

An Empirical Test of the Theory of Sales: Do Household Storage Costs Affect Consumer and Store Behavior?

David R. Bell *

and

Christian A. L. Hilber **

This Version: August 12, 2004

We thank seminar participants at the Marketing Science Conference 2003 at College Park and at the Tjalling C. Koopmans Workshop “Issues on the Economics of Pricing” 2004 at the Utrecht University for helpful comments. We also thank Sangyoung Song for research assistance. Any errors, of course, are our own.

* The Wharton School, University of Pennsylvania, 700 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104, davidb@wharton.upenn.edu.

** London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom, c.hilber@lse.ac.uk

An Empirical Test of the Theory of Sales: Do Household Storage Costs Affect Consumer and Store Behavior?

Abstract

We revisit and test Salop and Stiglitz (1982) Theory of Sales. Equilibrium predictions are that higher consumer storage costs lead to: (1) higher average prices, (2) fewer promotions, and (3) shallower promotions. Empirical estimates of storage cost are developed for approximately 1,000 households using the American Housing Survey (1989), United States Census (1990), and Stanford Market Basket Database (1991-1993). A test of the key assumption finds consumers with higher storage costs shop more often and purchase smaller quantities per visit; moreover, all three equilibrium predictions are supported. The estimated quantitative effects on shopping frequency and prices are economically important.

KEY WORDS: *Consumer Behavior, Retail Prices, Price Promotion, Storage Costs.*

JEL CLASSIFICATION: D12, D40, M3.

Price variation for homogenous goods is a commonly observed feature of many product markets. Typical explanations relate to differences in firm characteristics (Simester 1995), or some facet of consumer heterogeneity in price knowledge or willingness to undertake effort to obtain lower prices (e.g., Narasimhan 1984; Varian 1980). This gives rise to many important questions regarding how consumers might respond to disparate prices, and implicitly, how firms should set prices in a given market context. It is usual for empirical studies to isolate one side of this issue and address in detail, either: (1) consumer response to a particular pricing profile, or (2) equilibrium firm behavior given assumed market characteristics and consumer behaviors, but not both. One relatively overlooked theoretical conjecture for price variation stems from Salop and Stiglitz (1982) who show that the presence of consumer storage costs is sufficient to generate price variation, specifically a two-price equilibrium in which ex ante identical firms charge different prices. In this paper we investigate whether the presence of storage costs influences the purchasing behavior of consumers, and if so, how this in turn affects the pricing decisions of retailers. While such issues are of theoretical interest, it is also possible that the economic consequences for consumers and firms are quantitatively important.

In order to identify instances where storage costs might affect consumer behavior, we first delineate conditions under which consumers might be differentially motivated to both purchase excess inventory and to hold it for future consumption. Implicit in most empirical marketing studies of consumables is the notion that consumers are largely unconstrained in their response to price variation – that is, consumer response is predominantly a function of observed prices. Our contention however is purchase and consumption patterns may differ across consumers as follows. Households who are constrained in their ability to store products (i.e., they have prohibitively high storage costs) may exhibit purchase patterns such that the observed purchase volume does not vary much between promotion (low) and regular (high) price points. Conversely, households with very low storage costs are able to stockpile should they be confronted with a low price. This stockpiling behavior leads to a longer time until these consumers return to the market.¹

Some empirical evidence for this stylized fact – that price responsiveness of consumers is related to their cost of storage – is provided in Bucklin and Gupta (1992). They show price sensitivity in purchase incidence decisions for liquid laundry detergent is much higher for consumers who live in houses than for those who live in apartments. They interpret this finding

to imply that house dwellers are able to manage inventory and stockpile, while apartment dwellers have relatively prohibitive costs of storage. Hendel and Nevo (2003) also analyze detergent purchases (but using data from a different metro area and time period) and find that larger households and households in suburban locations hold relatively larger volumes of inventory. They also conclude that these “low storage cost” types of households can benefit from non-linear pricing schedules in which larger package sizes are offered at a per unit discount.

Price Response by Consumers. The nature of consumer response to price has generated a vast literature in marketing. Empirical research in this stream includes seminal work by Gupta (1988) on the decomposition of promotion response through a model that accounts for the separate behavioral elements of purchase incidence, brand switching, and purchase quantity. Gupta shows that approximately 85% of the effect of a price cut on ground coffee occurs on secondary demand (brand switching). Chiang (1991) and Chintagunta (1993) address the same behaviors and find similar substantive results, however their model specifications allow for co-variation among the three decisions and for unobserved consumer heterogeneity, respectively. Bucklin, Gupta and Siddarth (1998) use a finite mixture model to provide evidence of distinct market segments with very different stockpiling profiles.

A number of others have addressed these same issues with aggregate data (e.g., Pauwels, Hanssens, and Siddarth 2002; van Heerde, Leeflang, and Wittink 2000; van Heerde, Gupta, and Wittink 2003) and have found that short and long term response decompositions can be very different (the latter is more likely to show a heavier emphasis on primary demand effects). Overall, the weight of evidence suggests that the possibility of primary demand expansion through stockpiling and purchase acceleration may be considerably higher than previously thought. Analytical results and empirical findings in van Heerde, Gupta, and Wittink (2003) imply that sales expansion resulting from price discounting is not necessarily correlated with the magnitude of the price elasticity for the individual behavioral components. That is to say, while the largest short term price elasticity is typically that for brand switching (Bell, Chiang, and Padmanabhan 1999; Chiang 1991; Chintagunta 1993; Gupta 1988), the predominant component of the increase in *sales* may come from purchase acceleration and stockpiling.

Collectively, these findings point to the importance of stockpiling as a key driver of price response. As such, it seems necessary to advance the state of the literature with respect to

determinants of stockpiling behavior at the consumer level, particularly in light of the dearth of research that addresses this fundamental issue directly. In fact, a common implicit assumption in most empirical studies in marketing is that the observed purchase response (e.g., the volume of product purchased at a particular price point) is a function of price only.² A deeper investigation suggests that storage costs may play a significant role in the behavior of consumers, and also influence the equilibrium price-setting of firms. In this paper we argue that some consumers may also operate under real external constraints, specifically their ability to take advantage of low prices may be hindered or furthered by their capacity to store excess inventory.

Price Setting by Firms. A number of other authors have examined the second half of the consumer/firm dichotomy to provide insights into why price variation arises in a price-setting equilibrium. In a seminal paper that derives retail prices as outcomes of mixed strategy equilibria, Varian (1980) assumes the presence of informed and uninformed consumers to obtain his results. In related studies, price variation for homogenous goods can arise due to: (1) the presence of price discrimination through mechanisms such as coupons (Narasimhan 1984), and (2) heterogeneity in consumer brand loyalty (Raju, Lal, and Srinivasan 1990). Lal and Matutes (1994) and Lal and Rao (1997) explore price variation across firms for a bundle of items in order to explain loss leader pricing and price-positioning strategies, respectively. Again, the sustainability of the competitive pricing equilibrium (where, for example, firms either randomize price discounts across goods or decide to offer a fixed advertised discount on the market basket) rests on consumer heterogeneity in price sensitivity induced by differences in travel costs or willingness to search.

Empirical Tests of Pricing Theories. Several authors have attempted to test the predictions of some of these theories. Recent among them is Lach (2002) who provides an empirical test of Varian's theory using data from both durables (refrigerators) and consumables (chicken, flour, paper products). He finds that: (1) cross-sectional variation in store prices remained approximately constant over time for these goods, and (2) stores however constantly rearrange their relative positions within the cross-sectional price distribution in a random manner – consistent with the predictions of Varian (1980). Sorenson (2000) utilizes data on cross-sectional price variation for prescription drugs and finds drugs with a higher usage velocity show smaller

cross-sectional differences between the minimum and maximum posted prices in a particular time period. This empirical finding is consistent with the classic conjecture of Stigler (1961) on consumer motivations to search for lower prices.

In a departure from these previous empirical tests, the phenomenon of interest in our research is the extent to which the presence of heterogeneity in storage costs: (a) influences demand patterns across categories and (b), how it affects the optimal pricing behavior of competing firms.³ Consumer storage costs are posited to interact with category characteristics – the absence of stockpiling behavior for some households should be more pronounced for product categories that require greater volumes of space. That is, we not only test the price-setting implications of consumer storage costs, but also propose and implement an empirical method to first isolate the storage costs themselves.

Overview and Contribution. This paper provides three new contributions. First, we offer an empirical approach to assessing consumer storage costs. Our measure of storage cost is validated by showing that while it predicts average purchase quantities for stockpiled products (e.g., bathroom tissue, detergents, paper towels) it has no impact on purchase behavior for pills in capsule form – a product that requires essentially no space. Second, we develop empirical tests of the predictions emanating from the theoretical model of Salop and Stiglitz (1982). We show that all three key predictions for average prices and price variation are supported by the data. Finally, we present an analysis of the quantitative effects of storage costs and show that consumer movement to higher storage cost regions has non-trivial implications for shopping behavior and retail price variation established by firms. The remainder of the paper is organized as follows. In the next section we provide a brief overview of related research and present a summary of the key results from Salop and Stiglitz (1982). We then describe our data and empirical approach. Subsequently, we proceed to test both the assumptions and predictions of our modeling framework. The paper concludes with a discussion and implications for further research.

1 Background and Model

Interest in consumer response to price and the influence of market characteristics on the price-setting behavior of firms has a long and varied history. This is understandable given the

primacy of price in decision making and the importance of the price-mechanism in market performance. While many factors such as informational differences or heterogeneity in willingness to search have received a good deal of attention, relatively little emphasis has been placed on the role of storage costs in both (consumer) purchase behavior and (firm) price-setting behavior.

1.1 Storage Costs

Blattberg, Eppen, and Lieberman (1981) is the first marketing article to introduce storage costs as an explanation for retail price promotions. If retailers have storage costs that exceed those of the customers it may be optimal for them to offer deals that induce these consumers to stockpile. In this way, individuals bare small incremental increases in storage costs while the retailer benefits from the aggregated reduction in the cost of holding inventory. Data from four product categories (aluminum foil, facial tissue, liquid detergent, and waxed paper) are presented to show that consumers indeed buy higher product volumes when prices are low and take longer to replenish subsequent supplies subsequent to purchases on promotion.⁴

Bell, Chiang, and Padmanabhan (1999) study multiple product categories and present evidence that bathroom tissue, ground coffee, liquid detergent, and paper towels are stockpiled by consumers. That is, when deal prices are offered, average volumes increase, but the average time until a subsequent purchase also increases. The net result is that average consumption rates are not different between deal and non-deal purchases. Collectively, the two studies point to the viability of retail promotions as a mechanism for transferring product storage costs from retailers to consumers, and to differences in product categories with respect to consumer willingness to stockpile.

1.2 The Salop and Stiglitz Model

Consumers and Firms. In Salop and Stiglitz (1982) there are T consumers who have two-period consumption and planning cycles. The purchase decisions are for goods that are not advertised explicitly, such that consumers cannot know the actual price charged by a particular store, but do know the distribution of prices, $f(p)$. All consumers are homogenous with respect to their reservation price, u , for each unit of the product under consideration. Furthermore, price uncertainty in the market implies that consumers select stores at random.

