopping area	1345	Talpiyyot - Industrial & Commercial Area, Yad Haruzim st.					
Givat Shaul shopping area	924	Giv'at Sha'ul Industrial Area, Menuhot Cemetery, Kanfei Nesharim, Giv'at Sha'ul B'					
Malcha shopping center	1146	Teddi Stadium, Biblical zoo, Jerusalem Mall					
Romema shopping area	914	Romema, Industrial Area, Etz Haim, Central Bus Station					
Central Bus Station							
Mahane Yehuda	853	Beit Yaakov, Kelal Centre, Mahane Yehuda Market					

Notes: The table presents our 46 subquarters (neighborhoods), and provides the statistical areas that are included in each neighborhood. For residential neighborhoods, the statistical areas included follow the ICBS definitions. For commercial neighborhoods (bold type), the included statistical areas were determined by the authors and their explicit names are provided. Residential neighborhoods marked with an * mean that the neighborhood shares portions of the same statistical areas with preceding neighborhood. A common statistical area was divided into two subquarters according to the zipcodes of the expenditure data.

Table A2: Demographics, housing prices and number of supermarkets

Subquarter	Population (000s)	Mean household size	Mean housing price	Percentage driving to work	Percentage car ownership	Percentage senior citizens	Number of supermarkets
Neve Yaaqov	18.3	3.9	9.5	21.2	28.6	7.6	2
Pisgat Zeev North	17.7	3.3	8.8	48.3	66.5	10.4	2
Pisgat Zeev East	21.7	3.6	9.7	59.2	73.5	7.6	1
Pisgat Ze'ev east (north - west) & Pisgat Ze'ev west	21.7	3.6	9.2	59.2	73.5	7.6	1
Ramat Shlomo	14.1	6.1	12.2	23.8	35	1.1	1
Ramot Allon north	23.1	4.9	11.9	32.7	39.9	2.5	2
Ramot Allon	16.6	4.1	12.2	51.4	61.3	5.6	1
Ramot Allon South	16.6	4.1	12.0	51.4	61.3	5.6	1
Har H-hozvim, Sanhedriyya, Tel-Arza	15.8	5.3	15.7	9.9	14.7	4.6	1
Ramat Eshkol, Giv'at-Mivtar	10.2	3.9	15.2	27.5	34.4	12.1	1
Ma'a lot Dafna, Shmuel Ha-navi	8.7	4	13.3	17.1	21.8	7	1
Giv'at Shapira	9.3	2.3	10.7	56.3	65.9	10.6	3
Mamila, Morasha	13	3.3	15.6	9.9	12.4	10.7	1
Ge'ula, Me'a She'arim	28.7	4.6	13.9	7.5	6.9	5.9	1
Makor Baruch, Zichron Moshe	13	3.3	13.2	9.9	12.4	10.7	1
City Center	6.2	1.9	13.7	13.6	24	15.4	3
Nahlaot, Zichronot	9.1	2.1	15.5	27.4	35.7	12.5	1
Rehavya	7.5	2	21.1	42.5	57.6	25.6	2
Romema	21.1	4.5	15.8	11.4	10.7	7.5	2
Giv'at Sha'ul	10.5	4.2	13.0	33.8	40.6	7	1
Har Nof	15.8	4.3	13.8	36.1	49.2	6.4	2
Qiryat Moshe, Bet Ha-Kerem	23.3	2.7	15.8	49.8	62.4	16.7	3
Nayot	23.3	2.7	15.1	49.8	62.4	16.7	2
Bayit va-Gan	18.1	3.4	15.9	30.7	39.1	12.3	1
Ramat Sharet, Ramat Denya	8.5	3.3	14.9	68.1	85.4	8.9	1
Qiryat Ha-Yovel north	10.6	2.7	11.9	46	54.6	16.9	1
Qiryat Ha-Yovel south	10.6	2.4	11.5	44.8	49.4	16.3	2
Qiryat Menahem, Ir Gannim	17.5	3.3	11.8	57	62.5	10.2	2
Manahat slopes, Qedoshe Struma st, Ha-Ayal st	17.5	3.3	14.9	57	62.5	10.2	1
Gonen (Qatamon) A - I	23.5	2.8	11.7	39.7	50.7	11.9	1
Rassco, Giv'at Mordekhay	13.5	2.4	15.1	51.5	62.9	14.4	2
German Colony, Gonen (Old Qatamon)	10	2.5	19.7	52	69.6	16.3	1
Qomemiyyut (Talbiya), YMCA Compound	10	2.5	20.7	52	69.6	16.3	1
Baq'a, Abu Tor, Yemin Moshe	11	2.9	15.0	51.7	67	16.4	2
Talpiyyot, Arnona, Mekor Hayyim	13.8	2.8	13.6	55.5	67.9	18	1
East Talpiyyot	13.9	2.9	9.5	55.3	60.8	9.5	1
East Talpiyyot (east)	13.9	2.9	9.5	55.3	60.8	9.5	1
Homat Shmuel (Har Homa)	9.8	4	10.4	66.7	89.3	2.3	1
Gilo east	18.7	3.1	9.4	53.2	65.5	11.6	1
Gilo west	10.4	3.4	9.3	63.7	77.6	8.9	1
Talpyiot shopping area	11	2.9	9.5	51.7	67	16.4	5
Givat Shaul shopping area	10.5	4.2	13.0	33.8	40.6	7	3
Malcha shopping center	17.5	3.3	14.9	57	62.5	10.2	1
Romema shopping area	21.1	4.5	15.8	11.4	10.7	7.5	3
Central Bus Station	21.1	4.5	15.8	11.4	10.7	7.5	0
Mahane Yehuda	13	3.3	13.2	9.9	12.4	10.7	2

