An Empirical Analysis of Search Costs and Price Dispersion

Joshua Sherman[*]   Avi Weiss[**]

Abstract

We exploit cross-sectional and temporal differences in search intensity in order to examine the relationship between search costs and price dispersion using a hand-collected panel data set from Jerusalem's *Shuk Mahane Yehuda* outdoor market. We present empirical evidence that price dispersion increases with the cost of search using several different measures of price dispersion, however, our interpretation of this finding is sensitive to the search proxy in question. We also address several acute difficulties facing empiricists seeking to test theoretical price-dispersion models in which consumers are heterogeneous in their search behavior.

JEL Classifications: L11, L13

This Version: May 17, 2014

[*]Department of Economics, University of Vienna. Email: joshua.sherman@univie.ac.at.
[**]Department of Economics, Bar-Ilan University and IZA. Email: avi.weiss@biu.ac.il.

I.   Introduction

In this study we utilize a hand-collected dataset of daily prices in Jerusalem's outdoor market, *Shuk Mahane Yehuda*, in order to empirically investigate the relationship between price dispersion and search costs. Specifically, we employ several proxies for search costs in order to capture the effect of cross-sectional and temporal differences in search intensity on the extent of price dispersion that is observed. Examining various measures of price dispersion, we provide empirical evidence that price dispersion increases with the cost of search. In addition, we exploit our empirical environment in order to demonstrate difficulties facing empiricists seeking to test theoretical models of price dispersion in which consumers are heterogeneous in their search behavior. In our opinion, some of these difficulties have not received sufficient attention in previous literature.

Empiricists have encountered obstacles testing the theoretical relationship between search costs and price dispersion, largely because much of search theory predicts a non-monotonic relationship between the intensity of search and price dispersion. This non-monotonicity is easily understood by considering two cases. Suppose that all consumers have identical demand and that firms, each differentiated in its location, have symmetric costs. In one case, all consumers costlessly search prices of every firm, where costless search implies costless acquisition of information as well as costless travel from one store to the next. In the other case, all consumers have positive search costs. Previous studies have shown that when all consumers costlessly search the setting reduces to a Bertrand price setting game and when all consumers have positive search costs, every firm acts as a monopolist over its consumers. In both cases, there exists a pure strategy equilibrium in which all firms charge the same price. However, if a market contains both costless searchers (shoppers) and consumers with positive search costs

2

(non-shoppers), firms have dueling incentives to charge the lowest price in the market in order to attract all shoppers and to charge the monopoly price in order to maximize the surplus that can be obtained from non-shoppers. A well-known result in such a case is that no pure strategy equilibrium exists in a static setting.[1] It follows that in such a theoretical environment, the extent to which firms randomize prices will first increase and then decrease with the proportion of shoppers. Therefore, any empirical finding of a monotonic relationship between search intensity and price dispersion may be consistent or inconsistent with theory depending on the actual proportions of shoppers in the market.[2]

Another reason that empiricists have encountered difficulties testing the theoretical relationship between search costs and price dispersion is that it is not easy to link the level of sophistication of consumers in any particular market with the assumptions of search-theoretical models. For example, it is difficult to discern the extent to which consumers' search patterns endogenously adjust to changes in variables such as costs or the number of firms. Recent theoretical models assume the consumer has an ability to understand how such changes should affect one's incentives to search, however, it is not clear that consumers are sophisticated enough to do so. Therefore, the question of whether search is exogenous or endogenous in a given empirical setting is rarely clear. This is an important issue because predictions of equilibrium price dispersion vary given the chosen assumption. We believe that these difficulties have not received sufficient attention in previous studies.

In practice, the empirical literature dealing with consumer search has centered on retail gasoline markets, supermarkets, and retail goods and services sold on the internet, most likely

---

[1] For example, see Varian (1980).

[2] As noted by Chandra and Tappata (2011) and others.

due to the fact that it is difficult to find proxies for search costs in other industries that offer sufficient variation for purposes of understanding their effect on prices. In contrast, this study analyzes a large outdoor market that occupies a single physical location, which generates several advantages relative to previous studies. One advantage of this empirical setting is that it allows us to identify firm pairs for which both consumer search costs and consumer travel costs are zero – two conditions required in theory for a Bertrand homogeneous goods pricing result to prevail. That is, vendors in our environment are situated along pedestrian walkways, and, therefore, not only may customers observe prices of two firms located across from one another (or adjacent to one another) without travelling, such vendors are located literally only two to three steps away from one another. Another advantage of our unique empirical setting is that we may exploit previously unexplored types of variations in search costs, such as variation in pedestrian congestion, as well as variation previously examined in different empirical settings, such as different opportunity costs of searching (due to presumed differences in consumer incomes), in order to understand the influence of consumer search costs on price dispersion.

The literature on price dispersion is vast and has been a topic of considerable interest in the economics literature for at least the past fifty years. Baye et al. (2006) provide a thorough survey of both the theoretical and empirical literature that has sought to understand and explain the phenomenon.[3] Search-theoretic models that predict equilibrium price dispersion vary in their assumptions regarding simultaneous vs. sequential search, exogenous vs. endogenous search, the number of prices observed by the consumer, cost symmetries amongst firms, the nature of

---

[3] Theoretical models of price dispersion begin with Stigler's (1961) seminal work in which he concludes that dispersion of prices results from costly search. In the monopolistic competition model of Salop and Stiglitz (1977), spatial price dispersion is observed despite the homogeneity of the product sold. Previous literature has commonly referred to spatial price dispersion as an indication that some firms persistently sell at lower prices than others. Shilony (1977), on the other hand, constructs a model in which temporal price dispersion is observed. A small sample of subsequent theoretical studies includes Rosenthal (1980), Varian (1980), Burdett and Judd (1983), Carlson and McAfee (1983), Stahl (1989), Dana (1994), Janssen and Moraga-González (2004), Janssen, Moraga-González, and Wildenbeest (2005), and Lach and Moraga-González (2011).

consumer demand, and whether the search setting is static or dynamic. Empirical studies that provide evidence of temporal price dispersion include Lach (2002) and Hosken, McMillan, and Taylor (2008), amongst others. These studies find that sellers change their relative position in the price distribution frequently over a period of time. Several empirical studies have also empirically investigated the relationship between competition and price dispersion.[4] Barron et al. (2004) analyze four geographic gasoline markets in the United States and find that the amount of price dispersion decreases with the level of competition, contrary to predictions of search-theoretic models which the authors thought comparable to the empirical environment that they examine. Lewis (2008), on the other hand, finds that the relationship between competition and price dispersion in the data of gasoline stations he examines depends on whether dispersion is measured city-wide or locally. That is, if one measures price dispersion relative to nearby stations rather than the city as a whole, he finds that price dispersion increases with competition. When price dispersion is measured relative to the entire city, however, he finds the opposite result. Chandra and Tappata (2011), also using gasoline station data, find that price dispersion increases with the number of firms in the market and with search costs and decreases with production costs. De los Santos et. al. (2012) draw inferences regarding consumer search using data on the web browsing and purchasing behavior of consumers. Zhou (2006) studies Chicago supermarkets and finds a positive correlation between search costs and price dispersion, Tang et al. (2010) find that increased use of shopbots results in less price dispersion, and Orlov (2011) finds that increases in Internet penetration by airlines lead to higher intrafirm price dispersion. Most recently, the empirical relationship between search costs and price dispersion has also been

---

[4] Borenstein and Rose (1994) analyze price dispersion amongst airline carriers using cross-sectional data from 1986 and find that price dispersion increases with competition. However, Gerardi and Shapiro (2009) responded to their study with a panel data study on the airline industry from 1993 to 2006 and conclude the opposite result. Png and Reitman (1994), studying gasoline station data from Massachusetts, find that prices are more dispersed in markets with relatively more competitors.

explored by Dubois and Perrone (2012), who analyze French supermarkets, and Pennerstorfer et al. (2014), who analyze retail gasoline stations in Austria. The former study finds that search costs partially explain price dispersion, and also concludes that the vast majority of consumers in their setting do not search. The latter study finds that the empirical relationship between information and price dispersion is consistent with theoretical predictions of Stahl's (1989) model of sequential search.

The paper is structured as follows. Section 2 provides a short background of *Shuk Mahane Yehuda* and describes the data. Section 3 presents the different proxies for search costs available in our empirical setting that will be used as the explanatory variables of primary interest in our analysis. Section 4 discusses the different measures of price dispersion we seek to explain and the associated estimation procedures. In Section 5 we present, discuss and interpret the empirical results. Section 6 concludes.