All consumers have unit demand in each period for a total of two units over the consumption cycle. Consumers who are able to stockpile will do so if the “pivot price” (i.e., the price that makes them indifferent between purchasing for storage and future consumption instead of current consumption only) is sufficiently low. A consumer who encounters such a price, \hat{p} , will not reject it in favor of additional search for an even lower price. The pivot price is therefore \hat{p} , where $\hat{p} + h \leq \bar{p}$. The variables \bar{p} and h represent the expected price from the stationary price distribution, $f(p)$, and the per unit storage cost, respectively. Consumer risk neutrality is assumed such that it is the expected price that is relevant in the purchase decision. Consumers maximize their full inter-temporal utility and in equilibrium, their expectations regarding prices will be fulfilled.⁵ That is, their behavior is rational and the stockpiling strategy is undertaken only in the event that the surplus from doing so outweighs that of a simple per-period current consumption strategy. The market consists of n firms who are ex ante identical (so that price differences observed in equilibrium will be driven by the internal workings of the market and not by heterogeneity in firm characteristics). Retailers selling in the market have constant and equal marginal costs which are normalized to zero.

Equilibrium and Model Predictions. The equilibrium price distribution, $f^*(p)$, is derived as follows. Consumers are fully rational in their selection of purchase strategies, choosing options that maximize total inter-temporal utility, and furthermore, their expectations concerning the price distribution are fulfilled. Market entry only occurs when consumers get non-negative surplus.⁶ Likewise, retailers are rational and choose pricing strategies to maximize prices, given the behavior of their retail competitors. The further conditions supporting the equilibrium price distribution are: (1) retailers price to maximize profits and (2) all firms earn the same profits.

The equilibrium price distribution contains at most two prices, denoted by p_h and p_l with $p_h > p_l$. Random selection of stores by consumers implies that all stores will receive the same number of consumers (T/n consumers). The mixed strategy profile contains exactly two prices. The higher of the two prices equals the reservation price, u , and will be charged with probability λ . It is also straightforward to see that the lower of the two prices should be exactly equal to the pivot price, \hat{p} .⁷ Explicit solutions for the two unknown parameters of the price distribution, \hat{p} and λ , are obtained by characterizing the demand functions for the two types of stores.

Each type of store receives the same number of customers and also encounters a mix of “old” (re-entering) and “new” (first time) customers. Defining sales at the high and low priced stores as S_h and S_l yields $S_h = T/n [1 + \lambda]$ and $S_l = T/n [2 + \lambda]$. The equal profits condition in the mixed strategy equilibrium requires that $p_h \cdot S_h = p_l \cdot S_l$. As the pivot price for the stockpiling strategy can be expressed as a function of the expected price: $\hat{p} = p_l = \lambda u + (1 - \lambda) p_l - h$ it is possible to solve for the two unknowns, p_l and λ . Given no cost of re-entry, the equilibrium price distribution, $f(p)^*$, can be characterized as follows (see Salop and Stiglitz 1982, pp. 1124-1127):

Proposition 1

A necessary and sufficient condition for the existence of a symmetric mixed strategy equilibrium is $h < u / 3$ and the equilibrium prices are $p_h = u$ and $p_l = (u + h) / 2$. Stores charge the high price with probability $\lambda = 2h / (u - h)$.

Proof: The necessary and sufficient conditions for a mixed strategy equilibrium with promotional pricing to exist are $0 < \lambda < 1$ and $p_l + h \leq u$. The upper constraint on λ gives $h < u / 3$ while $h < u$ is required for $\lambda > 0$. Furthermore, $p_l + h < u$ is equivalent to $h < u / 3$ as well.

Proposition 1 encompasses the three results that we test empirically. Namely, an increase in consumer storage costs leads to:

- (1) An increase in average prices as $\bar{p} = (3h + u)/2$ is increasing in h (Prediction 1),
- (2) a reduction in the probability of promotion as $(1-\lambda) = (u-3h)/(u-h)$ is decreasing in h (Prediction 2), and
- (3) a decrease in promotion depth as the size of the discount (high price minus low price) relative to the regular (high) price $= (u-h)/2u$ is decreasing in h (Prediction 3).⁸

2 Empirical Analysis

We begin with a description and integration of separate datasets on (1) consumer purchasing behavior, (2) store pricing behavior, and (3) consumer storage costs. We subsequently elaborate on the estimation procedure for the household-level storage cost proxy and the estimation results for this variable. That is, we seek to establish the validity of the storage cost proxy as a determinant of consumer behavior. The section concludes with tests of the store-level predictions that follow from the theory.

2.1 Data

Our data are derived from the Stanford Market Basket Database consisting of scanner data for 1042 panelists in the Chicago Metropolitan area, collected between June 1991 and June 1993 in two submarkets. We have information on a number of demographic characteristics for each household in the panel (see Table 1). The two submarkets are in downtown Chicago (five stores) and at the urban fringe in the South West of Chicago (four stores). 548 panelists shop in the downtown market and 494 shop in the urban fringe. For reasons of confidentiality the five downtown stores are coded as 1419, 1420, 1422, 1423, and 1424 (stores 1423 and 1424 are owned by the same chain). In the urban fringe, the four stores are 1521, 1522, 1542, and 1558. In addition to knowing the exact location of each of the nine stores, we also know the zip code location of the households in the panel. From this information we are able to compute an estimate of the travel distance for each store-household pair.⁹

The empirical analysis requires both identification of storage costs for each household in the sample and sufficient variation in storage costs within and across the two trade areas. A map of both markets which illustrates the spatial distribution of storage costs and of consumers by zip codes is provided in Appendix 1. In this map, storage costs are represented by the following measure: Housing cost per square footage of living area. We will introduce the method of imputation for this measure in the next section and in subsequent sections discuss a number of other candidates for the storage cost proxy.¹⁰ Each shaded zip code represents a region that contains panelists who shop in the submarket and the location and concentration of panelists themselves is shown by the dots on the map. There is no cross-shopping in the sense that panelists either shop in the downtown stores, or in the urban fringe stores. As evidenced in the map and discussed subsequently there are substantial differences between the two submarkets in terms of the level of household storage costs: The average housing unit value per square foot is approximately \$400 in the downtown market and \$40 in the urban fringe. While this is somewhat evident from the map in Appendix 1, the distributions of consumer storage costs by submarket are illustrated in Appendix 2.

In the empirical analysis we utilize household-level location information to control for the distance households must travel to reach a store in conjunction with other demographic control variables. Summary statistics for the travel, demographic, and behavioral characteristics of the consumer panelists are summarized in Table 1. Of the original 1,042 panelists fewer than five

percent had missing values for demographics or location information. In the empirical results presented in Tables 3 – 8 we will utilize two variations of the storage cost proxy under consideration. The first will assume that all households within the same zip code share the same storage cost value (e.g., all households in the same zip code have the same “housing unit value per square foot”). The second will use hedonic regression estimates in combination with household-level values of demographic variables to produce a household-specific measure of the storage cost proxy. Under these two variations 996 households remain in the dataset in which each household within a particular zip code shares the same storage cost value and 991 remain when this measure is adjusted according to individual household characteristics.

[----- **Table 1 About Here** -----]

Purchase Data and Pricing Data. Our analysis of purchasing behavior is focused on five product categories. In four of these product categories (bathroom tissue, ground coffee, liquid detergent, and paper towels) the tendency of some consumers to stockpile has been previously documented. All four categories were analyzed by Bell, Chiang, and Padmanabhan (1999, Exhibit 3, p. 517) and Blattberg, Eppen, and Lieberman (1981, Tables 1-2, pp. 124-125) also report stockpiling in liquid detergents. A fifth category that uses essentially no space – pills in capsule form – is used as a control category to help establish the empirical validity of our storage cost proxies. Descriptive information in terms of number of SKUs, average prices and average volumes for the five product categories — summarized by store — is provided in Table 2. As noted in the Table, all categories are summarized according to the IRI definition of a “standard unit” for the category in question. We identify our effects using both variation across categories in requirements for space, and variation across markets in household storage cost.

[----- **Table 2 About Here** -----]

2.2 Estimation of Storage Cost Proxy

Central to the contribution of this research is the creation of a reliable estimate of household storage costs which can then be introduced as an explanatory variable in our model of consumer and firm behavior. While our measure and the method of construction is somewhat new to the marketing literature, there is precedent for this approach in recent work in urban economics (e.g., Glaeser, Gyourko, and Hilber (2001) and Glaser and Gyourko (2001 and 2002)).¹¹

Our preferred proxy measure for household storage cost is the housing cost per square footage of living area. We need to impute this measure as the value and living area of the panelists' homes are not directly observed. Specifically, we use location and demographic information of the panelists and hedonic housing unit value and living area equations to impute the panelists' housing cost per square foot of living area. We report regressions for two alternative measures of housing cost per square foot. Both measures are imputed from the following three data sources: the Stanford Market Basket Database, the US Census 1990, and the American Housing Survey (AHS) 1989.¹²

The first measure is the median value of a housing unit in a zip code (derived from 1990 Census data) per square foot (estimated based on AHS 1989 information). The second measure goes one step further and uses the demographic information of the panelists from the Stanford Market Basket Database (in addition to their location information) to impute the housing value per square foot of living area for each panelist.¹³ The methods to impute the two housing value per square foot measures are described in detail in Appendix 3.¹⁴ Summary statistics of the two alternative storage cost measures (zip code and household level) are provided in Table 1.

The two measures represent the notion of “opportunity cost of housing space”. That is to say, for any given housing unit, consumers must choose how much space to allocate to storage and how much space to allocate to other purposes. In this context, any decision to stockpile household consumables is costly, as it gives up the opportunity to use the space for alternative purposes. Of course, the opportunity costs are much larger in an expensive downtown unit with a high per square foot value of space compared to a housing unit in the urban fringe or in rural areas, where housing space is relatively inexpensive.

While the notion of “opportunity cost of housing space” seems quite intuitive, one could also make the case that for consumer packaged goods, housing unit *size* alone may be a more suitable proxy for storage costs. This argument relies on the premise that households only take into

account the size of their housing unit but not the cost per square foot. That is, households in small units cannot stockpile because of space limitations, while households in large units will always find some space to stockpile consumables. In order to test whether our findings are robust to alternative measures of storage cost itself (rather than simply alternative levels of aggregation – zip code or household level) we also report results that utilize the *housing unit size* as a proxy for storage cost. Again, we compute a zip code level specific measure and a measure that also takes into account the demographic characteristics of the panelist. All our results hold when utilizing these two alternative measures.