Notes: demographic data for the 46 neighborhoods. Commercial neighborhoods appear in bold type and have associated demographics because they also contain a small residential neighborhood. Housing prices = the 2007-2008 average price per square meter. Driving to work = percentage of those aged 15 and over 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 those aged 65+ . Source for demographic variables: http://www1.cbs.gov.il/census/census/pnimi_page_e.html?id_topic=12.

B Products and prices

Table B1: Definition of products

1	**Waffles**	simple packed waffles, non-coated,same brand
2	**Mayonnaise**	low-fat mayonnaise, same brand
3	**Cottage cheese**	250 gr container of same brand
4	**Sugar**	packed sugar, same brand, 1kg
5	**Chocolate bar**	regular milk chocolate, same brand
6	**Mineral water**	in plastic bottle, 1.5 liter
7	**Coca cola**	in plastic bottle, 1.5 liter
8	**Ketchup**	same brand
9	**Tea**	regualr tea, teabags, same brand
10	**Turkish coffee**	packaged roasted and ground turkish coffee, same brand
11	**Cocoa powder**	instant chocolate powder, same brand
12	**Green peas (can)**	garden variety, same brand
13	**Hummus (salad)**	hummus salad, not fresh, same brand
14	**Cucumbers**	fresh standard cucumbers, type A, 1kg
15	**Onion**	dry onion, type A, 1kg
16	**Carrots**	medium size fresh carrots, type A, 1kg
17	**Eggplants**	medium size fresh eggplants, type A, 1kg
18	**Cabbage (white)**	white fresh cabbage, 1kg
19	**Cauliflower**	fresh cauliflower, type A, 1kg
20	**Potatoes**	fresh potatoes, type A, 1kg
21	**Tomatoes**	round tomatoes, type A, 1kg
22	**Apples**	granny smith apples, type A, 1kg
23	**Bananas**	type A, 1 kg
24	**Lemons**	fresh, type A, 1kg
25	**Fabric softener**	same brand
26	**Dishwasher detergent**	in plastic bottle, same brand
27	**Shaving cream/gel**	same brand

Table B2: List of products and their prices (in NIS)

Product	Mean price	Coefficient of Variation	Number of stores
Waffles			
Sep-07	10.4	0.14	24
Nov-07	10.2	0.18	22
Nov-08	11.1	0.24	20
Mayonnaise			
Sep-07	7.6	0.12	22
Nov-07	9	0.21	21
Nov-08	9.6	0.14	16
Cottage cheese			
Sep-07	5.3	0.04	23
Nov-07	5.8	0.03	25
Nov-08	6	0.05	22
Sugar			
Sep-07	3.6	0.22	24
Nov-07	3.6	0.22	23
Nov-08	3.4	0.26	24
Chocolate bar			
Sep-07	4.4	0.11	23
Nov-07	4.5	0.11	23
Nov-08	5.1	0.12	23
Mineral water			
Sep-07	12.8	0.11	21
Nov-07	12.7	0.15	20
Nov-08	12.3	0.28	20
Coca cola			
Sep-07	5.5	0.18	25
Nov-07	5.5	0.18	25
Nov-08	5.9	0.17	24
Ketchup			
Sep-07	11.1	0.14	24
Nov-07	10.9	0.14	24
Nov-08	11	0.15	23
Tea			
Sep-07	15.8	0.15	22
Nov-07	16.2	0.15	23
Nov-08	17.1	0.15	20

Product	Mean price	Coefficient of Variation	Number of stores
Turkish coffee			
Sep-07	5.8	0.09	23
Nov-07	5.7	0.11	23
Nov-08	7	0.07	23
Cocoa powder			
Sep-07	10.3	0.12	23
Nov-07	10.5	0.12	23
Nov-08	10.7	0.11	22
Green peas (can)			
Sep-07	5.2	0.10	16
Nov-07	5.2	0.10	16
Nov-08	5.9	0.12	14
Hummus (salad)			
Sep-07	9	0.11	17
Nov-07	9.2	0.05	18
Nov-08	10.6	0.10	14
Cucumbers			
Sep-07	4.6	0.28	37
Nov-07	5.8	0.17	37
Nov-08	4.8	0.29	35
Onion			
Sep-07	2.8	0.32	37
Nov-07	3.2	0.34	36
Nov-08	3.7	0.35	35
Carrots			
Sep-07	4.9	0.18	37
Nov-07	5.1	0.18	36
Nov-08	5.6	0.38	32
Eggplants			
Sep-07	4	0.40	38
Nov-07	3.7	0.41	35
Nov-08	4.7	0.34	33
Cabbage (white)			
Sep-07	4.7	0.51	33
Nov-07	3.7	0.57	32
Nov-08	5.1	0.61	31