II.   The data

The Mahane Yehuda market, located in the commercial center of western Jerusalem, dates back to the late 19th century and underwent a substantial renovation in the early 2000s.[5] It is one of the largest outdoor markets in Israel, consisting of hundreds of independent vendors. The largest percentage of vendors sell vegetables and fruits; other products sold include meats, poultry, fish, baked goods, and clothing. The market is open Sunday through Friday, although stores close in the late afternoon on Fridays prior to the Jewish Sabbath. Prices are generally

---

[5] In what follows we provide a brief description of the Mahane Yehuda market. A more detailed description may be found in Sherman and Weiss (forthcoming).

6

steady between 9:30 a.m. – 3:30 p.m. (9:30 a.m. – 1:00 p.m. on Fridays and holiday eves), after which prices for most perishable goods begin to decline.

The majority of product sold by retailers at Mahane Yehuda is bought at the Jerusalem wholesale market.[6] There is minimal advertising in the market with the exception of a sign with the firm's name (in some instances) and occasional shouts highlighting particular products, prices, or both. Also, it is uncommon for customers to negotiate with vendors on price for vegetables in Mahane Yehuda. However, it is possible that vendors who do not post their prices for particular products engage in some form of price discrimination for these products.

The market is neither uniform in product quality nor in terms of environmental conditions (such as cleanliness, lighting, etc.). Quality differences across different areas of the market are largely driven by differences in customer demographics across these areas. For example, the southern area, which we will refer to as Area A, features the highest quality product and is closest to the parking lots, where customers who arrive by car enter the market (a hand-drawn map of the market showing the different areas is provided in Figure 1). On the other hand, the area to which we will refer as Area D features the lowest quality produce and is located at the northern end of the market near the bus lines and a neighborhood where relatively low income residents live (although a gentrification process began in Area D after the completion of the data collection). The area to which we will refer as Area B has entrances at the northern and eastern ends of the market and the area to which we will refer as Area C has entrances at the northern and western ends of the market – the quality of product sold in Areas B and C is positioned

---

[6] There are a few vendors that purchase their product directly from farmers, purchase from the Tel Aviv wholesale market, or receive their product directly from the West Bank. With this, we note that any cost and quality differences amongst retailers who purchase product from different sources at the Jerusalem wholesale market or in general are more pronounced across geographic areas of the market than within geographic areas. Such differences across geographic areas will be related to our geographic market definition in the empirical analysis, discussed in detail in Section 3.

somewhere between the quality of product found in Areas A and D. Along these lines, we assume that the marginal costs of firms in a given area are comparable while we expect that the marginal costs of firms across areas will differ.

Consequently, we divide the market into these four geographical areas, which generally correspond to four levels of quality of product sold.[7] While one might suggest alternate ways of defining geographic markets in this setting, distance-based measures will be misleading because they would not take into account abrupt changes in the physical environment (cleanliness, for example) that signal differences in the quality of product sold. The chosen partition is supported by the average product prices for each of our ten products in each of the four areas, presented in Table 1. As seen in the Table, there are large differences in average price across areas for each of the goods.[8]

Data collection

Prices were hand-collected daily between 10:00 a.m. – 12:00 p.m. during the period April 2, 2009 – August 4, 2009 for the following ten products: beets, cauliflower, cucumbers, corn, kohlrabi, red cabbage, red peppers, sweet potato, white cabbage, and yellow peppers. These observations originate from the same data set analyzed in Sherman and Weiss (forthcoming) and were chosen in order to restrict the analysis to homogeneous, nondurable products. Ten products were chosen for purposes of keeping the daily hand collection of data manageable. In addition,

---

[7] We also note that market insiders have indicated to us that a given firm is typically not disciplined by firms in other areas of the market for reasons previously cited.

[8] There are three firms that are not in the immediate vicinity of the other firms in Area D but are classified as Area D firms due to the extremely low quality of product that they sell. Most low income consumers with low search costs would presumably be aware of these three firms. In the data set we analyze there are a total of 12 firms in Area A, 15 firms in Area B, 17 firms in Area C, and nine firms in Area D that sell at least one of the products for which we collected data more than once during the data collection period.

fruits were not chosen for observation as many fruits were subject to Jewish sabbatical year dietary laws during the period and therefore measures of competition would have been more difficult to interpret for these products.[9]

On a given day the majority of firms that sell also post prices for most, if not all, of their products.[10] When price was not posted, the instance was noted, but vendors were not surveyed due to concerns for the researcher's anonymity (the number of firms posting price and the number of firms selling overall on a given day is highly correlated). In Table 2 we display the frequency of the number of firms posting price in a particular area, by product.

The market is characterized by a large number of vendors in very close proximity. On rare occasions, the dividing line distinguishing one vendor from another was indistinguishable to the researcher, and this became particularly difficult when vendors which typically did not sell products that we tracked chose to sell a tracked product for a period of only one to two days. In some cases, inquiring as to which vendor a particular product belonged would have jeopardized the researcher's anonymity. As a result, 350 observations were not included in our analysis.[11] In extremely rare cases, one stand sold a particular product at two different posted prices due to obvious quality differences. The lower priced product was likely product that did not sell out from the previous day, so these lower priced observations were not recorded by the researcher. In all other cases, firms post one price per product or choose to not post price. Due to oversight,

---

[9] Vendors differ with respect to their level of compliance with these religious laws, thereby appealing to different clientele. This heterogeneity also results in an increased level of cross-sectional price dispersion as stricter compliance generally coincides with higher prices.

[10] There are a total of 7,491 posted price observations out of a total of 11,983 product observations in the dataset. Generally speaking, a given firm posts prices for nearly all of its products or for none of its products.

[11] In Sherman and Weiss (forthcoming) only nine such observations needed to be omitted for this reason because the remaining 341 observations could not be used in that analysis due to a requirement in that study that firms post price and retain their identities (rival or isolated) for several consecutive periods.

one stand that typically sold sweet potatoes was occasionally missed. In addition, there was one stand that began selling vegetables in the middle of the data period but was not noticed until somewhat later. Once noticed, it was learned anecdotally that this stand hardly ever posted price for its products.

Observations of prices on Sundays raise particular concerns. Many stands are closed on Sundays and a portion of the stands that are open sell low quality product remaining from the previous Friday. It becomes particularly difficult to interpret prices on Sunday over the time period because a given stand might sell fresh product on certain Sundays and leftover product on other Sundays at much lower prices.   Therefore, price dispersion on Sundays would be overestimated in our regressions since firm×product fixed effects could not account for such differences. For these reasons, Sunday observations are excluded from this analysis.

III.   Proxies for search

As stated above, our empirical setting contains several proxies for search costs. These proxies will serve as the variables that we shall analyze for purposes of explaining the various measures of price dispersion presented in the next Section. A more detailed discussion of these proxies and their implications will be discussed in Section 5 when presenting and discussing the results. We utilize the following six proxies:

a) Firm status: Our first variable of interest relates to whether a given firm is situated immediately adjacent to another firm or immediately across from another firm (firms that we shall refer to as *rivals*). Rival firms are unique because any consumer located at one such firm incurs literally no cost in order to learn the rival firm's price. Since consumer search costs are

zero at such firms we would expect that temporal volatility of relative prices (temporal price dispersion) amongst rival firms should be relatively low. Thus, we expect rival firms to exhibit less price dispersion than firms not facing such immediate competition. Any other firm will face a strictly positive proportion of customers who incur a search cost in order to compare prices at another firm.

b) Day of the week: The markets in Israel become more congested as the week progresses, and they are the most congested on Fridays as Friday is the only shopping day of the Israeli weekend and immediately precedes the Jewish Sabbath. The high concentration of customers at the end of the week increases the time and effort required to travel from one firm to another, and it may also be the case that time is more precious to customers on Fridays if they feel pressured to prepare for the Jewish Sabbath during the day. Thus, we expect search costs to increase and the percentage of shoppers to correspondingly decrease as the week progresses.

c) Geographic area: The underlying differences between the different geographic areas yield a proxy for cross-sectional differences in search costs. Since higher income consumers are likely to have a higher opportunity cost to search than lower income consumers, a relatively high proportion of consumers in Area A are likely to possess high search costs whereas a relatively low proportion of consumers in Area D are likely to possess high search costs.

d) Number of firms: There clearly must be at least two firms in an area in order for price dispersion to exist. As the number of firms increases, lowering price in an attempt to draw consumers for whom search is costless becomes less worthwhile because the probability of being the lowest-price firm decreases. Therefore, the extent to which firms randomize prices is

influenced by the number of firms that compete in an area, which, in turn, influences price dispersion.

e) Wholesale price level: Theoretically speaking, the extent of price dispersion should increase with the size of price-cost margins. However, the empirical manifestation of this effect is far from obvious for several reasons. One of these reasons stems from the fact that consumers often do not know the current marginal cost, and not all firms respond to cost changes with the same immediacy. In particular, Sherman and Weiss (forthcoming) show that retail prices may not respond to changes in wholesale prices immediately in the Mahane Yehuda market due to the existence of consumer information lags. Furthermore, the nature of consumer demand (e.g. unit, logarithmic, linear), of which consumers themselves may not be aware, will dictate whether there is a positive or a negative relationship between the cost level and the size of price-cost margins. Later in the study we shall address these complications in more detail as well as additional difficulties in forming empirical predictions regarding the relationship between wholesale price and the extent of price dispersion.

f) Product identity: Product dummies will also be included as controls in our analysis. However, while the nature of the product sold may influence the extent of price dispersion that is observed, such predictions are not obvious and will be discussed in Section 5.