Moreover, our main results hold even when we use ‘less precise’ proxy measures, albeit at lower significance levels. The list of tested alternative measures includes a dummy variable for “storage constrained households”, which is one if the housing unit is below a certain threshold size; the percentage of housing units in a zip code that are in apartment buildings; the percentage of single family detached housing units in a zip code; the average number of rooms per housing unit in a zip code; the population density of the zip code; whether the panelist has a dog or not; the number of dogs of each panelist; or a logarithmic specification of the house value per square foot measure. This set of empirical results is consistent with and complementary to those reported by Hendel and Nevo (2004) in their analysis of the Stanford Market Basket Database. They find that in “Market 1” (the downtown market) households are less likely to avail themselves of sales. They attribute this to higher relative storage costs and note that these households live in smaller homes (compared to those in the other submarket). They also find that dog ownership (but not cat ownership) is positively correlated with the frequency with which a household buys a storable product on deal, and again attribute this to relatively low storage costs for these households: Households with dogs are conjectured to have larger homes (see Hendel and Nevo 2004, pp. 21-22).

2.3 Empirical Analysis of Model Assumptions (Consumer Behavior)

(a) *Purchase Frequency.* First, we test the assumption that consumers visit stores more often if they have higher storage costs. Our basic estimating equation for the purchase frequency of consumer i is as follows:

$$(1) \quad \# \text{ trips}_i = \beta_0 + \beta_1 \text{ storage cost}_i + \beta_2 \text{ weighted distance to store}_i + \beta_3 \text{ demographics}_i + \beta_4 \text{ submarket}_i + \varepsilon.$$

The coefficient of interest is β_1 . We also include the distance to stores (weighted by the number of trips to each store) and a number of demographic characteristics of the consumers (see Table 1 for a description of these variables). Finally, the equation controls for unobservable characteristics that are unique to the submarket (downtown or urban fringe).

Alternatively, we include store dummy variables. The dummy variable for a store is zero unless it is consumer i 's closest store. The equation is as follows:

$$(2) \quad \# \text{trips}_i = \beta_0 + \beta_1 \text{storage cost}_i + \beta_2 \text{weighted distance to store}_i + \beta_3 \text{demographics}_i + \beta_4 \text{store dummies (closest store)}_i + \varepsilon.$$

Table 3 summarizes the regression results for both estimating equations and for both storage cost proxy measures: (1) the zip-code level storage cost measure, in which each household in the zip code shares the same storage cost value, and (2) the individual-level measure, in which household-specific demographic information is used to adjust the zip-code level estimate. As noted above both measures are obtained from the median housing values (from the 1990 Census) per square foot (from the 1989 AHS) and the results in Table 3 are robust to a number of alternative ways of imputation (see section 2.2). Table 3 (and all subsequent tables) reports robust standard errors, using a Huber and White sandwich estimator of variance.

Our coefficient of interest, β_1 , has a positive sign and is highly statistically significant in all four cases – that is for both storage cost measures and whether the estimating equation is (1) or (2). The other coefficients have plausible signs. For example, higher income households are visiting stores less often, likely because of higher opportunity costs. Household size shows the same effect, again likely due to the opportunity costs of time and the need to buy larger baskets of goods per store visit. Total spending in all stores is included as a control variable to proxy for overall consumption: Households that consume more visit stores more often. Another variable of particular interest is the weighted distance to stores. Although the estimated coefficient is negative, it is not statistically significant in any of the reported specifications. This finding implies that the average distance to stores has no appreciable effect on the purchasing frequency. However, this result does not imply that distance does not matter for consumer behavior.¹⁵

[----- Table 3 About Here -----]

The results in Table 3 offer strong support for the assumption that higher storage costs will lead consumers to take more shopping trips. Yet, as noted in section 2.2., one could perhaps question our preferred storage cost proxy: Housing unit value per square foot. In order to test whether our findings are robust to an alternative definition of storage cost, we re-estimated equations (1) and (2) using housing unit size as the independent variable representing storage costs.

In this instance, our coefficient of interest, β_1 , should have a negative sign: Households with larger housing units (lower storage costs or lower stockpiling constraints), should shop less often. Table 4 shows this finding is strongly supported, and as before the coefficients for the control variables have plausible signs and magnitudes similar to those in Table 3. In the interest of demonstrating the robust nature of our results we will continue to report findings for the preferred measure (housing unit cost per square foot) and the alternative proxy (housing unit size) when we examine firm behavior and the quantitative effects implied by our estimates.

[----- Table 4 About Here -----]

(b) *Purchase Quantity per Trip*. Our second testable assumption is that a consumer i will purchase smaller quantities per trip if they face higher storage costs. Here we report separate regression results for all five product categories. Our estimating equations parallel equations (1) and (2) and are as follows:

$$(3) \quad \begin{aligned} \text{purchase quantity}_i = & \beta_0 + \beta_1 \text{ storage cost}_i + \beta_2 \text{ unit price of product category} + \\ & \beta_3 \text{ weighted distance to store}_i + \beta_4 \text{ demographics}_i + \\ & \beta_5 \text{ submarket}_i + \varepsilon \end{aligned}$$

$$(4) \quad \begin{aligned} \text{purchase quantity}_i = & \beta_0 + \beta_1 \text{ storage cost}_i + \beta_2 \text{ unit price of product category} + \\ & \beta_3 \text{ weighted distance to store}_i + \beta_4 \text{ demographics}_i + \\ & \beta_5 \text{ store dummies}_i + \varepsilon. \end{aligned}$$

In estimating these regressions, recall that we have five product categories as described in Table 2: Bathroom tissue, paper towels, liquid detergents, ground coffee, and pills in capsules. The results in Table 3 show that storage costs have the expected macro-level effect on consumer shopping behavior: Households with higher storage costs shop more often, all else equal. The

analysis of purchase behavior in individual product categories that differ with respect to storage requirements offers a further opportunity to validate our measure of household storage cost. Storage requirements for bathroom tissue, paper towels, and liquid detergents are relatively high, whereas pills in capsules and ground coffee consume very little space. Even households with very high housing values per square foot (or very small housing units) should be able to store an additional fifty pills in capsules. Conversely, it may be relatively difficult for them to hold an additional 32oz of laundry detergent or six rolls of bathroom tissue.

In reporting the results for equations (3) and (4) we separate the findings into “high storage use items” of paper towels, bathroom tissue, and liquid detergents (Table 5) and “low storage use items” of ground coffee and pills in capsules (Table 6). Again, our coefficient of interest is β_1 and the prediction is that – at least for high storage use items – the coefficient will be negative: Consumers with higher storage costs purchase smaller quantities on average, per shopping trip. As before, our equations contain a number of control variables (including the average unit price) and again we report our results for the zip-code level and the individual-level storage cost proxy.

Table 5 reveals that storage costs have a strong negative effect on purchase quantities in all high-storage cost categories (paper towels, bathroom tissue, and liquid detergent). All effects are statistically different from zero with 99 percent confidence. Table 6 reports results for products that use much less storage space; coffee and pills in capsules. Interestingly, here the effects of storage costs on purchase quantity are statistically insignificant (although still negative).

[----- Tables 5 and 6 About Here -----]

Thus, the findings pertaining to the key model assumption – storage costs influence consumer behavior – are unequivocal. Higher levels of storage cost not only cause consumers to shop more often, but also cause them to purchase smaller volumes of product per shopping trip (i.e., the effects can be observed at the product category level). The face validity of our storage cost proxy is further enhanced by the negative and significant effect on purchase quantities for high storage use items (paper towels, bathroom tissue, and liquid detergent) and insignificant impact on low storage use items (ground coffee and pills in capsules). Finally, it is worth noting that results are similar when we use the housing unit size as the storage cost proxy rather than the housing unit value per square foot.

Having established the validity of the consumer behavior assumption and our ability to measure storage costs, we now turn our attention to the predictions of Salop and Stiglitz' Theory of Sales.

2.4 Empirical Analysis of Model Predictions (Firm Behavior)

Salop and Stiglitz (1982, p. 1122) note that the price variation captured by their model "... may be across stores, across brands of the same product, or at a single store over time." The substance of the model and this interpretation has important implications for empirical testing as it suggests both cross-sectional and temporal approaches are equally legitimate. In our data we have relatively few cross-sectional units (nine stores in two submarkets) but potentially many more observations over time for each SKU (as many as 104 weeks per store). The measures of price variation (average prices, promotion probabilities, promotional depth) are therefore calculated within stores, over time.

(a) *Prediction 1: Average Product Offer Price.* The specification for the average product (SKU) offer price for product p in store s is as follows:

$$(5) \quad \begin{aligned} \text{offer price}_{p,s} = & \beta_0 + \beta_1 \text{ consumer storage cost faced by store}_s + \\ & \beta_2 \text{ demographics of store's customers}_s + \\ & \beta_3 \text{ weeks on shelf}_p + \beta_4 \text{ SKU fixed effects}_p + \varepsilon. \end{aligned}$$

The theoretical model presented in section 1 predicts that the coefficient β_1 will be positive (holding everything else constant) as the comparative static result shows that \bar{p} is increasing in h . The set of control variables mirror those used in the consumer behavior regressions. In particular, we characterize the store's customer base according to their average income, age, and family size. While the store owners can observe the age (distribution) of the consumers quite easily, scanner data allows them to make reasonable assumptions about their income as well. Higher income consumers may be less sensitive to prices compared to low income customers. The same may be true for elderly consumers, as searching for better deals may be more cumbersome for them. Similar to the regressions on the consumer side, we calculate the percentage of unemployed, Blacks, and Hispanics in the customer base faced by each store.

To the extent that different age and income (and other) groups have different elasticities, it will be important to control for the impact this has on the stores' pricing decisions. We also control for differences in product "supply" — summarized by the number of weeks the product is available on the shelf. Product availability is likely to be correlated to popularity and therefore also the probability that the store will promote the product to drive store traffic or match retail competitors.

As shown in Table 7, Prediction 1 of the theory has strong support: Columns (1a) and (1b) show that average prices are higher in stores that face a distribution of customers who have higher storage costs (as measured by the house value per square foot and as measured by the housing unit size). This substantive finding is robust to a number of different specifications as outlined in the note below Table 7. In addition, the control variables are significant and have expected and plausible signs. For example, prices are higher if stores face households that have higher income, are older, or are larger. However, prices are lower if stores face households with unemployed members. All these results are consistent with search cost theory. We also include SKU-level fixed effects to control for all variation in prices that could be due to idiosyncratic differences in the product characteristics themselves.

[----- Table 7 About Here -----]

(b) *Prediction 2: Probability of Promotion.* The empirical estimation strategy is consistent with that employed for the average price prediction, apart from the fact that the dependent variable is log-transformed to ensure that it is symmetric about zero. The basic estimating equation is as follows:

$$(6) \quad \log\left(\frac{\Pr(\text{promotion})_{p,s}}{1 - \Pr(\text{promotion})_{p,s}}\right) = \beta_0 + \beta_1 \text{ consumer storage cost faced by store}_s + \beta_2 \text{ demographics of store's customers}_s + \beta_3 \text{ weeks on shelf}_p + \beta_4 \text{ SKU fixed effects}_p + \varepsilon.$$

The theoretical model presented in section 1 predicts that the coefficient β_1 will be negative (holding all else constant) as the comparative static result shows that $(1-\lambda)$ is decreasing in h . $\Pr(\text{promotion})_{p,s}$ is defined empirically as the proportion of times the product is offered at a price

discount (taken relative to the number of weeks the product is available on the shelf). We report results for two alternative storage cost proxy measures: Housing unit value per square foot (Column 2a) and housing unit size (Column 2b).