Product	Mean price	Coefficient of Variation	Number of stores
Cauliflower			
Sep-07	7.3	0.32	25
Nov-07	5.9	0.19	22
Nov-08	6.6	0.24	23
Potatoes			
Sep-07	4	0.23	37
Nov-07	4.2	0.26	37
Nov-08	4.8	0.25	35
Tomatoes			
Sep-07	6.1	0.33	37
Nov-07	5	0.34	37
Nov-08	6.9	0.33	35
Apples			
Sep-07	9	0.20	36
Nov-07	9.1	0.12	34
Nov-08	9.6	0.18	33
Bananas			
Sep-07	6.3	0.13	35
Nov-07	5.6	0.30	35
Nov-08	7.8	0.23	33
Lemons			
Sep-07	11.7	0.22	38
Nov-07	8.1	0.25	36
Nov-08	10.4	0.37	35
Fabric softener			
Sep-07	20.8	0.08	21
Nov-07	19.9	0.16	25
Nov-08	22.1	0.07	22
Dishwasher detergent			
Sep-07	10.8	0.12	16
Nov-07	11.9	0.10	19
Nov-08	11.1	0.20	23
Shaving cream/gel			
Sep-07	22.1	0.20	22
Nov-07	23.2	0.22	16
Nov-08	23.5	0.16	18

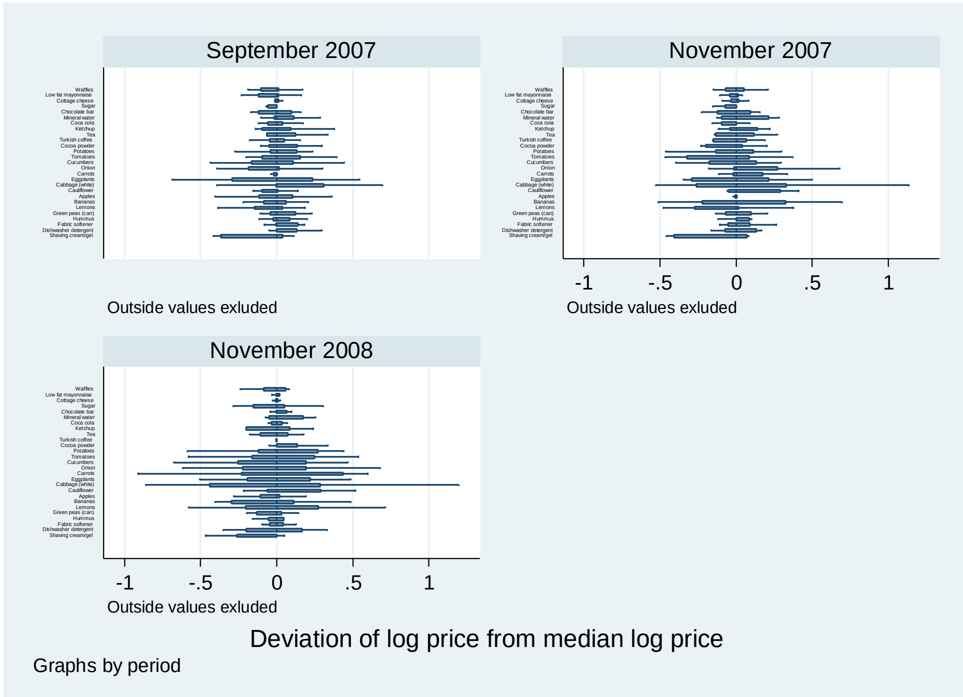

Figure B1: Dispersion of log prices

The "box" starts at the 25th percentile of the log price distribution and ends at the 75th percentile (for expositional clarity, each plot is centered on the product's median log price).

C Computational details on counterfactuals

To perform the counterfactual exercise, one must be able to compute the left hand side of (19), namely $(p - c) \bullet d(p)$ given any price vector p. Computation of $(p - c)$ is, of course, trivial since p is given and c is held fixed during the exercise. The critical task is, therefore, the computation of $d(p)$. Examining the terms inside this vector, we note that x_j (observed data) and α (an estimated parameter) are also held fixed. The terms that need to be calculated are then the choice probabilities $\pi_{jn}(p)$, and the quantities $Q_{jn}(p)/Q_n(p)$ for each j and n. We now explain how these are calculated.

We begin by explaining how to calculate $\pi_{jn}(p)$ for any j, n and a generic value for p. Recall that the model implies equation (5):

$$\pi_{jn}(\mathbf{p}; \theta) = \frac{D_{jn}^{1-\sigma}}{\sum_{m \in \mathbb{N}} D_{jm}^{1-\sigma}}$$

where $\theta = (\alpha, \beta, \kappa, \sigma)$ are the model's parameters, and the term D_{jn} is defined by:

$$D_{jn} = \sum_{s \in n} e^{(\delta_{jsn} + \gamma^{-1} \ln y_j x_j \alpha)/(1-\sigma)}$$

Imposing price symmetry within the neighborhood (which, again, holds by assumption in the observed equilibrium and in any counterfactual equilibrium), we can write

$$D_{jn} = e^{(\gamma^{-1} \ln y_j x_j \alpha)/(1-\sigma)} \cdot L_n \cdot e^{(\delta_{jn})/(1-\sigma)}$$

where, again, L_n denotes the number of symmetric retailers located in neighborhood n, and the symmetric mean utility is

$$\delta_{jn} = \nu_c + \nu_j + \nu_n + hp_j \cdot \nu_n - \ln p_n \cdot x_j \alpha - d_{jn} \cdot x_j \beta + \kappa \cdot h_{jn}$$