IV.   Measures of price dispersion

In this Section we present the various measures of price dispersion that will be investigated in this paper and discuss our estimation procedures. Since we will use virtually the same variables (those presented in the previous Section) to explain each measure of price

dispersion, we do not present any results in this Section; rather, we present the tests only and defer discussion of the results until Section 5.

a. Measuring dispersion using residual prices

Using actual prices in order to measure the extent of price dispersion can be problematic due to unobserved heterogeneities across firms. Therefore, consistent with an approach taken in the literature to address this matter, we begin by estimating a firm's unexplained price variation using the residuals from a first stage regression in which we explain price, whereby each residual signifies a firm's homogenized price for that observation. The homogenized prices obtained from this regression will then be used in order to examine price dispersion in a second stage regression.[12]

As stated earlier, we examine prices for ten different fresh vegetables in our analysis. Therefore, to the extent that price differences amongst firms selling a particular product within a given geographic area are due to seller or product heterogeneity, such differences may be dealt with either by employing controls for observed product and seller characteristics or by using fixed effects for each firm×product combination. We employ firm×product fixed effects in order to control for all observed and unobserved heterogeneities that remain constant over time, as this is likely the more appropriate approach for purposes of accounting for heterogeneities in an empirical setting in which sellers engage in personal interaction with buyers of fresh vegetables. Note that while the quality of a product sold by a particular retailer may certainly vary over time in an absolute manner for a variety of reasons that the retailer cannot control (e.g.

---

[12] Wildenbeest (2011) points out that random residual prices are generated by construction if store fixed effects are not additively separable from their costs. In such a case, random residual prices could be interpreted as evidence of mixed strategies. This is one well-known drawback of analyzing price dispersion via residual prices.

weather), we presume that the index of quality of all firms in the market relative to one another does not change over time.

We must also account for fluctuations in price over time due to changes in marginal cost, in our case wholesale price. One possible approach would be to include the current period's wholesale price for the product in question as an explanatory variable. However, as noted previously, retail prices may not respond to changes in wholesale prices immediately in the Mahane Yehuda market due to the existence of consumer information lags. Therefore, we employ a date×product interaction variable in order to account for aggregate time-varying effects common to a particular product on a given date, including wholesale price prior to and on that date. We shall consider the implications of consumer information lags in the context of predicting the extent of price dispersion in more detail in Section 5.[13]

Given the differences in the four geographical areas discussed earlier in the paper, we run first stage regressions separately for each of the four areas of the market. The justification for this approach is perhaps most clearly understood by considering the case in which only one firm sells a particular product in each geographic area on a particular date. Given that we have defined markets according to geographic area, in such a case price dispersion cannot exist. However, if we were to include observations from all areas in the same first stage regression in which there are firm×product fixed effects, residuals for such observations would not be equal to zero. Put differently, if we were to include all observations in the same first stage regression, we would be making the implicit assumption that the relevant geographic market is the entire market.

---

[13] In particular, Tappata (2009) discusses predictions of price dispersion in a dynamic environment given consumers' expectations of marginal cost. We will return to these predictions later in the paper in our discussion of price dispersion in a dynamic environment.

In summary, first we claim that observed prices may be explained by the product (*PROD*) g in question, quality and service offered by each particular firm f, and the date t, which accounts for aggregate time-varying effects common to a particular product. Note that we interact *PROD* with firm indicator variables as well as date indicator variables in order to distinguish amongst products sold by a given firm and amongst distinct time-varying effects for different products:

$$p_{fgt} = \alpha + \sum_g \gamma_g PROD_g + \sum_g \sum_f \theta_{fg} FIRMPROD_{fg} + \sum_g \sum_t \tau_{gt} DATEPROD_{gt} + u_{fgt} \quad (1)$$

This regression is run for each of the four areas of the market separately, and the estimated residuals from the regression may be interpreted as homogenized prices net of all time-invariant and time-varying effects. Understanding the nature of these residual prices will allow us to further understand relative price movement amongst sellers, or price dispersion. There are a total of 7,491 posted price observations in our dataset, however, the residual price associated with 862 of these observations is zero by construction, and therefore must be excluded from our analysis.[14]

In order to examine the behavior of the residuals from each of the four first stage regressions, we refer to Harvey's (1976) general formulation for the structure of the variance of

---

[14] Specifically, there are 766 observations in which there is only one firm in the geographic area posting price for a given product on a particular date and there are 65 observations in which a given firm sells a particular product only once throughout the time period. As there are 16 observations that meet both of the preceding criteria, 815 such observations are eliminated. Furthermore, given a firm×product combination for which there are only two observations, one of which is on a date when it is the only firm selling in the area, the residual price associated with the other firm×product observation on a different date will also necessarily equal zero. Likewise, given an area×product×date combination for which there are only two observations, one of which is the only instance in which a firm×product instance occurs in the dataset, the residual price associated with the other area×product×date observation will also equal zero. There are an additional 47 observations meeting at least one of the two preceding classifications.

the error term, in our case $\sigma^2_{u_{fgt}}$ from (1).[15] By using $\hat{u}^2_{fgt}$ as an estimate of $\sigma^2_{u_{fgt}}$ and taking a

logarithmic transformation of Harvey's general formulation of a model with multiplicative

heteroskedasticity, the residuals $\hat{u}_{fgt}$ from the untransformed first-stage OLS regression may be

used in order to obtain coefficient estimates of the variables that influence the extent of price

dispersion observed.[16] In our case, we first run Equation 1 separately for each of the four areas

of the market and obtain estimated residuals from each regression. We then run the following

OLS and Prais-Winsten AR(1) specifications using the log of the squared residuals obtained

from the first stage regression as the dependent variable:

$$\ln\left(\hat{u}^2_{fgt}\right) = \vartheta + \psi x_{fgt} + \varepsilon_{fgt} \tag{2}$$

where x_{fgt} is the vector of search proxies that were presented in Section 3, ϑ is a constant, and

ε_{fgt} is the error term. Results are presented in Table 3 and discussed in Section 5. Given that

we interpret price dispersion as randomization of prices, it may be more appropriate to

exclusively examine non-rival firms. In Table 3 we therefore also display the results of a

specification that excludes observations of rivals.

b.   Measuring dispersion using actual prices

We also account for unobserved heterogeneities across firms by analyzing the actual

prices charged across firm×product pairs over time.[17] Denoting the two firms that constitute a

---

[15] See Genesove (1995) for an additional application involving multiplicative heteroskedasticity.

[16] For more details on the replacement of σ^2_{fgt} with $\hat{u}^2_{fgt}$, see Amemiya (1985).

[17] In order to address the phenomenon of temporal price dispersion several recent studies have separated the market into price percentile ranges and analyzed firms' price movements across ranges over time (e.g., Lach (2002) and Lewis (2008)). However, such an exercise would be inappropriate in our case due to the nature of our data. Namely, firms' entry and exit during the data period limits our ability to interpret a firm's price movements relative to the market over time as entry and exit of firms may distort relative price movements.

given firm×product pair selling a particular product g on date t as firms i and j, we examine the difference (in absolute value) between the price spread of a given firm×product pair on a particular date $(spread_{gijt})$ and the average price spread of that firm×product pair over the entire time period $(\overline{spread}_{gij})$. This measure, which we denote as s_{gijt}, is therefore defined as:

$$s_{gijt} = \left| spread_{gijt} - \overline{spread}_{gij} \right| \tag{3}$$

Such a measure of price dispersion is useful because it accounts for service or product heterogeneity between vendors as well as price differences due to differences in market power, which may exist as a result of a given firm's proximity to other competitors. We then examine s_{gijt} as defined in (3) as a function of the search variables introduced in Section 3. Results may be found in Table 4.[18]

Given that it is easy for consumers to compare prices when search costs are low, we may also expect to observe firm pairs charging the same price with a higher probability when it is easy to compare prices. While this measure of price dispersion is more directly related to the extent of product/service homogeneity or difference in market power between two firms rather than temporal price dispersion, it is also informative to understand the extent of spatial price dispersion in the market.[19] Along these lines, Table 4 reports the results of a logit specification in which the dependent variable assumes a value of 1 if the prices of a given firm×product pair on a given date are equal. Alongside the logit regression results we also report average marginal

---

[18] Throughout this section, firm×product pairs that sell on the same date only once throughout the time period are not included in the analysis, as such pairs are not useful for purposes of studying temporal price dispersion.