Again, we find strong support for the Salop and Stiglitz model: Columns (2a) and (2b) indicate a negative effect of consumer storage costs on the likelihood of firms charging a high price. It is worth noting that results with regard to the storage cost proxy measures are little changed if we employ a linear instead of a log-transform specification. This empirical finding supports the intuition that in markets where the costs of storage are relatively high, it does not make sense for stores to attempt to induce stockpiling.

(c) *Prediction 3: Promotional Depth.* The basic estimating equation for the promotion depth of product p in store s is as follows:

$$(7) \quad \log\left(\frac{\text{promotion depth}_{p,s}}{1 - \text{promotion depth}_{p,s}}\right) = \beta_0 + \beta_1 \text{ consumer storage cost faced by store}_s + \beta_2 \text{ demographics of store's customers}_s + \beta_3 \text{ weeks on shelf}_p + \beta_4 \text{ SKU fixed effects}_p + \varepsilon.$$

The prediction from theory is that all else equal, the relative size of the discount decreases in storage costs because the comparative static reveals that $(p_h - p_l)/p_h$ is decreasing in h . Empirically, promotional depth is the relative size of the price cut, expressed as a percentage of the regular price taken over the number of weeks the product is available on the shelf. A log-transform model is estimated to ensure that the dependent variable is symmetric about zero. Again, we report results for two alternative storage cost proxy measures: Housing unit value per square foot (Column 3a) and housing unit size (Column 3b). Each specification yields a storage cost parameter estimate that is negative and therefore consistent with theory: Stores that face customers with higher storage costs, not only promote less often (Columns 2a and 2b), but when they do, they offer shallower discounts (Columns 3a and 3b).

2.5 Summary and Quantitative Effects

Our empirical results confirm the key assumption and three main predictions of the Salop and Stiglitz (1982) model in the sense that the estimated relationships between storage costs on

the one hand and consumer and store behavior on the other hand are statistically significant. We further confirm our results by demonstrating that the consumer storage costs matter for relatively storage intensive products (bathroom tissue, liquid detergents, paper towels) but that the effects are not significant in the case of low storage use items (coffee and pills). Statistical significance, however does not tell us anything about the quantitative importance of these effects. In particular, we are interested to see how large the difference in consumer and store behavior is in two markets that differ substantially in storage costs (see Appendix 1).

As the summary statistics for our preferred storage cost proxy reveal, a typical downtown household pays around 400 dollars per square foot of living space, whereas in the urban fringe, where households consume larger but cheaper housing units, the house value per square foot is around only 40 dollars. The alternative storage cost proxy – housing unit size – reveals an average value of around 1600 square feet in the low storage cost urban fringe and 1000 square feet in downtown.

Quantitative effects for key measures of consumer behavior – shopping frequency and product quantities purchased per trip – are reported in the first six columns of Table 8. These effects are computed under the scenario where a typical household moves from the urban fringe (\$40 per square foot of living space; 1600 square feet of living space) to the downtown (\$400 per square foot of living space; 1000 square feet of living space) neighborhood.

[----- **Table 8 About Here** -----]

Shopping Frequency. Column (1) and rows 1-4 of Table 8 focus on shopping frequency and reveal that an urban fringe household moving to the downtown neighborhood would on average increase the percentage of shopping trips taken by approximately 10 to 23%, depending on the measure of house value per square foot employed (see Table 3). That is, whether the value is determined at the zip code level, imputed to the household level, or based on regressions that use store dummies. A qualitatively identical finding is shown in rows 5-8. In this case the percentage of shopping trips taken increase on average by approximately 10 to 21%, depending on the measure of housing unit size employed in the analysis (see Table 4). The level of increase in shopping trip frequency of approximately 15%, obtained by averaging across regression results for both storage cost proxies – housing value per square foot and housing unit size – and across

slightly different specifications of these two main proxies, is economically quite important. Differences in the storage cost environment play a large role in shopping trip frequency, all other things held constant.

Purchase Quantities. Columns (2) through (6) report the effect on average purchase quantities for all five product categories. As shown in rows 9-12, the shopping constraint imposed by the increase in storage costs leads the household to purchase significantly smaller volumes on average, for each shopping trip. The percentage changes should be interpreted with respect to the definitions of a standard unit for each category as introduced in Table 2 and reproduced at the bottom of Table 8. Focusing on paper towels as an illustrative example, we see that a household moving from a low to a high storage cost environment would reduce the average purchase quantity of paper towels by about 18% (averaged across the four different measures of housing value per square foot). The average number of paper towels bought per purchase occasion would fall from 1.7 to approximately 1.4. Similarly, the average number of rolls of bathroom tissue per purchase occasion would decline to 4.7 from 5.3. Detergent volumes per purchase occasion would fall from 5.7 to 5.2 standard units (or from 92 to 83 ounces). Such changes in purchasing behavior should, in theory, also have important corresponding implications for the way in which firms set product prices and decide upon the frequency and depth of promotion.

Prices, Frequency, and Depth of Promotion. Columns (7) to (9) focus on the store level implications when household storage costs increase in this same manner. The magnitude of the effect on average prices is on average 3%. That is, a downtown store sets about 3% higher prices on average than a store in the urban fringe, as a consequence of higher consumer storage costs, all else equal. An average household purchasing an average volume of bathroom tissue would see the price increase from approximately \$1.64 to \$1.69. Corresponding negative effects are also observed for the probability of promotion and the depth of promotion. In absolute terms the average likelihood of promotion declines by 3% and the depth of promotion falls by just over 10%.

Summary. Thus, we find storage costs not only produce robust and statistically significant effects on firm and consumer behavior, but also have important quantifiable economic consequences for both.

3 Conclusion

While Sorensen (2000) and Lach (2002) provide empirical tests of the more prevalent theories of consumer price search (Stigler 1961) and price variation that results from informational differences (Varian 1980) respectively, our study is the first to examine in detail the storage cost explanation posited by Salop and Stiglitz (1982) in their Theory of Sales. A critical underlying assumption of their theoretical model is that the storage of goods is costly to consumers and this will not only affect their purchase behavior but also drive variation in retail prices. This assumption implies that an increase in storage costs causes consumers to purchase goods more often and to buy smaller quantities at a time. We find this assumption to be supported in data for household consumables that can potentially be stockpiled. Of equal importance, we find that for low storage use items like pills in capsule form, household level storage costs have no appreciable effect on consumer purchase behavior.

Of particular interest are three empirically testable predictions that can be derived from Salop and Stiglitz' equilibrium framework: An increase in consumer storage costs will lead to: (1) an increase in the average price, (2) a reduction in the probability of promotion, and (3) a decrease in promotion depth. The assumptions and predictions are tested using and combining data from two remotely adjacent fields; marketing (market basket data) and urban and real estate economics (housing cost data). Our empirical analysis provides evidence for both the validity of the assumptions and the predictions of the Theory of Sales. Furthermore, we show that the investigated effects are not only statistically significant but also quantitatively important economically. For example, our results imply that a downtown consumer who lives in a small but expensive housing unit makes about 15 percent more shopping trips compared to an identical consumer who lives in a larger and less expensive housing unit in the urban fringe. Storage costs are not only affecting consumer behavior but also a store's pricing behavior. Our results suggest that, holding everything else constant, a downtown store is about 60 percent less likely (3% in

absolute terms) to promote any particular bathroom tissue SKU compared to a store in the urban fringe. In addition, promotional depth is significantly lower (10% in absolute terms). We demonstrate that storage costs also affect the store's product offer price, although our estimates suggest that this effect is fairly limited in scope. That is, all else equal, downtown stores set offer prices that are on average about 3% higher than those set by urban fringe stores, as a consequence of higher consumer storage costs.

Future research will address whether storage costs are related to price dispersion in the way suggested by Salop and Stiglitz (1982). Their equilibrium framework implies that price dispersion is first increasing and then decreasing in storage costs. We will derive an econometric specification that will allow us to test this prediction. Initial results for both submarkets (downtown and urban fringe) indicate that within-store price dispersion is indeed first increasing and then decreasing in storage costs, confirming Salop and Stiglitz' main prediction with regard to price dispersion. A different research path will attempt to advance the urban and real estate economics literature, by investigating the relationship between the customer's distance to a store and the consumer's shopping and the store's monopolistic price setting behavior.

Notes

- ¹ While storage costs are primarily a *consumer* specific phenomenon related to the size and value of the housing unit, there are also potential interactions with product categories. We explore this possibility in our empirical analysis and compare and contrast categories that require essentially no space (e.g., pills in capsules) with other categories such as detergents and paper towels.
- ² If one takes an expansive view of the construct “price”, this general observation extends to studies that examine the influence of reference prices and deviations from price expectations. Krishnamurthi, Mazumdar, and Raj (1992), for example, show that consumers are loss averse in their purchase quantity decisions; however their data do not allow an investigation of the role of storage costs in affecting purchase quantities.
- ³ Bell, Iyer, and Padmanabhan (2002) derive and test the implications of a model in which consumers who purchase additional inventory at low prices may also be induced to consume at a higher rate. They do not however empirically assess the effect of storage costs on consumer or firm behavior.
- ⁴ Actual storage costs (or suitable proxies) are not available to the authors. They partially circumvent this problem by looking at the frequency of dealing for products of the same type but of different sizes -- they make the reasonable assumption that stockpiling larger sizes should be more costly to the consumer and therefore require greater inducements. This intuition is consistent with findings in recent work by Hendel and Nevo (2003).
- ⁵ As noted by Salop and Stiglitz (1982) one could further generalize this notion of an “expected price” to include consumer beliefs about prices that reflect commonly observed biases and heuristics. Examples might include modifications suggested by Prospect Theory (Kahneman and Tversky 1979) or behavioral decision anomalies such as the Frequency Heuristic (Alba et al 1994).
- ⁶ Salop and Stiglitz abstract away from this issue by assuming that the cost of re-entry is zero. This facilitates ease of exposition and does not affect the qualitative results.
- ⁷ The solution cannot have $\hat{p} < p_l < p_h$ because in this instance both the low and high priced stores sell only one unit to both sets of customers, yet the store charging p_h necessarily makes higher profits.
- ⁸ Salop and Stiglitz also show that price dispersion is non-linear – first increasing and then decreasing – in relative storage costs u / h .
- ⁹ Bell, Ho, and Tang (1998) augmented the Stanford Market Basket Database with household-store distance measures. These measures were derived assuming that households are uniformly distributed across the zip code (details are contained in their paper, pp. 359-360).
- ¹⁰ In addition to providing alternative measures of household storage cost (e.g., housing unit value per square foot, housing unit size, etc.) we also provide estimates at different levels of aggregation, for the *same* measure. That is, the housing unit value per square foot (for example) is constructed under the assumptions that: (1) all households living in the same zip code have the same storage cost value, and (2) household-specific storage cost values can be calculated from hedonic regression estimates in combination with household demographic information found in the Stanford Market Basket Database. See Tables 3-8.