The choice probability simplifies to:

$$\pi_{jn}(\mathbf{p}; \theta) = \frac{L_n^{1-\sigma} e^{\delta_{jn}}}{\sum_{m \in \mathbb{N}} L_m^{1-\sigma} e^{\delta_{jm}}} \tag{20}$$

To compute these probabilities in the various counterfactuals we need estimates of the mean utility levels δ_{jn}. While the terms $\ln p_n \cdot x_j \alpha, d_{jn} \cdot x_j \beta$ and $\kappa \cdot h_{jn}$ are known to us given the data, the estimated parameters and the current guess for p, the terms v_c, v_j and v_n are not known to us, since the fixed effects actually used in estimation are the terms ϕ_j, ϕ_n. In other words, unlike typical applications, our treatment of measurement errors implies that our estimation strategy does not deliver estimates that allow the direct computation of the mean utility terms δ_{jn} given any price vector.

Our strategy for dealing with this challenge is as follows: we begin by noting again that, under maintained Assumption 2 – the ratio (τ_{jn}/λ_{jn}) is fixed over all j and n – the choice probabilities in *the observed equilibrium* are equivalent to the observed credit card expenditure shares. We can use this fact, along with the inversion principle from Berry (1994), to calculate the mean utility levels δ_{jn} in the observed equilibrium. We then hold these values, denoted δ_{jn}^{obs}, fixed and calculate the counterfactual level of δ_{jn}, given any price vector p, by $\delta_{jn}(p) = \delta_{jn}^{obs} - x_j \alpha (\ln p_n - \ln p_n^{obs})$. Counterfactuals that change distances or demographics work similarly by appropriately adjusting the observed mean utility levels.

To compute δ_{jn}^{obs} for all j and n, we first recall a result derived in Section 3.2,

$$\ln \left(\frac{E_{jn}}{E_{j0}} \right) = (1 - \sigma) \ln L_n + \delta_{jn}$$

We further note that
$$\frac{E_{jn}}{E_{j0}} = \frac{\tilde{E}^{cc}_{jn}(\lambda_{jn}/\tau_{jn})}{\tilde{E}^{cc}_{j0}(\lambda_{j0}/\tau_{j0})} = \frac{\tilde{E}^{cc}_{jn}}{\tilde{E}^{cc}_{j0}}$$
where the first equality holds by definition, and the second equality follows from Assumption 2. Recall that we rely on Assumption 2 for the computation of elasticities and counterfactuals but not for estimation. We can now obtain an estimate for δ^{obs}_{jn}

$$\delta^{obs}_{jn} = \ln(\tilde{E}^{cc}_{jn}/\tilde{E}^{cc}_{j0}) - (1 - \widehat{\sigma}) \ln L_n$$

where $\widehat{\sigma} = 0.7$ is our estimate for the correlation parameter σ. It is, therefore, easy to calculate δ^{obs}_{jn} for all j and n. This enables, as explained above, the calculation of $\delta_{jn}(p)$ given any price vector, and the calculation of $\pi_{jn}(p)$ then follows easily from (20).

It remains to show how to calculate $Q_{jn}(p)/Q_n(p)$ for each j and n and any price vector p. Note first that $Q_{jn}(p) = H_j \pi_{jn}(p) q_{jn} = H_j \pi_{jn}(p) \gamma y_j / p_n$, and that $Q_n(p) = \sum_{j=1}^{N} Q_{jn}(p)$. As a consequence, we have:

$$Q_{jn}(p)/Q_n(p) = \frac{H_j \pi_{jn}(p) \gamma y_j / p_n}{\sum_{\tau=1}^{N} H_\tau \pi_{\tau n}(p) \gamma y_\tau / p_n} = \frac{\gamma y_j H_j \pi_{jn}(p)}{\sum_{\tau=1}^{N} \gamma y_\tau H_\tau \pi_{\tau n}(p)} \tag{21}$$

We next note that, *in the observed equilibrium*, the following identity holds: $\tilde{E}^{cc}_{jn} = (\tau_{jn}/\lambda_{jn}) E_{jn}$, where $\tilde{E}^{cc}_{jn}$ are the observed credit card expenditures. Substituting in the definition of E_{jn}, we get that $\tilde{E}^{cc}_{jn} = (\tau_{jn}/\lambda_{jn}) H_j e_{jn} = (\tau_{jn}/\lambda_{jn}) H_j \pi^{obs}_{jn} \gamma y_j$, implying that:

$$\gamma y_j H_j = \frac{(\lambda_{jn}/\tau_{jn}) \tilde{E}^{cc}_{jn}}{\pi^{obs}_{jn}}$$

By Assumption 2, the ratio (τ_{jn}/λ_{jn}) is fixed over all j and n. Substituting into (21), we then get:

$$Q_{jn}(p)/Q_n(p) = \frac{\widetilde{M}_{jn} \cdot \pi_{jn}(p)}{\sum_{s=1}^{N} \widetilde{M}_{sn} \cdot \pi_{sn}(p)}$$

where $\widetilde{M}_{jn} = \tilde{E}^{cc}_{jn}/\pi^{obs}_{jn}$.