[19] We use the term *spatial* with regards to price dispersion as the term has been used in the literature to distinguish between price dispersion due to product or service heterogeneities or differences in the number of competitors a firm faces, for example, and temporal price dispersion due to firms' alternating positions in the price distribution over time.

effects; for indicator variables these are differences in predicted probabilities due to a discrete change in the value of the indicator variable from 0 to 1.

We also examine two additional measures of price dispersion that are best considered together. The first is another measure of spatial price dispersion: the absolute value of the spread of firm pairs, or $\left| spread_{gijt} \right|$. In particular, we expect that the average absolute spread between rivals will be smaller than the average absolute spread between non-rivals due to rivals' homogeneity in physical location. The second measure is the likelihood with which a given firm i charges a higher price than firm j, given that firm i charges a lower price than firm j *at least* 50% of the time. Chandra and Tappata (2011) were the first to examine this measure, which they refer to as *rank reversals*.[20] While theoretically speaking we would expect to observe fewer rank reversals between rival firm×product pairs because a pure strategy equilibrium exists for such firm pairs (as discussed earlier), we also expect that the number of rank reversals of a given firm×product pair will decrease with that pair's average absolute price spread. That is, if two firms charge substantially different prices on average, they are unlikely to exhibit frequent rank reversal. The average spread is analyzed in Table 5 using OLS and rank reversals are analyzed in Table 5 using a logit regression in which the dependent variable assumes the value of 1 if a rank reversal occurs for a given firm×product pair on that date.

c. Extensions

We also examine whether the relationship between search costs and price dispersion is robust to measures of price dispersion which more appropriately represent the extent of cross-

---

[20] Chandra and Tappata (2011) examine rank reversals and the standard deviations of price differences between firms in order to verify the existence of temporal price dispersion in their dataset of gasoline stations. By distinguishing between stations at the same corner and stations located further away from each other, they show that temporal price dispersion is a result of costly consumer search.

sectional price dispersion, noting that these measures do not control for persistent and unobservable within-area "service" differences amongst firms. Along these lines, we examine several additional specifications in which price dispersion is explained by right hand side variables comparable to those examined in the previous specifications. These measures are calculated at the $area \times product \times date$ level and include (1) the value of information (VOI), measured as $[E(p) - \min(p)]$, (2) range, (3) standard deviation, (4) the gap, measured as the difference between the second lowest and lowest price, and (5) the coefficient of variation. An observation for any of the aforementioned measures of price dispersion relates to all firms in a particular geographic area (A, B, C, or D) selling a given product on a given date. Table 6 displays OLS results using actual prices and Table 7 displays OLS results using residual prices (note that the coefficient of variation is not calculated for residual prices because the average residual price for any $area \times product \times date$ combination is zero). Consistent with our criteria for analyzing price dispersion in Table 3, we do not analyze $area \times product \times date$ combinations that solely consist of one or more of the 862 observations for which the residual price is zero by construction, as discussed earlier in this section. As a result, 794 of 2,474 $area \times product \times date$ combinations are excluded from this analysis, the vast majority of which can be classified as instances in which only one firm sells a particular product in an area on a given date. After this exclusion there are 1,680 $area \times product \times date$ combinations analyzed in Tables 6 and 7. In these specifications, we use the percentage of rivals for a given $area \times product \times date$ combination (RIV_PCT) as an explanatory variable.

Finally, the mere presence of rival firms may have ramifications for the existence of pure strategies for other firms in the same geographic area. For example, consider a scenario in which there are three firms; two rival firms and a third firm located at some distance from the rival firm

pair. Furthermore, suppose that consumers are heterogeneous in their search costs in the sense that a proportion of consumers incur a cost to search and some proportion does not incur any cost to search. From a theoretical standpoint, due to the fact that a pure strategy price equilibrium always exists for two or more non-capacity constrained symmetric firms (i.e., rival firms in our empirical context) with constant marginal costs facing homogeneous consumers, it can be shown that a pure strategy will also exist for the third firm.[21] Thus, we would expect less temporal dispersion in the presence of rivals than when they are absent. As the extent of price dispersion also depends on the number of firms, in Table 8 we provide summary statistics for our two principle measures of temporal price dispersion, $\ln(\hat{u}_{fgt}^2)$ (Equation (2)) and s_{gijt} (Equation (3)), according to the number of firms observed in a given $area \times product \times date$ for a given observation. Recall that when measuring $\ln(\hat{u}_{fgt}^2)$, observations are recorded at the $firm \times product \times date$ level whereas when measuring s_{gijt}, observations are recorded at the $firm\ pair \times product \times date$ level.

V.   Results and interpretation

We will now proceed to analyze the results of the estimation procedures discussed in the previous Section, and presented in Tables 3-8. We will discuss each of the proxies offered in Section 3 in turn.

---

[21] If consumers do not know the firms' current marginal cost and if search costs are sufficiently low for the proportion of consumers who find it costly to search, this may not hold (see Sherman and Weiss, forthcoming, for more details). If yet additional firms are located at a distance from a pair of rival firms, using similar reasoning one can also show that under certain circumstances a pure strategy will also exist for such firms as well. A formal proof of this result is beyond the scope of this study.

In light of the empirical researcher's difficulty in testing the non-monotonic relationship between search costs and price dispersion, it is useful empirically when we are able to directly compare a situation in which it is known that search is nearly costless for all customers with a situation in which some customers find it costly to search. Rival firms provide such a sound benchmark for a situation in which search is costless for all customers, since consumers at one such firm can observe posted prices at another such firm without travelling. And, indeed, we find that rival firms exhibit significantly less price dispersion than other firms across all specifications, and this finding is robust to every manner in which we measure price dispersion. In particular, residual prices are significantly less dispersed for rivals (Table 3) and rival pairs exhibit significantly lower spread volatility (measured by s_{gijt}) than other firm pairs (Table 4). Furthermore, rival pairs are predicted to charge identical prices with a substantially higher probability than other firm pairs (Table 4).

Examining Table 5 we notice that the size of the spread is significantly larger for non-rival firm×product pairs than for rival firm×product pairs. The most likely explanation for this is that rival firm pairs lack the ability to exercise market power based on geographic isolation alone, whereas a firm that does not face rivals is likely to charge a higher price than a firm that does face a rival. We keep this in mind when examining the results of our regression in which the occurrence of a rank reversal is our dependent variable in Table 5. Here we see that the difference between the predicted probability that rival firm×product pairs reverse rank versus the predicted probability that non-rival firm×product pairs reverse rank is approximately 0.02. Therefore, it appears that these two forces essentially offset one another. That is, the symmetry in market power across pairs of rivals and the similar prices charged by rival firm pairs as a result likely explains the small but positive and significant rank reversal estimate of 0.02.

Tables 6 and 7 provide additional support for the notion that the extent of price dispersion decreases in the percentage of firms in an area that are classified as rival firms. Finally, Table 8 provides evidence that less price dispersion is observed amongst other firms when rival firms compete in the same area, consistent with the idea that a pure strategy equilibrium will exist for any firm whose most proximate competition is a pair (or a cluster of any size) of rival firms.[22]

Our second set of proxies for search intensity is the day of the week, as the market becomes the most congested prior to the weekend, and particularly on Fridays. Therefore, we would suspect that there is less search on Fridays in all areas of the market. However, a difficulty with understanding our weekday variables in the context of theory is that we do not know the baseline proportion of searchers and non-searchers on the least costly or most costly search days. A similar issue arises when examining our geographic area variables (our third set of proxies); that is, we do not know the proportion of searchers and non-searchers in Area A, where average consumer income is presumably the highest and therefore the proportion of searchers is the lowest, nor do we know the proportion of searchers and non-searchers in Area D, where average consumer income is presumably the lowest.