- ¹¹ Glaeser, Gyourko, and Hilber (2001) and Glaeser and Gyourko (2001 and 2002) use the US Census and the American Housing Survey to impute the house value per square footage, similar to the way described in Appendix 3. They then use construction cost information per square footage to compute the hypothetical value of residential land. This methodology is used to analyze housing affordability, urban decline, and the impact of zoning on housing affordability.
- ¹² The National AHS consists of approximately 55,000 households and numerous household, housing unit, and neighborhood specific variables. The dataset also includes two measures of particular interest to our study: The living area of a housing unit and the house value. These two measures and the numerous other housing unit and household specific characteristics allow us to impute the house value per square foot, our storage cost proxy measure of interest. The most disaggregated location information available is the MSA level.
- ¹³ Three alternative proxy measures of storage cost were imputed. The first measure uses the average house value instead of the median house value for computations. The median — as opposed to the average — house value may create a better proxy as outlier values are less relevant. However, results are very similar if we use the average instead of the median house value per square footage. The second and third measures use AHS data from 1991 instead of 1989 in order to compute the zip code specific median and average house value per square footage. The AHS information from 1989 may be more closely related to the 1990 Census information since the Census data was collected during 1989. However, again, results are virtually unchanged if AHS 1991 data are used for imputations rather than 1989 data.
- ¹⁴ The imputation method is similar to the method first developed in Glaeser, Gyourko, and Hilber (2001). For applications of this method see Glaeser and Gyourko (2001 and 2002).
- ¹⁵ We investigated this further and found that for low price stores the correlation between distance to a store and probability of store visits is quite weak, while distance is highly negatively correlated to this same probability for a high price store. The explanation from urban economics is straightforward (e.g., DiPasquale and Wheaton 1996, p. 124 onwards): Consumers typically only buy small quantities in high price stores, thus, the transport costs per unit are relatively high and the retailer's market area is therefore quite small. That is, consumers are highly unlikely to visit a high price store if the distance exceeds a certain threshold. On the other hand, low price stores have much larger market areas and the probability of a store visit is much less sensitive to distance, as distance is in relative terms less costly. We intend to explore this issue in greater detail in a separate study.

References

- Alba, Joseph W., Susan M. Broniarczyk, Terence A. Shimp, and Joel E. Urbany. (1994). "The Influence of Prior Beliefs, Frequency Cues, and Magnitude Cues on Consumers' Perceptions of Comparative Price Data", *Journal of Consumer Research* 21, 219-235.
- Bell, David R., Jeongwen Chiang, and V. Padmanabhan. (1999). "The Decomposition of Promotional Response: An Empirical Generalization", *Marketing Science* 18, 504-526.
- Bell, David R., Ganesh Iyer, and V. Padmanabhan. (2002). "Price Competition Under Stockpiling and Flexible Consumption", *Journal of Marketing Research* 39, 292-303.
- Bell, David R., Teck-Hua Ho, and Christopher S. Tang. (1998). "Determining Where to Shop: Fixed and Variable Costs of Shopping", *Journal of Marketing Research* 35, 352-369.
- Blattberg, Robert C., Gary D. Eppen, and Joshua Lieberman. (1981). "A Theoretical and Empirical Evaluation of Price Deals for Consumer Non-durables", *Journal of Marketing* 45, 116-129.
- Bucklin, Randolph E. and Sunil Gupta. (1992). "Brand Choice, Purchase Incidence and Segmentation: An Integrated Modeling Approach", *Journal of Marketing Research* 29, 201-215.
- Bucklin, Randolph E., Sunil Gupta, and S. Siddarth. (1998). "Determining Segmentation in Sales Response Across Purchase Behaviors", *Journal of Marketing Research* 35, 189-197.
- Chiang, Jeongwen. (1991). "A Simultaneous Approach to the Whether, When and How Much to Buy Questions", *Marketing Science* 10, 297-315.
- Chintagunta, Pradeep K.. (1993). "Investigating Brand Choice, Purchase Incidence and Purchase Quantity Decisions of Households", *Marketing Science* 12, 184-208.
- DiPasquale, Denise and William C. Wheaton. (1996). *Urban Economics and Real Estate Markets*. Englewood Cliffs: Prentice Hall.
- Glaeser, Edward L., Joseph Gyourko, and Christian A. L. Hilber. (2001). "Housing Affordability and Land Prices: Is There a Crisis in American Cities?", Mimeo, The Wharton School, University of Pennsylvania.
- Glaeser, Edward L. and Joseph Gyourko. (2001). "Urban Decline and Durable Housing", NBER Working Paper Number w8598, November.

- Glaeser, Edward L. and Joseph Gyourko. (2002). "The Impact of Zoning on Housing Affordability", Harvard Institute of Economic Research, Discussion Paper Number 1948, March.
- Gupta, Sunil. (1988), "Impact of Sales Promotions on When, What, and How Much to Buy", *Journal of Marketing Research* 25, 342-355.
- Hendel, Igal and Aviv Nevo. (2003). "Measuring the Implications of Sales and Consumer Stockpiling Behavior", Mimeo, University of California, Berkeley.
- Hendel, Igal and Aviv Nevo. (2004). "Sales and Consumer Inventory", NBER Working Paper 9048.
- Kahneman, Daniel and Amos Tversky. (1979). "Prospect Theory: An Analysis of Decision under Risk", *Econometrica* 47, 263-291.
- Krishnamurthi, Lakshaman, Tridib Mazumdar, and S.P. Raj. (1992). "Asymmetric Response to Price and Consumer Brand Choice and Purchase Quantity Decisions", *Journal of Consumer Research* 19, 387-400.
- Lach, Saul. (2002). "Existence and Persistence of Price Dispersion: An Empirical Analysis", *Review of Economics and Statistics* 84, 433-445.
- Lal, Rajiv and Carmen Matutes. (1994). "Retail Pricing and Advertising Strategies", *Journal of Business* 67, 345-370.
- Lal, Rajiv and Ram C. Rao. (1997). "Supermarket Competition: The Case of Everyday Low Pricing", *Marketing Science* 16, 60-80.
- Narasimhan, Chakravarthi. (1984). "A Price Discrimination Theory of Coupons", *Marketing Science* 3, 128-147.
- Pauwels, Koen, Dominique M. Hanssens, and S. Siddarth. (2002). "The Long Term Effects of Pricing and Promotions on Category Traffic, Brand Choice and Purchase Quantity", *Journal of Marketing Research* 39, 421-439.
- Raju, Jagmohan S., Rajiv Lal, and V. Srinivasan. (1990). "The Effects of Brand Loyalty on Competitive Price Promotional Strategies", *Management Science* 36, 276-304.
- Salop, Steven and Joseph E. Stiglitz. (1982). "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents", *American Economic Review* 72, 1121-1130.

- Simester, Duncan I. (1995). "Signaling Price Image Using Advertised Prices", *Marketing Science* 14, 166-188.
- Sorenson, Alan T. (2000). "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs", *Journal of Political Economy* 108, 833-850.
- Stigler, Joseph. (1961). "The Economics of Information", *Journal of Political Economy* 69, 213-225.
- van Heerde, Harald J., Peter S. H. Leeflang, and Dick R. Wittink. (2000). "The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data", *Journal of Marketing Research* 37, 383-395.
- van Heerde, Harald J., Sunil Gupta, and Dick R. Wittink. (2003). "Is 75% of the Sales Promotion Bump due to Brand Switching? No, Only 33% Is", *Journal of Marketing Research* 40, 481-491.
- Varian, Hal R. (1980). "A Model of Sales", *American Economic Review* 70, 651-659.

Summary Statistics and Regression Tables

Table 1
Consumer Characteristics: Variable List and Means
 N=996

Variable	Mean	Standard Deviation	Minimum	Maximum
Estimated housing costs per square foot (in \$), based on zip code characteristics (Household storage cost proxy no. 1)	188.6	228.5	23.3	1392.3
Estimated housing costs per square foot, based on individual consumer and zip code characteristics (Proxy no. 2), <i>N=991</i>	266.2	342.8	25.1	1131.7
Estimated housing unit size in square foot, based on zip code characteristics (Household storage cost proxy no. 3)	1399.2	390.0	125.7	2454.2
Estimated housing unit size in SF, based on individual consumer and zip code characteristics (Proxy no. 4), <i>N=991</i>	1217.9	444.1	192.7	2494.2
Total number of shopping trips	148.3	83.4	38	709
Weighted distance to stores (in miles, weight=trips to each store)	1.6	2.5	0	59.8
Total spending in all stores	4088.0	2099.4	1809.2	15652.0
Household income	34684.2	21776.2	5000	75000
Household size	2.6	1.4	1	6
Average household age > 65 (dummy)	.39	.49	0	1
At least one household member is unemployed (dummy)	.26	.44	0	1
Household is Black (dummy)	.16	.36	0	1
Household is Hispanic (dummy)	.04	.20	0	1
Dummy for household shops in 'downtown market'	.52	.50	0	1
Store dummies (= 1 if consumer residence is closest to store)				
- Store ID = 1419	.12	.32	0	1
- Store ID = 1420	.14	.35	0	1
- Store ID = 1422	.038	.19	0	1
- Store ID = 1423	.23	.42	0	1
- Store ID = 1424	.14	.35	0	1
- Store ID = 1521	.33	.47	0	1
- Store ID = 1522	.090	.29	0	1
- Store ID = 1542	.035	.18	0	1
- Store ID = 1558	.026	.16	0	1

Sources: Stanford Market Database 1991-1993, American Housing Survey 1989 (national sample), and US Census 1990.

Table 2
Description of Five Product Categories by Store

		Downtown Market					Urban Fringe Market			
		Store ID 1419	Store ID 1420	Store ID 1422	Store ID 1423	Store ID 1424	Store ID 1521	Store ID 1522	Store ID 1542	Store ID 1558
Paper Towels	Number of SKUs	32	32	25	29	31	33	15	26	29
	Average Volume	2.4	2.0	1.6	1.6	1.3	1.7	1.5	1.6	2.1
	Average Unit Price	0.73	0.80	0.77	0.95	0.94	0.72	0.83	0.76	0.83
Bathroom Tissue	Number of SKUs	25	28	26	24	22	27	17	27	25
	Average Volume	6.4	6.1	4.9	5.0	4.9	4.8	4.9	5.3	5.4
	Average Unit Price	0.31	0.32	0.32	0.36	0.30	0.32	0.37	0.28	0.31
Liquid Detergents	Number of SKUs	50	54	48	37	42	51	21	48	43
	Average Volume	5.6	5.9	4.3	5.1	3.8	5.5	4.8	5.1	5.2
	Average Unit Price	0.93	0.96	1.07	1.14	1.22	0.93	1.09	1.00	1.15
Ground Coffee	Number of SKUs	41	55	39	40	38	27	19	45	32
	Average Volume	28.3	29.3	27.4	25.6	28.9	34.0	25.6	29.1	32.5
	Average Unit Price	0.17	0.20	0.20	0.23	0.25	0.14	0.19	0.20	0.17
Pills in Capsules	Number of SKUs	65	55	33	22	21	29	8	73	28
	Average Volume	72.2	72.2	44.4	51.4	44.3	36.6	60.7	73.3	77.7
	Average Unit Price	0.086	0.080	0.108	0.114	0.111	0.117	0.089	0.088	0.090

Source: Stanford Market Database 1991-1993. All averages are weighted by the number of purchases for each product, in each store, in each week. The downtown market includes two every-day-low-price stores (ID 1419 and ID 1420). Volume is in IRI-defined standard units: Rolls (paper towels, bathroom tissue), 16-ounce packs (liquid detergents), ounces (ground coffee), and individual capsules (pills in capsules). Prices are in US dollars.