$\widetilde{M}_{jn}$ is treated as a constant which is easy to calculate since $\tilde{E}^{cc}_{jn}$ is observed and $\pi^{obs}_{jn} = s^{cc}_{jn}$. Since $s^{cc}_{jn} = \tilde{E}^{cc}_{jn} / \sum_{\tau=1}^{N} \tilde{E}^{cc}_{j\tau}$, we finally get that $\widetilde{M}_{jn} = \sum_{\tau=1}^{N} \tilde{E}^{cc}_{j\tau}$. That is, this constant is equal to the total observed expenditures by residents of location j and does not actually vary by n, that is, $\widetilde{M}_{jn} = \widetilde{M}_j = \sum_{\tau=1}^{N} \tilde{E}^{cc}_{j\tau}$. The $\widetilde{M}$ constants are therefore computed from direct data and are held fixed during the iterative process that solves the FOCs. The other terms that appear in $Q_{jn}(p)/Q_n(p)$ are choice probabilities $\pi_{jn}(p)$, and we already explained above how to obtain those given any p. As a consequence, the final form of $d(p)$ is:

$$d(p) = \begin{bmatrix} \sum_{j=1}^{N} \left[\frac{\widetilde{M}_j \cdot \pi_{j1}(p)}{\sum_{s=1}^{N} \widetilde{M}_s \cdot \pi_{s1}(p)} \left[1 + x_j \alpha \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma}(1/L_1) - \pi_{j1}/L_1 \right)\right] \right] \\ \sum_{j=1}^{N} \left[\frac{\widetilde{M}_j \cdot \pi_{j2}(p)}{\sum_{s=1}^{N} \widetilde{M}_s \cdot \pi_{s2}(p)} \left[1 + x_j \alpha \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma}(1/L_2) - \pi_{j2}/L_2 \right)\right] \right] \\ \vdots \\ \sum_{j=1}^{N} \left[\frac{\widetilde{M}_j \cdot \pi_{jN}(p)}{\sum_{s=1}^{N} \widetilde{M}_s \cdot \pi_{sN}(p)} \left[1 + x_j \alpha \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma}(1/L_N) - \pi_{jN}/L_N \right)\right] \right] \end{bmatrix}$$