With this in mind, we note that our weekday variables imply that price dispersion increases as search becomes more costly. F-tests indicate that residual prices are significantly more dispersed on Fridays than on any other weekday (Table 3) and that s_{gijt} is significantly

---

[22] We also checked these differences in a regression framework in which the relevant measure of price dispersion was regressed on a dummy indicating whether rivals sold in that area, the number of firms in the area N and N^2. The measure s_{gijt} was found to be 0.126 lower (with significance at the 0.01 level) when rivals sold in the same area and the measure $\ln(\hat{u}_{fgt}^2)$ was found to be 0.267 lower (with significance at the 0.05 level) when rivals sold in the same area. Consistent with all other analyses in the paper, errors were clustered at the date×product level.

higher on Fridays and holiday eves (Table 4).[23] Examining Table 5, we find that the probability of rank reversal increases on Fridays; this finding provides particularly strong evidence of increased price dispersion on Fridays in light of the fact that Table 5 also indicates that the size of the spread for firm×product pairs is significantly larger on Fridays. *F*-tests also indicate that prices are more dispersed on Fridays than on any other weekday across nearly all measures of price dispersion reported in Tables 6 and 7.

Examining the effect of geographic area on price dispersion, our results are not quite as robust across all specifications. On the one hand, in Table 4 we find that s_{gijt} is significantly lower in Areas C and D relative to Areas A and B, and in Table 5 we also find that the probability of rank reversal is lower in Areas C and D than in Areas A and B. These results would appear to provide further support for the notion that price dispersion increases with the proportion of consumers who possess high search costs. However, examining results reported in Table 3, *F*-tests do not reject equality of price dispersion in Areas A and D, despite the presumption that the highest percentage of consumers with high search costs shop in Area A. Results in Tables 6 and 7 are also mixed in this regard.

We find that price dispersion is concave in our fourth proxy, the number of firms *N* that post price in a given geographic area, increasing at a decreasing rate. However, endogeneity is a concern when examining the relationship between *N* and the extent of price dispersion, and

---

[23] We should also note that while product sold on Mondays through Thursdays is purchased in the early morning of the same day, product sold on Fridays is typically bought at the wholesale market on the previous day (recall that the market closes early on Fridays). Since most firms apparently follow the same practice in this regard every week, there should not be a concern that time-varying idiosyncratic quality differences amongst different vendors (rather than higher search costs due to more congestion) would be responsible for the high level of price dispersion we observe on Fridays.

therefore we refrain from drawing conclusions regarding the estimated magnitudes of the coefficients.

But possible endogeneity is not the only difficulty we encounter when interpreting these estimates. That is, let us consider how consumers search with the following example. In their model, Janssen and Moraga-González (2004) assume exogenous firm entry and allow for endogenous search in a model with two types of consumers with unit demand, one who searches costlessly ("fully-informed") and one who must pay a fixed cost for each price observed ("less-informed"). They find that an equilibrium in which every less-informed consumer searches for exactly one price does not exist as N becomes large because the expected price increases with N holding the search intensity constant. Since consumers' search behavior is endogenous, as N becomes large less-informed consumers randomize between not searching and searching for exactly one price, which results in increased price dispersion as N continues to increase. Therefore, in their model, endogenous consumer search leads to an increase in price dispersion as N continues to increase.

On the other hand, when consumers do not account for the fact that continued entry should lead to a lower likelihood of any given firm charging the lowest price in the market given symmetric mixed strategies, previous theoretical models in which search is exogenous (such as Rosenthal (1980) and Varian (1980)) imply that the expected price in the market will increase as a higher proportion of firms target customers that do not search by charging the monopoly price. If this is the case, then price dispersion should first increase and then decrease as N becomes large due to the fact that nearly all firms will eventually charge the monopoly price in equilibrium. Unfortunately, the extent to which consumers internalize the effect of an increase in the number of firms on their expected gains from search is difficult to ascertain in our setting.

Furthermore, additional complications arise when seeking to interpret the relationship between the intensity of competition and price dispersion. Theoretical models predict that the number of firms at which price dispersion is maximized will depend on several parameter values that are difficult to calibrate empirically. We illustrate this difficulty by calculating the simulated values of the variance of price $(Var(p))$ implied by the Varian (1980) model, replacing his number of informed consumers with a proportion $0 \leq \lambda \leq 1$ of informed consumers, as done in several previous studies. In his model, consumers have unit demand and a reservation price r, each firm has a fixed cost k and a constant marginal cost c, and the market will support the maximum number of firms N such that profits are non-negative. It is natural to relate $Var(p)$ implied by the Varian (1980) model and one of our estimated measures of price dispersion, the variance of the error term in (1), $\sigma^2_{u_{fgt}}$(which we estimate using $\hat{u}^2_{fgt}$, as noted in Section 4). The numerical simulation presented in Table 9 illustrates the non-monotonicity implied by Varian's model with respect to the relationship between $Var(p)$ and N and between $Var(p)$ and λ. For given values of c, k, and r, one may observe in the table that $Var(p)$ in Varian's (1980) model may peak anywhere between $N = 3$ and $N = 8$ with associated proportions of informed consumers anywhere between 0.92 and 0.4.

Turning to our fifth proxy for search costs, we see that across nearly all specifications we estimate that price dispersion increases with the wholesale price, which is the primary component of firms' marginal cost. The relationship between marginal cost and the extent of price dispersion also serves to illustrate the challenge of linking the assumptions of search-theoretic models with the sophistication of consumers in a given empirical setting. Due to the fact that wholesale prices may change from one day to the next in our empirical environment, consumers presumably must form expectations regarding the current day's wholesale price based

on prices in previous periods when determining whether or not to search. However, when no pure strategy equilibrium exists, information asymmetry between the consumer and the firm complicates the relationship between the wholesale price, the primary component of the firm's marginal cost, and price dispersion.

Along these lines, Tappata (2009) extends the Varian (1980) framework in which there is no pure strategy equilibrium to incorporate endogenous consumer search and consumer uncertainty over production cost, allowing the static game conceived by Varian (1980) to be repeated over time. He notes that in equilibrium there is a negative relationship between current period marginal cost and the extent of price dispersion given any demand function for which the monopolist's absolute markup decreases with marginal cost. In a dynamic environment, however, he notes that the proportion of shoppers and non-shoppers in the current period is affected by consumers' expectations of the current period's production cost. For example, assume that firms and consumers interact in a simple two cost framework where costs are Markov-switching, and that $c_t = c_{t-1}$ with a probability strictly greater than 0.5, as in Tappata (2009). Now, suppose that cost was high in period $t-2$ and that cost was low in period $t-1$. Then in a dynamic environment consumers would expect an increase in price dispersion in the current period. As a result, we would expect a higher proportion of consumers to be shoppers due to greater gains from search in the current period. However, since the relationship between the proportion of shoppers and price dispersion is non-monotonic in many search-theoretic models (including in Varian (1980)), the effect of an increase in shoppers on price dispersion would depend on whether the proportion of shoppers was relatively low or high in period $t-1$.

In addition, if high cost is realized in the current period, then the extent of price dispersion would decrease given unit demand or linear demand due to the smaller difference

26

between the monopoly price (or reservation price) and marginal cost, but the opposite would occur given logarithmic demand (see Tappata (2006) and Tappata (2009) for a complete exposition). Additionally, these outcomes will be certain only if consumers can relate the magnitude of costs with the size of firms' margins, which requires consumers to know whether, for example, margins decrease with marginal costs (e.g. linear demand) or whether margins increase with marginal costs (e.g. logarithmic demand).

Taken together, in order to be able to predict current period price dispersion, the researcher and the consumer would be required to (1) know the nature of demand, (2) account for costs in at least the two previous periods, and (3) be able to surmise the baseline proportion of shoppers in multiple periods, noting that the effect of changes in the proportion of shoppers to non-shoppers on price dispersion is possibly non-monotonic in any given period. In addition, given that we examine several products in our empirical setting, one could not compare the wholesale price for one product with another without accounting for the fact that margins typically differ across products. Therefore, given our environment, an empirical analysis in which one were to analyze the response of the variance of price to cost changes given consumer information lags and search cost heterogeneity would be challenging.

Finally, we briefly comment on the coefficient estimates related to our final proxies, the product dummies (not reported). Generally speaking, there does not appear to be an obvious interpretation stemming from these findings. For example, across specifications we do not detect a clear distinction between the extent of price dispersion for the most expensive products by weight relative to the least expensive products (as shown in Table 1), although the extent of price dispersion observed for sweet potatoes, the most expensive product by weight, tends to be amongst the highest of all products in our residual price specifications in Table 4 and in Tables 6

and 7. We also find that equal prices across firm pairs are significantly more likely to be charged for cucumbers than for any other product, the absolute value of the spread of firm pairs is the lowest for cucumbers, and therefore, not surprisingly, we find that rank reversals are most likely for cucumbers. Cucumbers were purchased more frequently by the average Israeli household than any other product examined in our dataset (Israeli Central Bureau of Statistics, 2009 Household Expenditure Survey), and therefore this result is consistent with Sorensen's (2000) finding that prices for retail prescription drugs are relatively less dispersed for products that are purchased more frequently.