Table 3
Do Consumers with High Housing Costs per Square Foot Buy Goods More Often?
—Purchase Frequency Regression Results

Dependent Variable: Total Number of Trips to All Stores

Specification Explanatory Variable	Zip code level storage cost measure, 1990		Household specific storage cost measure, 1990	
	(1)	(2)	(3)	(4)
Housing costs per square foot (Preferred proxy for household storage costs)	.065 *** (.018)	.095 *** (.020)	.042 *** (.012)	.065 *** (.015)
Weighted distance to stores (in miles, weight=trips to each store)	-.76 (.54)	-.71 (.50)	-.67 (.53)	-.62 (.51)
Total spending in all stores	.0075 *** (.0016)	.0078 *** (.0016)	.0069 *** (.0014)	.0071 *** (.0015)
Household income	-.00064 *** (.00016)	-.00064 *** (.00017)	-.00055 *** (.00015)	-.00055 *** (.00016)
Household size	-8.3 *** (2.1)	-8.5 *** (2.1)	-7.7 *** (2.1)	-7.7 *** (2.1)
Average household age > 65	8.5 (6.9)	7.4 (7.1)	10.5 (6.9)	9.6 (7.1)
At least one member of household is unemployed (dummy equals 1 if true)	3.5 (6.5)	4.2 (6.6)	2.1 (6.3)	2.8 (6.4)
Race of household is Black	18.2 ** (8.3)	18.8 ** (8.3)	15.6 ** (7.8)	16.1 ** (7.7)
Race of household is Hispanic	-12.4 (8.9)	-13.5 (9.4)	-13.0 (8.8)	-13.6 (9.4)
Household shops in 'downtown market'	-22.5 *** (6.6)		-19.8 *** (6.4)	
Store dummies (equal 1 if consumer residence is closest to store)	No	Yes	No	Yes
Constant	155.3 *** (10.4)	124.3 *** (12.9)	152.7 *** (10.6)	115.8 *** (22.3)
Number of observations (consumers)	996	996	991	991
Adjusted R ²	.11	.12	.11	.12

Notes: Numbers in parentheses are robust standard errors. *** Significantly different from zero with 99 percent confidence. ** Significantly different from zero with 95 percent confidence. * Significantly different from zero with 90 percent confidence. Results are little changed if statistically insignificant or marginally significant variables are dropped.

Table 4
Do Consumers with Small Housing Units Buy Goods More Often?
—Purchase Frequency Regression Results with

Dependent Variable: Total Number of Trips to All Stores

Specification Explanatory Variable	Zip code level storage cost measure, 1990		Household specific storage cost measure, 1990	
	(1)	(2)	(3)	(4)
Housing unit size (Alternative proxy for household storage costs)	-0.027 *** (.0098)	-0.042 *** (.014)	-0.024 *** (.0091)	-0.035 *** (.012)
Weighted distance to stores (in miles, weight=trips to each store)	-0.87 (.56)	-1.1 ** (.53)	-0.81 (.54)	-1.0 ** (.51)
Total spending in all stores	.0078 *** (.0016)	.0078 *** (.0016)	.0070 *** (.0015)	.0071 *** (.0015)
Household income	-0.00055 *** (.00015)	-0.00059 *** (.00017)	-0.00046 *** (.00015)	-0.00047 *** (.00016)
Household size	-9.5 *** (2.1)	-9.1 *** (2.1)	-8.9 *** (2.1)	-8.6 *** (2.1)
Average household age > 65	10.4 (6.9)	9.3 (7.1)	12.4 * (6.8)	11.6 * (7.1)
At least one member of household is unemployed (dummy equals 1 if true)	4.9 (6.5)	4.7 (6.6)	3.6 (6.3)	3.2 (6.5)
Race of household is Black	22.3 *** (8.0)	21.4 *** (8.2)	19.1 ** (7.5)	17.8 ** (7.6)
Race of household is Hispanic	-15.2 * (9.1)	-15.6 * (9.5)	-15.9 * (9.1)	-16.3 * (9.5)
Household shops in 'downtown market'	-19.3 *** (7.1)		-18.3 *** (7.2)	
Store dummies (equal 1 if consumer residence is closest to store)	No	Yes	No	Yes
Constant	201.6 *** (19.2)	242.1 *** (37.1)	190.0 *** (16.9)	187.7 *** (27.1)
Number of observations (consumers)	996	996	991	991
Adjusted R ²	.11	.11	.10	.11

Notes: Numbers in parentheses are robust standard errors. *** Significantly different from zero with 99 percent confidence. ** Significantly different from zero with 95 percent confidence. * Significantly different from zero with 90 percent confidence. Results are little changed if statistically insignificant or marginally significant variables are dropped.

Table 5**Do Consumers with High Storage Costs Buy Smaller Quantities?—Average Quantity Regressions for High-Storage-Use Items**

Dependent Variable: Average purchase quantity per customer and trip

Explanatory Variable	Paper Towels (A)				Bathroom Tissue (B)			
	(1A)	(2A)	(3A)	(4A)	(1B)	(2B)	(3B)	(4B)
Storage cost proxy 1: Housing costs per sq. foot, median of zip code	-.00085 *** (.00019)		-.0012 *** (.00043)		-.0015 *** (.00053)		-.0025 *** (.00085)	
Storage cost proxy 2: Housing costs per sq. foot, household specific		-.00056 *** (.00012)		-.00085 *** (.00031)		-.0010 *** (.00035)		-.0018 *** (.00061)
Weighted distance to stores (in miles, weight=trips to each store)	-.0041 (.0069)	-.0041 (.0069)	-.012 (.0081)	-.013 (.0084)	.025 (.024)	.025 (.024)	.011 (.023)	.0091 (.023)
Unit price (average)	-1.0 *** (.18)	-1.0 *** (.18)	-1.0 *** (.19)	-1.1 *** (.19)	-10.4 *** (.62)	-10.4 *** (.62)	-10.3 *** (.62)	-10.4 *** (.62)
Total spending in all stores (x 10 ⁻³)	.0060 (.0065)	.0059 (.0067)	.013 ** (.0060)	.013 ** (.0061)	.015 (.013)	.014 (.014)	.025 * (.014)	.024 * (.014)
Household income (x 10 ⁻³)	.0034 ** (.0017)	.0028 (.0017)	.0036 ** (.0016)	.0028 * (.0017)	.0099 ** (.0045)	.0088 ** (.0044)	.0094 ** (.0046)	.0079 * (.0045)
Household size	.015 (.027)	.0012 (.027)	.013 (.027)	.0079 (.028)	.14 ** (.056)	.14 ** (.057)	.14 ** (.057)	.13 ** (.058)
Average household age >65	.016 (.084)	.0045 (.084)	.027 (.081)	.017 (.081)	-.018 (.19)	-.034 (.18)	.0053 (.19)	-.011 (.19)
At least one member of household is unemployed (dummy equals 1 if true)	-.058 (.056)	-.057 (.056)	-.073 (.058)	-.073 (.059)	.16 (.18)	.15 (.18)	.11 (.18)	.11 (.18)
Race of household is Black	-.028 (.070)	-.029 (.070)	-.053 (.073)	-.054 (.073)	.056 (.29)	.044 (.29)	.029 (.29)	.017 (.29)
Race of household is Hispanic	-.18 (.14)	-.18 (.14)	-.21 (.14)	-.22 (.14)	.31 (.48)	.31 (.47)	.27 (.48)	.24 (.48)
Household shops in 'downtown market'	.41 *** (.091)	.39 *** (.089)			1.2 *** (.23)	1.1 *** (.23)		
Store dummies (for closest store)	No	No	Yes	Yes	No	No	Yes	Yes
Constant	2.3 *** (.22)	2.3 *** (.22)	3.0 *** (.70)	2.9 *** (.41)	7.6 *** (.31)	7.6 *** (.31)	9.3 *** (.88)	9.2 *** (.76)
Number of observations (consumers)	954	949	954	949	981	976	981	976
Adjusted R ²	.070	.070	.094	.096	.23	.23	.24	.25

Notes: Numbers in parentheses are robust standard errors. *** / ** / * Significantly different from zero with 99 percent / 95 percent / 90 percent confidence. Results are little changed if statistically insignificant variables are dropped. Regressions with alternative storage cost proxies provide qualitatively very similar results.

Table 5—Continued

Do Consumers with High Storage Costs Buy Smaller Quantities?—Average Quantity Regressions for High-Storage-Use Items

Dependent Variable: Average purchase quantity per customer and trip

Explanatory Variable	Liquid Detergent (C)			
	(1C)	(2C)	(3C)	(4C)
Storage cost proxy 1: Housing costs per sq. foot, median of zip code	-.0011 *** (.00043)		-.0020 *** (.00077)	
Storage cost proxy 2: Housing costs per sq. foot, household specific		-.00075 *** (.00028)		-.0015 *** (.00052)
Weighted distance to stores (in miles, weight=trips to each store)	-.014 (.015)	-.014 (.015)	-.011 (.016)	-.013 (.016)
Unit price (average)	-2.9 *** (.15)	-2.9 *** (.14)	-2.8 *** (.15)	-2.8 *** (.14)
Total spending in all stores (x 10 ⁻³)	.030 * (.017)	.025 (.016)	.023 (.017)	.018 (.017)
Household income (x 10 ⁻³)	.0068 * (.0036)	.0060 * (.0035)	.0087 ** (.0036)	.0075 ** (.0037)
Household size	.17 ** (.068)	.16 ** (.068)	.17 ** (.069)	.16 ** (.068)
Average household age >65	-.042 (.18)	-.093 (.18)	-.032 (.18)	-.082 (.18)
At least one member of household is unemployed (dummy equals 1 if true)	.091 (.19)	.093 (.19)	.089 (.19)	.090 (.19)
Race of household is Black	-.033 (.19)	-.047 (.19)	-.047 (.19)	-.062 (.19)
Race of household is Hispanic	-.0024 (.47)	-.00084 (.47)	-.078 (.47)	-.099 (.47)
Household shops in ‘downtown market’	.49 ** (.21)	.47 ** (.21)		
Store dummies (equal 1 if consumer residence is closest to store)	No	No	Yes	Yes
Constant	7.6 *** (.29)	7.7 *** (.29)	8.5 *** (.58)	7.9 *** (.47)
Number of observations (consumers)	861	857	861	857
Adjusted R ²	.38	.38	.39	.39

Notes: Numbers in parentheses are robust standard errors. ***/**/* Significantly different from zero with 99/95/90 percent confidence. Results are little changed if statistically insignificant variables are dropped. Regressions with alternative storage cost proxies give qualitatively similar results, although at lower confidence levels.