D Observed vs. counterfactual expected price, all neighborhoods

Table D1: Counterfactual expected prices paid by origin neighborhhod

Retail location	Observed expected price	Scenario 1: disutility from travel Reduced 50%	Scenario 1: + Reduced κ	Scenario 2: improved amenities Talpiot only	Scenario 2: Three areas	Scenario 3: additonal entry
Neve Yaaqov	7.86	0.4%	0.0%	-2.2%	-3.4%	-2.6%
Pisgat Zeev North	7.48	-1.5%	-1.5%	-3.2%	-3.7%	-2.7%
Pisgat Zeev East	7.67	-4.2%	-4.2%	-6.3%	-6.8%	-0.8%
Pisgat Ze'ev east (north - west) & Pisgat Ze'ev west	7.46	-3.0%	-2.9%	-4.6%	-4.9%	-1.5%
Ramat Shlomo	8.20	-1.1%	-1.4%	-0.5%	-3.6%	-1.2%
Ramot Allon north	7.86	-3.3%	-3.5%	-5.2%	-6.7%	-1.6%
Ramot Allon	7.83	-3.6%	-3.8%	-5.5%	-6.9%	-1.1%
Ramot Allon South	7.75	-4.8%	-4.9%	-6.0%	-6.9%	-0.7%
Har H-hozvim, Sanhedriyya, Tel-Arza	8.29	-1.4%	-1.7%	-0.4%	-3.4%	-1.4%
Ramat Eshkol, Giv'at-Mivtar	8.12	-2.6%	-2.6%	-4.2%	-5.6%	-0.3%
Ma'a lot Dafna, Shmuel Ha-navi	8.07	-2.8%	-2.9%	-4.9%	-6.1%	-0.4%
Giv'at Shapira	7.85	-3.5%	-5.5%	-6.6%	-7.3%	-0.7%
Mamila, Morasha	7.80	-3.6%	-3.9%	-7.4%	-7.8%	-0.7%
Ge'ula, Me'a She'arim	8.18	-2.2%	-2.3%	-4.0%	-6.0%	-0.5%
Makor Baruch, Zichron Moshe	8.28	-2.5%	-3.1%	-2.6%	-6.0%	-1.1%
City Center	7.96	-3.6%	-3.9%	-7.4%	-8.2%	-0.8%
Nahlaot, Zichronot	7.93	-5.2%	-6.5%	-7.8%	-8.3%	-2.6%
Rehavya	7.98	-5.7%	-7.3%	-8.6%	-8.9%	-3.2%
Romema	8.24	-1.8%	-2.2%	-0.8%	-3.1%	-2.7%
Giv'at Sha'ul	7.97	-2.0%	-2.2%	-1.4%	-6.7%	-0.5%
Har Nof	7.62	-1.5%	-1.9%	-0.6%	-5.1%	-1.8%
Qiryat Moshe, Bet Ha-Kerem	7.67	-2.9%	-3.4%	-4.7%	-6.2%	-0.5%
Nayot	7.71	-3.1%	-3.4%	-5.1%	-6.6%	-0.9%
Bayit va-Gan	7.86	-3.0%	-3.3%	-6.0%	-7.4%	-0.9%
Ramat Sharet, Ramat Denya	7.71	-2.4%	-2.5%	-6.9%	-7.3%	-0.5%
Qiryat Ha-Yovel north	7.78	-3.4%	-3.5%	-6.7%	-7.3%	-0.7%
Qiryat Ha-Yovel south	7.72	-3.3%	-4.7%	-7.0%	-7.4%	-1.2%
Qiryat Menahem, Ir Gannim	7.86	-3.8%	-3.9%	-7.2%	-7.7%	-0.3%
Manahat slopes, Qedoshe Struma st, Ha-Ayal st	7.34	-1.7%	-1.8%	-4.6%	-4.7%	-0.5%
Gonen (Qatamon) A - I	7.41	-1.1%	-1.4%	-5.2%	-5.4%	-0.8%
Rassco, Giv'at Mordekhay	7.44	-1.8%	-3.2%	-5.4%	-5.6%	-1.6%
German Colony, Gonen (Old Qatamon)	7.28	-1.3%	-1.5%	-4.1%	-4.3%	-0.6%
Qomemiyyut (Talbiya), YMCA Compound	7.75	-3.0%	-3.4%	-7.4%	-7.7%	-0.8%
Baq'a, Abu Tor, Yemin Moshe	7.28	-1.1%	-1.2%	-4.1%	-4.2%	-0.3%
Talpiyyot, Arnona, Mekor Hayyim	7.21	-0.5%	-0.6%	-3.4%	-3.6%	-0.2%
East Talpiyyot	7.19	-0.9%	-1.0%	-3.1%	-3.3%	-0.2%
East Talpiyyot (east)	7.23	-1.2%	-1.3%	-3.5%	-3.7%	-0.2%
Homat Shmuel (Har Homa)	7.14	-1.3%	-1.3%	-2.6%	-2.8%	-0.1%
Gilo east	7.55	-2.4%	-2.4%	-6.4%	-6.6%	-0.2%
Gilo west	7.55	-2.7%	-2.8%	-6.3%	-6.6%	-0.2%
Talpyiot shopping area	7.14	0.0%	2.3%	-2.6%	-2.8%	-0.2%
Givat Shaul shopping area	7.51	-1.1%	0.8%	-1.9%	-4.7%	-0.4%
Malcha shopping center	7.29	-1.4%	-1.4%	-4.2%	-4.2%	-0.3%
Romema shopping area	8.34	-3.5%	-6.7%	-3.1%	-6.5%	-1.0%
Central Bus Station	8.06	-4.2%	-4.5%	-6.3%	-7.0%	-1.1%
Mahane Yehuda	7.79	-4.7%	-5.5%	-7.1%	-7.6%	-2.1%
Mean		-2.5%	-2.8%	-4.7%	-5.8%	-1.0%
Median		-2.5%	-2.8%	-4.8%	-6.2%	-0.8%
Price levels						
Mean price	7.72	7.53	7.50	7.36	7.27	7.65
Median price	7.75	7.51	7.43	7.28	7.20	7.67
Standard deviation of price	0.34	0.30	0.28	0.37	0.29	0.32

Notes: The table reports the percentage changes in expected prices charged at all 46 neighborhoods. See text for detailed explanations of each scenario. All analyses performed for the third time period (November 2008). The last three rows report statistics on the expected prices in levels rather than as percentage changes.

E Counterfactual analyses, $\sigma = 0.8$

Table E1: Percentage change in prices under counterfactual scenarios, sigma=0.8

Retail location	Observed price	Scenario 1: disutility from travel Reduced 50%	Scenario 1: + Reduced κ	Scenario 2: improved amenities Talpiot only	Scenario 2: Three areas	Scenario 3: additonal entry
Neve Yaaqov	8.01	2.6%	3.7%	0.0%	-0.2%	-2.4%
Pisgat Zeev North	7.36	0.5%	1.0%	-0.5%	-0.7%	-2.7%
Ramot Allon north	7.61	0.2%	0.4%	-0.2%	-0.3%	-2.9%
Giv'at Shapira	8.14	0.3%	0.7%	0.1%	0.1%	-1.1%
Rehavya	8.52	-6.3%	-9.3%	-2.7%	-1.3%	-5.9%
Romema	8.17	-1.0%	-2.0%	0.9%	1.3%	-3.8%
Har Nof	7.62	-0.4%	-1.0%	0.0%	-0.7%	-3.8%
Qiryat Moshe, Bet Ha-Kerem	7.85	-0.9%	-2.7%	0.2%	-0.5%	-1.6%
Qiryat Ha-Yovel south	8.19	-0.3%	-0.2%	-0.4%	-0.5%	-3.1%
Rassco, Giv'at Mordekhay	7.87	-1.2%	-2.5%	-0.3%	-0.5%	-4.0%
Baq'a, Abu Tor, Yemin Moshe	7.76	-0.2%	-0.2%	-0.1%	-0.2%	-2.7%
Talpiot shopping area	6.89	-0.2%	-0.2%	0.2%	0.0%	0.0%
Givat Shaul shopping area	7.07	-0.8%	-0.8%	0.2%	0.1%	0.0%
Romema shopping area	8.69	-0.8%	-0.8%	0.3%	0.2%	0.1%
Mahane Yehuda	7.20	-1.1%	-1.1%	0.1%	-0.1%	0.0%
Mean (residential)		-0.6%	-1.1%	-0.3%	-0.3%	-3.1%
Median (residential)		-0.3%	-0.2%	-0.1%	-0.5%	-2.9%

Notes: The table reports the percentage changes in prices charged at locations where prices are observed (11 residential neighborhoods and four commercial areas appearing in bold type) under the various policy interventions, computed at the third time period (November 2008). See text for expanations of each scenario. The last two rows report statistics that are computed over the 11 residential neighborhoods only.