VI.   Conclusion

In this paper we have exploited various ways in which consumers vary in their search intensity in the Mahane Yehuda market in order to investigate the relationship between consumer search costs and price dispersion. In particular, we find evidence that price dispersion increases with the cost of search. We should note, though, that most empirical environments, including Mahane Yehuda, are constrained in their ability to provide a direct test of many search-theoretic models due to data limitations and the researcher's knowledge of how consumers decide to search. While we believe that it is important that empiricists continue to seek out novel environments in which consumers vary in their search costs, such limitations would also seem to highlight the appeal of experiments that can address the most recent developments in consumer search theory in a more controlled environment. In particular, ideally such experiments would address whether consumers behave in a manner that is consistent with the search assumptions found in recent search-theoretic models.

References

Amemiya, T., 1985. *Advanced Econometrics*. (Cambridge: Harvard University Press, 1985).

De los Santos, B., A. Hortaçsu, and M.R. Wildenbeest, 2012. Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior, *American Economic Review* 102 (6): 2955-2980.

Barron, J.M., B.A. Taylor, and J.R. Umbeck, 2004. Number of Sellers, Average Prices, and Price Dispersion. *International Journal of Industrial Organization* 22: 1041-1066.

Baye, M., J. Morgan, and P. Scholten, 2006. 'Information, Search, and Price Dispersion,' in *Handbook of Economics and Information Systems,* T. Hendershott, ed. (Amsterdam: Elsevier Press, 2006)

Borenstein, S. and N.L. Rose, 1994. Competition and Price Dispersion in the U.S. Airline Industry. *Journal of Political Economy* 102 (4): 653-683.

Burdett K. and K. Judd, 1983. Equilibrium Price Dispersion. *Econometrica* 51 (4): 955-969.

Carlson, J.A. and R.P. McAfee, 1983. Discrete Equilibrium Price Dispersion. *Journal of Political Economy* 91 (3): 480-493.

Chandra, A. and M. Tappata, 2011. Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets. *RAND Journal of Economics* 42 (4): 681-704.

Dana, J.D., 1994. Learning in an Equilibrium Search Model. *International Economic Review* 35 (3): 745-771.

Dubois, P. and H. Perrone, 2012. Price Dispersion and Search Costs: The Roles of Imperfect Information and Product Differentiation, Working paper.

Genesove, D., 1995. Search at Wholesale Auto Auctions. *The Quarterly Journal of Economics* 110 (1): 23-49.

Gerardi, K., and A. Shapiro, 2009. Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry. *Journal of Political Economy* 117 (1): 1-37.

Harvey, A.C., 1976. Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica* 44 (3): 461-465.

Hosken, D., R. McMillan, and C. Taylor, 2008. Retail Gasoline Pricing: What Do We Know? *International Journal of Industrial Organization* 26: 1425-1436.

Janssen, M. and J.L. Moraga-González, 2004. Strategic Pricing, Consumer Search and the Number of Firms. *Review of Economic Studies* 71(4): 1089-1118.

Janssen, M., J.L. Moraga-González, and M. Wildenbeest, 2005. Truly Costly Sequential Search and Oligopolistic Pricing. *International Journal of Industrial Organization* 23: 451-466.

Janssen M., P. Pichler, and S. Weidenholzer, 2011. Oligopolistic Markets with Sequential Search and Production Cost Uncertainty. *RAND Journal of Economics* 42 (3): 444-470.

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

Lach, S. and J.L. Moraga-González, 2011. Heterogeneous Price Information and the Effect of Competition, Working Paper.

Lewis, M., 2008. Price Dispersion and Competition with Differentiated Sellers. *The Journal of Industrial Economics* (3): 654-678.

Orlov, E., 2011. How Does the Internet Influence Price Dispersion? Evidence from the Airline Industry. *The Journal of Industrial Economics* (1): 21-37.

Pennerstorfer, D., P. Schmidt-Dengler, N. Schutz, C. Weiss, and B. Yontcheva, 2014. Information and Price Dispersion: Evidence from Retail Gasoline, Working Paper.

Png, I.P.L. and D. Reitman, 1994. Service Time Competition. *RAND Journal of Economics* 25(4): 619-634.

Rosenthal, R.W., 1980. A Model in Which an Increase in the Number of Sellers Leads to a Higher Price. *Econometrica* 48 (6): 1575-1579.

Salop, S.C. and J.E. Stiglitz, 1977. Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. *Review of Economic Studies* 44: 493-510.

Sherman, J. and A. Weiss, forthcoming. Price Response, Asymmetric Information, and Competition. *The Economic Journal*.

Shilony, Y., 1977. Mixed Pricing in Oligopoly. *Journal of Economic Theory* 14: 373-388.

Sorensen, A., 2000. Equilibrium Price Dispersion in Retail Markets for Prescription Drugs. *Journal of Political Economy* 108 (4): 833-850.

Stahl, D.O., 1989. Oligopolistic Pricing with Sequential Consumer Search. *American Economic Review* 79: 700-712.

Stigler, G., 1961. The Economics of Information. *Journal of Political Economy* 69 (3): 213-225.

Tang., Z., M.D. Smith, and A. Montgomery, 2010. The Impact of Shopbot Use on Prices and Price Dispersion: Evidence from Online Book Retailing. *International Journal of Industrial Organization* 28 (2010): 579-590.

Tappata, M., 2009. Rockets and Feathers: Understanding Asymmetric Pricing. *The RAND Journal of Economics* 40 (4): 673-687.

Varian, H.R., 1980. A Model of Sales. *American Economic Review* 70: 651-659.

Wildenbeest, M.R., 2011. An Empirical Model of Search with Vertically Differentiated Products. *RAND Journal of Economics* 42 (4): 729-757.

Zhou, Y., 2006. Price Dispersion in the Grocery Market. *Journal of Business* 79 (3): 1175-1192.

Table 1: Average price per kilogram by product and area

Product / Area	A	B	C	D	Overall
Beet	5.06	4.47	4.16	3.2	4.19
St. Dev.	(2.01)	(1.19)	(1.35)	(0.95)	(1.33)
Obs.	37	424	204	145	810
Cauliflower	6.38	4.68	4.5	4.43	4.77
St. Dev.	(2.23)	(2.13)	(2.21)	(1.14)	(2.13)
Obs.	16	20	62	21	119
Corn	5.00	4.91	4.22	4.03	4.45
St. Dev.	(0.97)	(1.35)	(0.86)	(0.77)	(0.99)
Obs.	93	21	224	26	364
Cucumber	4.85	3.8	3.99	2.99	3.94
St. Dev.	(0.84)	(0.85)	(1.22)	(0.84)	(1.11)
Obs.	362	673	491	280	1,806
Kohlrabi	5.51	5.1	4.97	3.54	4.89
St. Dev.	(1.73)	(1.14)	(1.18)	(1.02)	(1.29)
Obs.	46	392	210	94	742
Red Cabbage	3.66	3.56	2.88	2.99	3.28
St. Dev.	(1.79)	(1.44)	(1.14)	(1.00)	(1.38)
Obs.	32	84	64	37	217
Red Pepper	7	6.12	4.64	4.22	5.85
St. Dev.	(1.31)	(1.43)	(1.18)	(1.14)	(1.64)
Obs.	408	573	364	110	1,455
Sweet Potato	7.96	9	4.39	4.84	7.48
St. Dev.	(1.60)	(1.82)	(1.09)	(0.83)	(2.49)
Obs.	218	434	198	58	908
White Cabbage	2.82	3.62	2.56	2.8	2.95
St. Dev.	(0.74)	(1.39)	(1.03)	(0.85)	(1.18)
Obs.	32	154	193	147	526
Yellow Pepper	7.15	5.83	4.67	4.01	5.63
St. Dev.	(1.18)	(1.65)	(1.65)	(1.09)	(1.79)
Obs.	93	284	110	57	542
Overall	6.15	5.42	4.16	3.41	4.93
St. Dev.	(1.86)	(2.17)	(1.38)	(1.12)	(2.02)
Obs.	1,337	3,059	2,120	975	7,491

Note: Prices are in shekels per kilogram. During the data collection period in 2009, on average 4.01 shekels = $1.00.