Table 6**Do Consumers with High Storage Costs Buy Smaller Quantities?—Average Quantity Regressions for Low-Storage-Use Items**

Dependent Variable: Average purchase quantity per customer and trip

Explanatory Variable	Ground Coffee (D)				Pills in Capsule Form (E)			
	(1D)	(2D)	(3D)	(4D)	(1E)	(2E)	(3E)	(4E)
Storage cost proxy 1: Housing costs per sq. foot, median of zip code	-.0014 (.0024)		-.0036 (.0030)		-.0081 (.010)		-.0092 (.011)	
Storage cost proxy 2: Housing costs per sq. foot, household specific		-.00082 (.0017)		-.0023 (.0024)		-.0050 (.0068)		-.0070 (.0079)
Weighted distance to stores (in miles, weight=trips to each store)	-.098 (.072)	-.099 (.072)	-.094 (.077)	-.096 (.078)	.0060 (.28)	.0066 (.28)	-.45 (.54)	-.46 (.54)
Unit price (average)	-71.7 *** (3.1)	-72.3 *** (3.2)	-72.0 *** (3.2)	-72.6 *** (3.3)	-819.7 *** (73.3)	-821.3 *** (73.5)	-808.2 *** (69.3)	-808.0 *** (69.5)
Total spending in all stores (x 10 ⁻³)	.011 (.067)	-.014 (.068)	.010 (.063)	-.016 (.064)	-.70 ** (.29)	-.69 ** (.31)	-.40 (.30)	-.40 (.32)
Household income (x 10 ⁻³)	-.0085 (.017)	-.0072 (.017)	-.0060 (.017)	-.0057 (.017)	.16 (.099)	.15 (.097)	.16 (.10)	.15 (.098)
Household size	.78 ** (.33)	.78 ** (.34)	.76 ** (.33)	.76 ** (.34)	-3.7 *** (1.0)	-3.7 *** (1.1)	-3.8 *** (1.1)	-3.9 *** (1.1)
Average household age >65	.72 (1.0)	.72 (1.0)	.80 (1.0)	.78 (1.0)	1.4 (3.6)	1.1 (3.6)	1.6 (3.5)	1.4 (3.5)
At least one member of household is unemployed (dummy equals 1 if true)	-1.1 (.77)	-1.1 (.77)	-1.1 (.77)	-1.1 (.78)	-1.1 (3.6)	-1.2 (3.7)	-.68 (3.6)	-.72 (3.6)
Race of household is Black	-.35 (1.7)	-.51 (1.7)	-.40 (1.6)	-.59 (1.7)	2.4 (6.7)	2.2 (6.7)	-.19 (6.1)	-.13 (6.2)
Race of household is Hispanic	-3.0 (1.8)	-2.9 (1.8)	-2.4 (1.9)	-2.4 (1.9)	6.4 (9.8)	6.5 (9.8)	3.6 (9.7)	3.6 (9.7)
Household shops in 'downtown market'	.11 (.89)	.13 (.89)			-2.4 (4.3)	-2.8 (4.2)		
Store dummies (for closest store)	No	No	Yes	Yes	No	No	Yes	Yes
Constant	41.4 *** (1.6)	41.5 *** (1.6)	42.1 *** (2.6)	42.9 *** (2.6)	148.5 *** (9.1)	149.1 *** (9.0)	140.8 *** (9.5)	188.1 *** (25.9)
Number of observations (consumers)	760	755	760	755	573	569	573	569
Adjusted R ²	.42	.41	.42	.42	.43	.43	.46	.46

Notes: Numbers in parentheses are robust standard errors. *** / ** / * Significantly different from zero with 99 percent / 95 percent / 90 percent confidence. Results are little changed if statistically insignificant variables are dropped. Regressions with alternative storage cost proxies provide qualitatively very similar results.

Table 7
Do Storage Costs and Constraints of Consumers Affect the Pricing Behavior of Stores?

Dependent Variables: (1) Average price, (2) probability of promotion, (3) promotional depth

	Average Purchase Price		Probability of Promotion – Log Transform		Promotional Depth – Log Transform	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Average of consumers' (zip code specific) housing costs per SF faced by store	.00039 ** (.00017)		-.0034 *** (.00068)		-.0039 *** (.00053)	
Average (zip code specific) housing unit size of consumers faced by store (x 10⁻³)		-.15 ** (.063)		1.2 *** (.23)		.96 *** (.19)
Average household income faced by store (x 10 ⁻³)	.024 *** (.0065)	.026 *** (.0066)	-.083 *** (.022)	-.088 *** (.022)	-.022 (.022)	-.00084 (.022)
Percentage of households with average age above 65 faced by store	2.1 *** (.38)	2.3 *** (.39)	-5.8 *** (1.3)	-6.5 *** (1.3)	-.057 (1.2)	.68 (1.3)
Average household size faced by store	.23 ** (.10)	.16 ** (.069)	-2.0 *** (.41)	-1.2 *** (.25)	-2.0 *** (.34)	-.63 *** (.23)
Percentage of households with at least one unemployed member faced by store	-1.5 *** (.45)	-.96 *** (.33)	5.8 *** (1.7)	.51 (1.2)	10.0 *** (1.5)	3.6 *** (1.1)
Percentage of Black households faced by store	-.17 (.13)	-.093 (.13)	1.4 *** (.41)	1.1 *** (.41)	.15 (.37)	-.0064 (.40)
Percentage of Hispanic households faced by store	-.032 (.41)	-.12 (.38)	-10.1 *** (1.4)	-8.7 *** (1.2)	-4.7 *** (1.4)	-2.1 * (1.3)
Number of weeks product is on shelf	-.00080 *** (.00027)	-.00089 *** (.00027)	-.0096 *** (.0011)	-.0085 *** (.0011)	.012 *** (.00079)	.013 *** (.00080)
SKU fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.0 *** .59	2.2 *** .50	7.2 *** (2.1)	4.6 *** (1.7)	1.9 (1.9)	-3.0 * (1.7)
Number of observations (SKUs x stores)	4760	4760	2031	2031	3669	3669
Adjusted R ²	.98	.98	.62	.62	.50	.49

Notes: Numbers in parentheses are robust standard errors. ***/**/* Significantly different from zero with 99/95/90 percent confidence. The qualitative results and the significance levels with regard to the coefficient on our storage cost proxy measures are little changed if the proxies are calculated based on the household specific information from the AHS. Specifications that use the shares of households with very high square footage housing costs or limited housing space as explanatory variables also lead to qualitatively similar results. Qualitative results are also similar if only SKU's are included with information for all 9 stores.

Table 8**Quantitative Effects: Effect of Consumer Move from Urban Fringe to Downtown**

Assumptions:

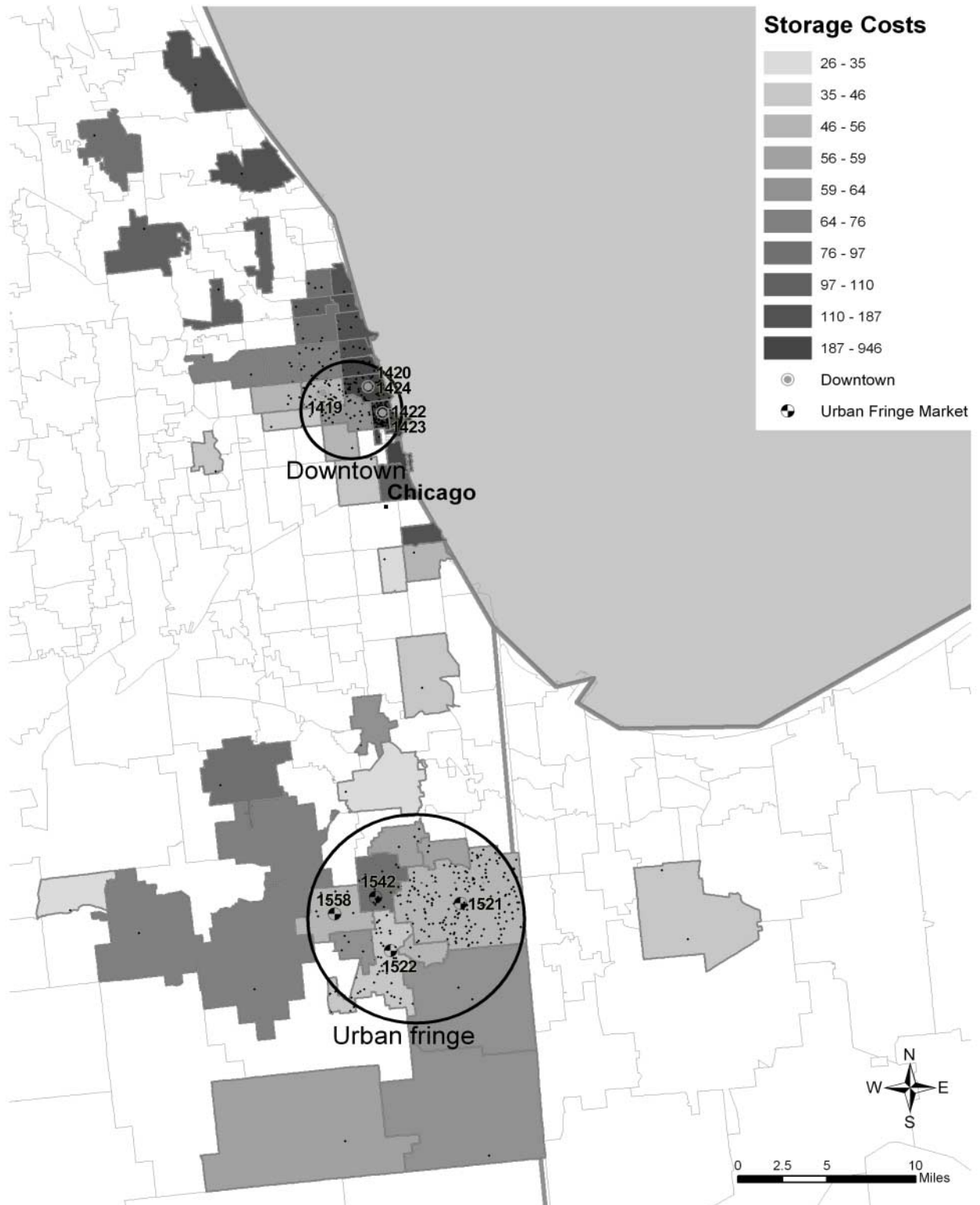
i) Estimated value of storage space per square foot increases from 40 to 400 US dollar