Table E2: Percentage change in expected prices under counterfactual scenarios, sigma=0.8

Retail location	Observed expected price	Scenario 1: disutility from travel Reduced 50%	Scenario 1: + Reduced κ	Scenario 2: improved amenities Talpiot only	Scenario 2: Three areas	Scenario 3: additonal entry
Neve Yaaqov	7.86	0.6%	0.3%	-2.1%	-3.3%	-2.1%
Pisgat Zeev North	7.48	-1.3%	-1.3%	-3.2%	-3.6%	-2.1%
Ramot Allon north	7.86	-3.1%	-3.2%	-5.3%	-6.8%	-1.3%
Giv'at Shapira	7.85	-3.4%	-5.4%	-6.8%	-7.4%	-0.6%
Rehavya	7.98	-4.9%	-6.7%	-8.6%	-8.9%	-2.8%
Romema	8.24	-1.5%	-1.8%	-1.0%	-3.2%	-2.3%
Har Nof	7.62	-1.3%	-1.6%	-0.6%	-5.3%	-1.5%
Qiryat Moshe, Bet Ha-Kerem	7.67	-2.7%	-3.1%	-4.8%	-6.4%	-0.4%
Qiryat Ha-Yovel south	7.72	-3.1%	-4.5%	-7.1%	-7.4%	-1.1%
Rassco, Giv'at Mordekhay	7.44	-1.6%	-3.0%	-5.6%	-5.7%	-1.4%
Baq'a, Abu Tor, Yemin Moshe	7.28	-0.9%	-1.0%	-4.3%	-4.3%	-0.3%
Mean		-2.1%	-2.9%	-4.5%	-5.7%	-1.4%
Median		-1.6%	-3.0%	-4.8%	-5.7%	-1.4%

Notes: The table reports the percentage changes in expected prices charged at the same 11 residential neighborhoods displayed in Table E1. See text for detailed explanations of each scenario. All analyses performed for the third time period (November 2008).

F Robustness of demand estimates to the computation of the composite good price

As explained in Section 2.2, we perform robustness checks to verify that our results are not driven by the way we computed the price for the composite good. Estimation results appear in Table F1. Elasticities are reported in Table F2.

Table F1: Robustness results

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline (Col 6 from Table 6)	No. of products in composite >= 9	Imputed prices	Fruits & Vegetables	Including Zero exp.	Supermarkets only
ln (price at destination)	4.727	3.090	4.107	1.75	5.349	4.061
	(1.304)	(1.200)	(1.763)	(0.458)	(1.766)	(1.344)
ln (price at destination) X housing prices	-0.232	-0.157	-0.176	-0.077	-0.219	-0.216
	(.078)	(0.064)	(0.127)	(0.034)	(0.132)	(.08)
Distance to destination	0.423	0.484	0.452	0.48	0.377	0.409
	(.12)	(0.097)	(0.090)	(0.103)	(0.170)	(.13)
Distance to destination Xsenior citizen	0.004	0.004	0.004	0.005	0.004	0.004
	(.007)	(0.006)	(0.005)	(0.007)	(0.012)	(.008)
Distance to destination Xdriving to work	-0.003	-0.004	-0.003	-0.004	0	-0.003
	(.002)	(0.001)	(0.001)	(0.001)	(0.002)	(.002)
Shopping at home	1.890	1.873	1.849	1.897	2.16	1.932
	(.426)	(0.294)	(0.259)	(0.297)	(0.485)	(.438)
# observations	1819	2354	2968	2091	2070	1633
R^2	0.784	0.767	0.769	0.757	0.704	0.776

Notes: 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. All regression includes fixed effects for origin, destination, periods and destination interacted with housing price at origin. Standard errors in parentheses are (2-way) clustered at the origin and destination levels.

First, we add locations having at least 9 prices out of the 27 prices for the 27 products. This increases the number of destinations from 15 to 20 in the first period and 19 in the second and third periods and the number of observations used in the regression to 2,354. Doing this decreases the price coefficient and the coefficient of its interaction with housing prices at origin, although they are still both significant (column 2). This attenuation of the estimates could reflect increased measurement error in prices brought about by the inclusion of locations with a different specification of the composite good. This attenuation translates into a decrease in own prices elasticities from a median elasticity of 4.95 to a median price elasticity of 3.18 (see Table F2). Remarkably, the estimates of the parameters related to distance remain basically unchanged. This will also hold for the other robustness checks.