Table 2: Frequency of number of firms N posting price in a given area by product

	1	2	3	4	5	6	7	8	9	10	11
Beet	70	67	51	35	22	14	10	5	1		
Cauliflower	73	17	4								
Corn	64	57	19	23	5	2					
Cucumber	3	19	45	64	63	62	32	19	24	4	5
Kohlrabi	85	60	46	28	27	17	6	1			
Red Cabbage	94	37	15	1							
Red Pepper	50	29	23	67	65	42	25	22	8	1	
Sweet Potato	74	70	60	43	19	17	15	5			
White Cabbage	102	99	56	12	2						
Yellow Pepper	151	55	26	15	12	7	5	1			

Table 3: Residual price dispersion:
Observations recorded at the firm×product×date level

Dep Var: $\ln(\hat{u}_{fgt}^2)$

	All observations		Excluding rival observations	
	OLS	Prais-Winsten	OLS	Prais-Winsten
Rival	-0.263***	-0.248***		
	(0.068)	(0.076)		
N	0.374***	0.406***	0.300**	0.361***
	(0.109)	(0.114)	(0.128)	(0.137)
N^2	-0.030***	-0.033***	-0.026**	-0.031**
	(0.010)	(0.010)	(0.012)	(0.013)
Mon	0.094	0.095	0.050	0.060
	(0.113)	(0.111)	(0.147)	(0.142)
Tues	0.170	0.161*	0.113	0.118
	(0.109)	(0.098)	(0.134)	(0.119)
Thurs	0.158	0.151	0.101	0.102
	(0.120)	(0.113)	(0.144)	(0.130)
Fri	0.624***	0.631***	0.571***	0.590***
	(0.115)	(0.111)	(0.143)	(0.135)
Hol Eve	0.413*	0.368*	0.368	0.341
	(0.215)	(0.197)	(0.244)	(0.231)
Area A	-0.324**	-0.364**	-0.323**	-0.331*
	(0.130)	(0.145)	(0.159)	(0.176)
Area C	-0.470***	-0.490***	-0.733***	-0.746***
	(0.106)	(0.119)	(0.122)	(0.135)
Area D	-0.389**	-0.412**	-0.507**	-0.515**
	(0.167)	(0.186)	(0.198)	(0.220)
Whl	0.287***	0.271***	0.299***	0.287***
	(0.035)	(0.039)	(0.045)	(0.049)
Constant	-4.658***	-4.668***	-4.212***	-4.320***
	(0.367)	(0.390)	(0.431)	(0.455)
Obs.	6,614	6,614	4,413	4,413
R^2	0.108	0.100	0.108	0.102

Note: Nine product dummy estimates not reported. Cluster-robust standard errors at the date×product level are reported for the OLS regressions. Clustered semi-robust standard errors at the date×product level are reported for the Prais-Winsten regressions. Fifteen observations for which the estimated residual is exactly equal to zero are excluded from each of the specifications reported in this table.

(***) indicates significance at the .01 level   (**) indicates significance at the .05 level   (*) indicates significance at the .1 level

Table 4: Firm pair price dispersion specifications:
Observations recorded at the firm pair×product×date level within each geographic area

	Dep Var: $s_{gijt} \left| spread_{gijt} - \overline{spread}_{gij} \right|$		Dep Var: Equality of firm pair prices (0,1)	
	OLS	Prais-Winsten	Logit	Average marginal effects
Rival	-0.154***	-0.139***	1.127***	0.207***
	(0.017)	(0.022)	(0.065)	(0.014)
N	0.001	-0.003	-0.288***	-0.003
	(0.028)	(0.028)	(0.077)	(0.004)
N^2	-0.001	-0.001	0.022***	
	(0.002)	(0.002)	(0.006)	
Mon	0.046	0.047	0.072	0.011
	(0.033)	(0.030)	(0.093)	(0.014)
Tues	0.064*	0.054**	-0.071	-0.010
	(0.032)	(0.026)	(0.091)	(0.013)
Thurs	0.049	0.044	0.057	0.008
	(0.034)	(0.029)	(0.103)	(0.015)
Fri	0.182***	0.176***	-0.169*	-0.024*
	(0.039)	(0.036)	(0.091)	(0.013)
Hol Eve	0.155***	0.127***	-0.161	-0.022
	(0.051)	(0.043)	(0.148)	(0.020)
Area A	-0.057	-0.073	0.422***	0.066**
	(0.042)	(0.048)	(0.152)	(0.026)
Area C	-0.124***	-0.146***	0.179*	0.027
	(0.029)	(0.033)	(0.107)	(0.016)
Area D	-0.134***	-0.169***	-0.031	-0.004
	(0.045)	(0.047)	(0.167)	(0.024)
Whl	0.099***	0.088***	-0.042*	-0.006*
	(0.011)	(0.013)	(0.024)	(0.003)
Constant	0.221**	0.283***	0.398	
	(0.102)	(0.103)	(0.280)	
Obs.	12,669	12,669	12,669	
R^2	0.142	0.128		

Note: Nine product dummy estimates not reported. Cluster-robust standard errors at the date×product level are reported for the OLS regressions. Clustered semi-robust standard errors at the date×product level are reported for the Prais-Winsten regressions. Delta-method standard errors are reported for the marginal effects. Marginal effects are for a discrete change from 0 to 1 for all indicator variables.

(***) indicates significance at the .01 level   (**) indicates significance at the .05 level   (*) indicates significance at the .1 level

Table 5: Firm pair price dispersion specifications:
Observations recorded at the firm pair×product×date level within each geographic area

| | Dep Var: $\lvert spread_{gijt} \rvert$ | | Dep Var: Rank Reversals (0,1) | |
	OLS	Prais-Winsten	Logit	Average marginal effects
Rival	-0.455***	-0.395***	0.172**	0.021**
	(0.022)	(0.034)	(0.082)	(0.010)
N	0.085	0.025	-0.174*	-0.001
	(0.034)	(0.032)	(0.100)	(0.003)
N^2	-0.009***	-0.003	0.013	
	(0.003)	(0.002)	(0.008)	
Mon	0.044	0.054	0.041	0.005
	(0.044)	(0.034)	(0.110)	(0.013)
Tues	0.089**	0.068***	0.149	0.018
	(0.040)	(0.024)	(0.110)	(0.013)
Thurs	0.041	0.026	0.044	0.005
	(0.038)	(0.030)	(0.116)	(0.014)
Fri	0.193***	0.184***	0.256**	0.031**
	(0.046)	(0.039)	(0.110)	(0.014)
Hol Eve	0.236***	0.192***	0.304**	0.039*
	(0.066)	(0.047)	(0.143)	(0.020)
Area A	-0.271***	-0.222***	0.184	0.022
	(0.061)	(0.071)	(0.121)	(0.015)
Area C	0.011	-0.009	-0.355***	-0.038***
	(0.097)	(0.062)	(0.111)	(0.011)
Area D	-0.209***	-0.230***	-0.366*	-0.037**
	(0.067)	(0.069)	(0.201)	(0.018)
Whl	0.055***	0.027	0.190***	0.022***
	(0.014)	(0.016)	(0.026)	(0.003)
Constant	0.053	0.244*	-1.418***	
	(0.140)	(0.136)	(0.339)	
Obs.	12,669	12,669	12,669	
R^2	0.150	0.142		

Note: Nine product dummy estimates not reported. Cluster-robust standard errors at the date×product level are reported for the OLS regressions. Clustered semi-robust standard errors at the date×product level are reported for the Prais-Winsten regressions. Delta-method standard errors are reported for marginal effects. Marginal effects are for a discrete change from 0 to 1 for all indicator variables.

(***) indicates significance at the .01 level   (**) indicates significance at the .05 level   (*) indicates significance at the .1 level

Table 6: Alternative measures of price dispersion (using actual prices): Observations recorded at the area×product×date level

	Dependent Variable				
	VOI	*Range*	*Standard Deviation*	*Gap*	*Coefficient of Variation*
Rival_Pct	-0.199***	-0.397***	-0.227***	-0.186**	-0.039***
	(0.049)	(0.095)	(0.053)	(0.072)	(0.012)
N	0.331***	0.799***	0.199***	-0.049	0.039***
	(0.031)	(0.063)	(0.034)	(0.047)	(0.007)
N^2	-0.021***	-0.056***	-0.019***	-0.001	-0.004***
	(0.003)	(0.006)	(0.003)	(0.004)	(0.001)
Mon	0.013	0.041	0.033	0.007	0.007
	(0.038)	(0.069)	(0.037)	(0.055)	(0.008)
Tues	0.050	0.122	0.078*	0.048	0.014*
	(0.040)	(0.074)	(0.040)	(0.059)	(0.008)
Thurs	0.035	0.061	0.033	0.023	0.004
	(0.037)	(0.066)	(0.036)	(0.058)	(0.008)
Fri	0.153***	0.324***	0.153***	0.102*	0.037***
	(0.038)	(0.074)	(0.039)	(0.056)	(0.009)
Hol Eve	0.076	0.200*	0.096	0.050	0.023
	(0.054)	(0.105)	(0.063)	(0.083)	(0.015)
Area A	-0.263***	-0.630***	-0.285***	-0.162**	-0.065***
	(0.055)	(0.111)	(0.059)	(0.080)	(0.011)
Area C	-0.275***	-0.609***	-0.300***	-0.253***	-0.008
	(0.042)	(0.087)	(0.045)	(0.058)	(0.009)
Area D	-0.245***	-0.555***	-0.323***	-0.258***	-0.019
	(0.052)	(0.103)	(0.056)	(0.075)	(0.013)
Whl	0.070***	0.112***	0.051***	0.068***	-0.006**
	(0.015)	(0.025)	(0.013)	(0.018)	(0.003)
Constant	-0.260**	-0.551**	0.371***	0.920***	0.133***
	(0.129)	(0.241)	(0.129)	(0.180)	(0.026)
Obs.	1,680	1,680	1,680	1,680	1,680
R^2	0.401	0.445	0.296	0.149	0.167

Note: Nine product dummy estimates not reported. Cluster-robust standard errors at the date×product level are reported for the OLS regressions. Instances of only one firm×product posting price in a given area on a given date are excluded.