ii) Estimated housing unit size decreases from 1600 to 1000 square feet

Specification	Change in Purchase Frequency	% Change in Average Purchase Quantity					% Change in Offered Product Price	Change in Promotion Probability	Change in Promotional Depth
		(A) Towels	(B) Tissue	(C) Detergent	(D) Coffee	(E) Pills			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Tab 3 (1) ⁱ	+15.8% (+23.4 trips)								
Tab 3 (2) ⁱ	+23.1% (+34.2 trips)								
Tab 3 (3) ⁱ	+10.2% (+15.1 trips)								
Tab 3 (4) ⁱ	+15.8% (+23.4 trips)								
Tab 4 (1) ⁱⁱⁱ	+11.0% (+16.2 trips)								
Tab 4 (2) ⁱⁱⁱ	+17.1% (+25.2 trips)								
Tab 4 (3) ⁱⁱⁱ	+9.7% (+14.4 trips)								
Tab 4 (4) ⁱⁱⁱ	+14.2% (+21.0 trips)								
Tab 5/6 (1) ⁱ		-18.1%	-10.2%	-7.0%	[-1.6%]	[-4.1%]			
Tab 5/6 (2) ⁱ		-11.9%	-6.8%	-4.8%	[-0.9%]	[-2.5%]			
Tab 5/6 (3) ⁱ		-25.6%	-16.9%	-12.7%	[-4.1%]	[-4.7%]			
Tab 5/6 (4) ⁱ		-18.1%	-12.2%	-9.5%	[-2.6%]	[-2.6%]			
Tab 7 (1a) ⁱ							+3.5%		
Tab 7 (1b) ⁱⁱⁱ							+2.3%		
Tab 7 (2a) ⁱ								-69.5% (-3.4% pts)	
Tab 7 (2b) ⁱⁱⁱ								-50.1% (-2.5% pts)	
Tab 7 (3a) ⁱ									-70.8% (-15.0% pts)
Tab 7 (3b) ⁱⁱⁱ									-34.7% (-7.4% pts)

Note: Percentage changes are measured at the urban fringe-sample averages. The average number of shopping trips of all customers within the two year period and within the urban fringe market is 147.8. The average unit sizes sold in the urban fringe market are as follows (see also Table 2): 1.7 units (rolls of paper towels), 5.3 units (rolls of bathroom tissue), 5.7 units (16oz packs of liquid detergent), 31.9 units (ounces of ground coffee), and 71.0 units (individual pills in capsules). The average basket item (SKU) price in the urban fringe market is \$3.97, the probability that an SKU is featured within a particular urban fringe store at a particular week is 4.9%, and the promotional depth in the urban fringe is 21.3%. Quantitative effects in [square brackets] are not statistically significant at the 90% level.

Appendix 1: Map of Chicago Downtown and Urban Fringe Submarkets



Appendix 2: Distribution of Consumer Storage Costs

Figure A3-1: Distribution of Zip Code Specific Imputed Consumer Storage Costs and Housing Unit Sizes for Submarkets

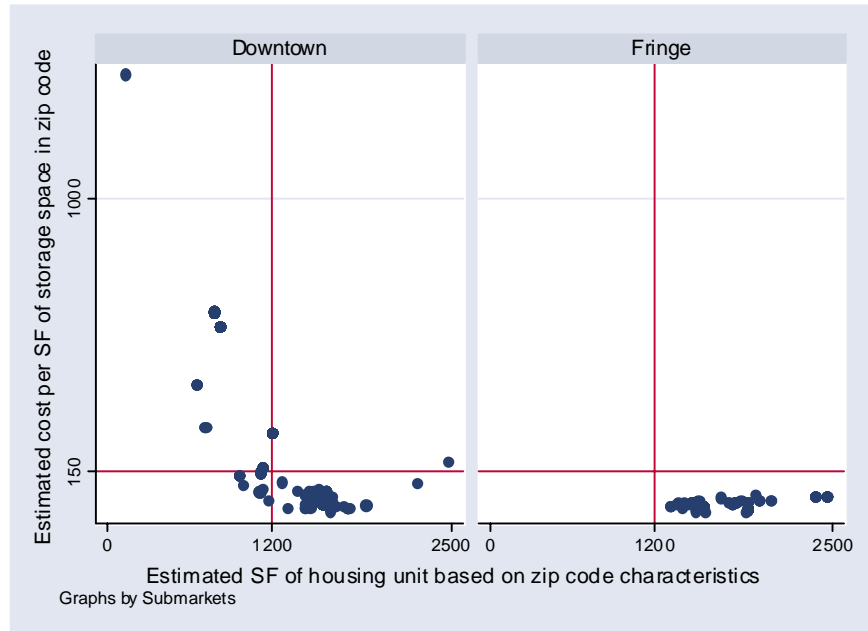
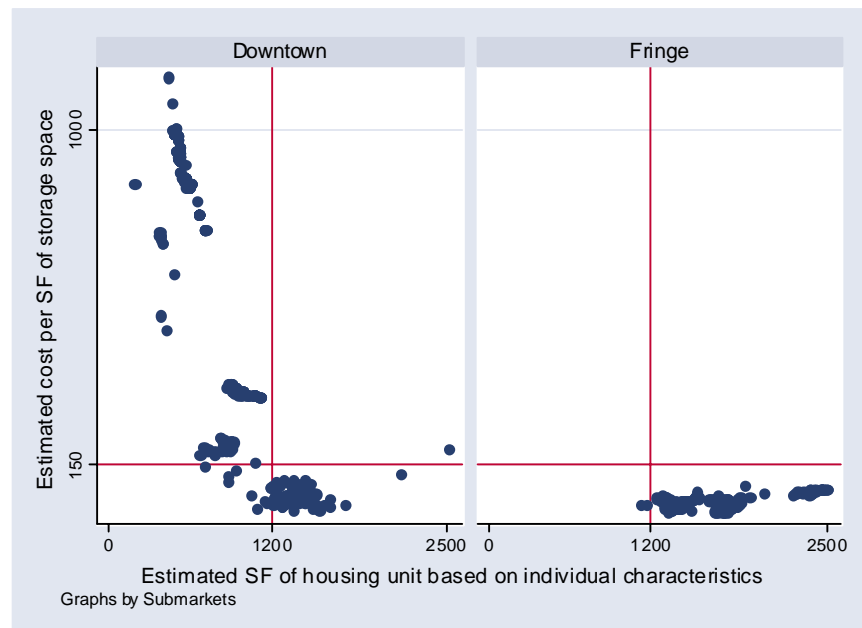


Figure A3-2: Distribution of Household's Imputed Storage Costs and Housing Unit Sizes for Submarkets



Appendix 3: Imputation Method for Storage Cost Proxy Measures

This Appendix provides details for the creation of our two variations (zip level and household level) of the proxy measures for storage cost. We begin by describing the zip code level proxy measure, and then discuss the method for obtaining the proxy measure that differs by panelist.

Zip Code Level Storage Cost Measure

The median value of a housing unit in a particular zip code is derived from the 1990 US Census. The average square footage of living area must be imputed for each zip code as it is not collected by the Census. Because the AHS contains square footage information, we begin by estimating square footage in that data set, using a number of variables that are common to the AHS and the Stanford Market Basket data. We use the AHS for 1989—the closest available year to the 1990 US Census—to estimate the square footage of living area of an average housing unit in an MSA j as the zip code of the occupants is not disclosed. The estimating equation is as follows:

$$(A1) \quad \phi \text{ size living area}_j = \beta_0 + \beta_1 \phi \text{ age of building}_j + \beta_2 \phi \text{ \# of rooms}_j \\ + \beta_3 \% \text{ units detached}_j + \beta_4 \% \text{ units attached}_j + \varepsilon.$$

The adjusted R^2 is 56.5%. We first adjust the size of the estimated coefficients from equation (A1) in order to control for sample differences between the US Census data and the AHS. We do so by computing the ratio of the predicted average median living area per zip code for all US zip codes to the median of the predicted square footage of living area from the national AHS sample and by then multiplying this adjustment factor with the coefficients. We then use these adjusted coefficients to impute the average square footage of living area for the zip codes of all panelists. Our first storage cost proxy measure can then be derived as the median value of a housing unit in zip code j divided by the predicted square footage of living area of zip code j .

Panelist Specific Storage Cost Measure

In a first step we use the National AHS 1989 to impute individual house values as a function of numerous demographic characteristics of the occupants, housing unit specific characteristics, and metropolitan area fixed effects. The estimating equation is as follows:

$$\begin{aligned}
\text{(A2)} \quad \text{house value}_i = & \beta_0 + \beta_1 \text{ race dummies}_i + \beta_2 \text{ education dummies}_i \\
& + \beta_3 \text{ age dummies}_i + \beta_4 \text{ children}_i + \beta_5 \text{ income category dummies}_i \\
& + \beta_6 \text{ \# of rooms}_i + \beta_7 \text{ \# of bathrooms}_i + \beta_8 \text{ age of building}_i \\
& + \beta_9 \text{ age of building}_i^2 + \beta_{10} \text{ housing type dummies}_i + \beta_{11} \text{ basement}_i \\
& + \beta_{12} \text{ garage}_i + \beta_{13} \text{ housing unit quality}_i + \beta_{14} \text{ neighborhood quality}_i \\
& + \beta_{15} \text{ MSA status dummies}_i + \beta_{16} \text{ MSA fixed effects}_i + \varepsilon.
\end{aligned}$$

The adjusted R^2 of the regression is 59.3%. We adjust the size of the estimated coefficients from equation (A2) in order to control for sample differences between the US Census data and the AHS. We do so by computing the ratio of the average median house values for all US zip codes to the median of the predicted house values from the national AHS sample and by then multiplying this adjustment factor with the coefficients.

We then use these adjusted coefficients to impute individual house values. We first calculate the differences between the zip code averages (from the 1990 Census) and the individual values for each panelist (from the market basket data). We then impute the predicted house value for each panelist i in zip code j as:

$$\text{(A3)} \quad \text{house value panelist}_i = \text{med house value zip code}_j(i) + \hat{\beta}_{adj.} \times (X_i - X_j).$$

The factor $\hat{\beta}_{adj.}$ denotes the adjusted predicted coefficient and the vectors of variables X_j and X_i denote the average demographic characteristics of the zip code and the demographic characteristics of the panelists. For multiple categorical variables adjustment only occurs for the categorical variable that is true for the panelist. For example, if the panelist belongs to income category 3, then adjustment only occurs for the coefficient for income category 3 but not for all other income categories. For binary categorical variables the adjusted coefficient is multiplied by the difference of the value that is true for the panelist (e.g., the household has children) minus the average value for the zip code (e.g., percentage of households with children in zip code).

In a next step we impute the panelists' square footage of living area using the same method as summarized previously for the imputation of the house value of the panelist. Finally, we can derive our individual measure by dividing the imputed house value of panelist i by the imputed square footage of living area of panelist i . We proceed in an analogous manner when we construct measures of housing unit size (the storage cost proxy used in Tables 4, 7, and 8).