A second check is to use our socioeconomic data to impute prices of products in locations where they are missing. For each subquarter we compute the mean price (over stores) for each product and period. We then regress each of these (mean) prices separately on a set of socioeconomic variables at the subquarter level, and compute predicted prices for each product and location.[46] In subquarters where prices of some products are missing we impute the predicted prices, and proceed as before to compute the price of the composite good for each of the destinations where some price data were available.[47] The price of the composite good is now a weighted average of all 27 products. Over all products and locations, the fraction of imputed prices is 31.5 percent. The estimated parameters are somewhat lower than in the baseline specification, again possibly consistent with attenuation bias due to the measurement error in prices brought about by the imputation exercise. The estimated own price elasticities are a bit smaller and more dispersed than in the baseline specification.

In a third robustness check, we estimate the baseline regression using fruits and vegetables only (11 items).[48] The estimated price elasticity is now about a half than in the baseline specification. This is not surprising since demand for fruits and vegetables is likely to be less price sensitive than for other products. Note, however, that the sensitivity to distance is about the same as for the full composite good. We also substitute a very small number (1 NIS) when expenditures are zero. We can now use the 2070 ($46 \times 15 \times 3$) observations. Results appear in column (5) of Table F1 and are a bit larger than in the baseline specification. The corresponding elasticities

[46] The socio economic variables used to predict prices are a subset of the following: number of family households, median age, percentage of married people aged 15 and over, average number of persons per household, percentage of households with 7+ persons in the household, percentage of households with 5+ children up to age 17 in the household, dependency ratio, percentage of those aged 15 and over in the annual civilian labor force, percentage of those aged 15 and over who did not work in 2008, percentage of Jews born abroad who immigrated in 1990-2001, percentage of households residing in self-owned dwellings, percentage of Jews whose origin is Israel, percentage of Jews whose continents of origin are America and Oceania, percentage of Jews whose continent of origin is Europe, percentage of those aged 15 and over with up to 8 years of schooling, percentage of those aged 15 and over with 9-12 years of schooling, percentage of those aged 15 and over with 13-15 years of schooling, percentage of those aged 15 and over with 16 or more years of schooling. In addition, we added an indicator for a commercial area and period dummies. The $R^{2\prime}s$ of these 27 regressions are quite high, ranging from 0.45 to 0.93 with a median value of 0.70.

[47] In 16 observations with missing prices where the imputed price was negative it was substituted for by the minimum imputed price for each product. In neighborhoods that were not sampled in the three periods we imputed prices only for the periods for which we had some price data (these are the neighborhoods with zero number of sampled stores in Table 3). Thus, for example, in November 2008 we imputed prices for 23 out of the 26 neighborhoods.

[48] In a few locations, the basket is composed of nine or ten fruits and vegetables.

Table F2: Distribution of estimated elasticities (absolute value)

Specification	mean	sd	min	p10	p25	p50	p75	p90	max	N
					Own price elasticity					
Specification	mean	sd	min	p10	p25	p50	p75	p90	max	N
Baseline (col 6 Table 6) σ = 0.7	4.82	0.92	3.00	3.86	3.99	4.95	5.87	5.95	6.13	15
Baseline (col 6 Table 6) σ = 0.8	6.43	1.37	3.78	5.01	5.31	6.54	7.94	8.32	8.47	15
Composite with 9 or more products	3.08	0.77	1.67	1.91	2.51	3.18	3.54	4.12	4.21	19
Imputed prices	4.40	1.26	2.30	2.67	3.01	4.52	5.34	5.89	6.34	23
Fruits and Vegetables	2.51	0.49	1.55	1.68	2.23	2.58	2.84	3.20	3.22	19
Including zero Exp.	6.60	0.96	4.75	5.55	5.68	6.59	7.29	8.02	8.16	15
Supermarkets only	3.84	0.87	2.15	2.94	3.09	3.88	4.75	4.96	5.20	15
					Distance semi-elasticity					
Specification	mean	sd	min	p10	p25	p50	p75	p90	max	N
Baseline (col 6 Table 6)	0.35	0.06	0.06	0.28	0.31	0.35	0.39	0.42	0.45	690
Composite with 9 or more products	0.37	0.07	0.06	0.29	0.33	0.37	0.43	0.48	0.50	874
Imputed prices	0.37	0.05	0.16	0.31	0.33	0.37	0.42	0.45	0.47	1,058
Fruits and Vegetables	0.37	0.07	0.13	0.28	0.32	0.37	0.44	0.48	0.50	798
Including zero Exp.	0.40	0.05	0.09	0.36	0.39	0.40	0.42	0.44	0.48	690
Supermarkets only	0.34	0.06	0.06	0.27	0.30	0.34	0.38	0.41	0.44	645

Notes: Elasticities are computed for November 2008. σ = 0.7 is used except in row 2 of top panel. Price elasticities are computed for each destination. Prices were imputed for 23 out of the 26 neighborhoods in November 2008. Distance semi-elasticities are computed for each oirign-destination pair (e.g., 46x15=690).

are shown in Table F2 and are somewhat larger than in the baseline case but, again, within the same order of magnitude. In a final check we use only price data from supermarkets and we find that estimated coefficients (column 6 of Table F1) and elasticities are very similar to the baseline results.

In sum, using different cuts of the price data does not alter the basic conclusion from Table 6 that prices and distance decrease utility in a way and in an order of magnitude that are economically sensible. These results, particularly those related to distance, are quite stable across the various subsamples.