(***) indicates significance at the .01 level   (**) indicates significance at the .05 level   (*) indicates significance at the .1 level

Table 7: Alternative measures of price dispersion (using residual prices): Observations recorded at the area×product×date level

	Dependent Variable			
	VOI	*Range*	*Standard Deviation*	*Gap*
Rival_Pct	-0.081**	-0.147**	-0.102***	-0.119**
	(0.036)	(0.068)	(0.036)	(0.056)
N	0.215***	0.414***	0.073***	-0.039
	(0.024)	(0.044)	(0.023)	(0.036)
N^2	-0.012***	-0.023***	-0.006***	0.002
	(0.002)	(0.004)	(0.002)	(0.003)
Mon	0.024	0.055	0.011	-0.017
	(0.034)	(0.059)	(0.031)	(0.050)
Tues	0.039	0.101*	0.041	0.012
	(0.034)	(0.059)	(0.030)	(0.050)
Thurs	0.057*	0.118*	0.041	-0.017
	(0.033)	(0.060)	(0.031)	(0.049)
Fri	0.187***	0.366***	0.168***	0.123**
	(0.038)	(0.068)	(0.033)	(0.054)
Hol Eve	0.057	0.140	0.041	0.063
	(0.057)	(0.093)	(0.054)	(0.096)
Area A	-0.066	-0.106	-0.041	0.145**
	(0.041)	(0.077)	(0.040)	(0.064)
Area C	-0.159***	-0.297***	-0.132***	0.005
	(0.030)	(0.052)	(0.027)	(0.045)
Area D	-0.082**	-0.160**	-0.093**	0.037
	(0.039)	(0.073)	(0.039)	(0.063)
Whl	0.109***	0.202***	0.090***	0.098***
	(0.012)	(0.022)	(0.010)	(0.016)
Constant	-0.428***	-0.767***	0.079	0.158
	(0.100)	(0.180)	(0.089)	(0.141)
Obs.	1,680	1,680	1,680	1,680
R^2	0.373	0.410	0.263	0.112

Note: Nine product dummy estimates not reported. Cluster-robust standard errors at the date×product level are reported for the OLS regressions. Instances of only one firm×product posting price in a given area on a given date are excluded. Coefficient of variation is not calculated for residual prices because the average value of the residual for a given area×date×product combination is zero.

(***) indicates significance at the .01 level   (**) indicates significance at the .05 level   (*) indicates significance at the .1 level

Table 8: Effect of the existence of rivals in a given area (for a given product) on the extent of price dispersion of other firms in the same area (for the same product)

Number of firms N posting price in a given area by product

	2	3	4	5	6	7	8
$\ln(\hat{u}_{fgt}^2)$							
Rivals exist							
Average	-3.901	-3.152	-2.889	-2.703	-2.805	-2.535	-2.41
St. Dev.	(2.457)	(2.496)	(2.363)	(2.442)	(2.371)	(2.274)	(2.290)
Obs.	346	473	745	870	850	618	399
No rivals exist							
Average	-3.335	-2.994	-2.969	-2.272	-2.498	-2.225	-1.562
St. Dev.	(2.465)	(2.418)	(2.818)	(2.475)	(2.677)	(2.696)	(2.474)
Obs.	616	538	395	198	113	28	24
s_{gijt}							
Rivals exist							
Average	0.519	0.600	0.633	0.700	0.614	0.705	0.771
St. Dev.	(0.647)	(0.667)	(0.661)	(0.687)	(0.579)	(0.698)	(0.736)
Obs.	169	449	1,097	1,709	2,103	1,831	1,386
No rivals exist							
Average	0.654	0.651	0.712	0.870	0.772	0.879	0.973
St. Dev.	(0.696)	(0.757)	(0.775)	(0.885)	(0.714)	(0.778)	(0.714)
Obs.	303	530	582	387	279	84	84

Note: Smaller magnitudes imply less price dispersion. For example, in the case of $\ln(\hat{u}_{fgt}^2)$, a value of -2 implies less price dispersion than a value of -1. Rivals exist for all observations associated with $N > 8$.

Table 9: Maximum allowable λ according to N and associated $Var(p)$
$(r = 3, c = 0)$

k/r		$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$	$N = 7$	$N = 8$	$N = 9$	$N = 10$
0.01	Max λ	0.98	0.97	0.96	0.95	0.94	0.93	0.92	0.91	0.9
	Var (p)	0.0711	0.5170	0.9811	1.2844	1.4518	1.530	1.5544	1.5461	1.5187
0.05	Max λ	0.9	0.85	0.8	0.75	0.7	0.65	0.6	0.55	0.5
	Var (p)	0.2329	0.7642	1.0553	1.1529	1.1507	1.1009	1.0295	0.9491	0.8658
0.10	Max λ	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	$\lambda = 0$
	Var (p)	0.3211	0.7389	0.8487	0.7968	0.6746	0.5190	0.3424	0.1493	
0.20	Max λ	0.6	0.4	0.2	$\lambda = 0$	$\lambda < 0$	$\lambda < 0$	$\lambda < 0$	$\lambda < 0$	$\lambda < 0$
	Var (p)	0.3282	0.4330	0.2291						

Figure 1: Map of Mahane Yehuda[24]

Buses run along
Jaffa Road

Area D

Map of Mahane Yehuda
© 2008, 2009 all rights reserved

To get the latest version of this map go to
http://israeltours.wordpress.com/

Restaurant/Cafe
① HaAgas 1
② Ichakanda
③ Emil
④ Zedkiyahu
⑤ Tzakhko
⑥ Azura
⑦ Mizrachi
⑧ Rachmo's
⑨ Topolino

Bakery
❶ Teller
❷ B'khira Tiv'it
❸ Marzipan
❹ HaOfeh
❺ Pe'er
❻ Tusya

Clothes
✿ Bordeaux
✳ 30 HaEgoz Avi Ben
✳ Vintage

Central bus station
ATM
Nisan-coffee, tea, spice
Area C
Iraqi Shuq
pharmacy
Jaffa Road
Bakery-pizza
Pickles
Ez HaHaim
Halva&baklava
Pickles
Area B
Halapuach
Danai
HaAdai
HaEshkol
Bakeries
HaAfarsek
JOAD
MaZetim
HaShezif
AbuDaud
Basher
Parking Lot
Area A
HaShaked
Gifts&clothes
Yemenite food&breads
mushrooms
Hummus
Hgiva
Pereg spices
petfood
HaDavon
burekas
Fri HaAdama
Ceramics
Avinoam
KippaMan
Aviv
Agrippas
English Cafe
Ezra Kahel
TexsNet organic produce
HaLemon
Rika Moroccan Gallery
Olam HaTeva

Exotic fruit
Flowers
Felafel
Gifts&clothes
Wine
Kikar Davidka
King George/Straus
Juice
Ice cream
Uzi Eli, Etrog man
Shuq on Thu & Fri
Synagogue
Cheese+
Basher-cheese, wine, chocolates, salads
Zedkiyahu-cheese, olives, salads
MaZetim-cheese, olives, wine
AbuDaud-cheese&deli
Ethiopian
Roasted nuts
Dried fruit
Parking Lot
Parking Lot
September 25, 2009

For guided tours contact
shmuel.browns@gmail.com
052 311-1265

---

[24] We thank Shmuel Browns for providing us with this hand-drawn map. We have added notation to indicate the locations of Areas A-D, the bus lines, and additional parking lots.