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

David R. Bell *

and

Christian A. L. Hilber **

This Version: December 20, 2005

We thank seminar participants at the Marketing Science Conference 2003 at College Park, the Tjalling C. Koopmans Workshop "Issues on the Economics of Pricing" 2004 at Utrecht University, Singapore Management University, The Department of Justice (Washington, DC), and Tilburg University for helpful comments. We are very grateful to the Editor and two anonymous *QME* Reviewers for many constructive suggestions that improved the paper and also to Sangyoung Song for research assistance. Maciej Szymanowski also provided detailed comments. 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 Constraints Affect Consumer and Store Behavior?

This Version: December 20, 2005

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

Abstract

We revisit and test Salop and Stiglitz (1982) Theory of Sales. Equilibrium comparative static predictions are that greater consumer storage constraints lead to: (1) higher average prices, (2) fewer promotions, and (3) shallower promotions. In equilibrium, price dispersion is nonlinear in storage constraints, first increasing then decreasing. Empirical estimates of storage constraints are developed for approximately 1,000 households using the American Housing Survey (1989), United States Census (1990), and Stanford Market Basket Database (1991-1993). We find consumers with greater storage constraints shop more often and purchase smaller quantities per visit; moreover, the comparative static predictions are supported and evidence consistent with the equilibrium dispersion prediction is observed. Estimated quantitative effects are economically important.

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

JEL CLASSIFICATION: D12, D40, M3.

Price variation for homogenous goods is common in many product markets. Typical explanations relate to differences in firm characteristics (Simester 1995), or consumer heterogeneity in price knowledge or willingness to undertake effort to obtain lower prices (e.g., Narasimhan 1984; Varian 1980). This leads to important questions regarding how consumers might respond to prices, and implicitly, how firms should set prices. 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, 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 article, we provide an empirical test of the consumer behavior assumptions and firm-level pricing decisions embodied in Salop and Stiglitz (1982) Theory of Sales. In our empirical work we use housing unit size as an (inverse) proxy for storage constraints faced by the household.[1] We investigate whether the storage constraints influence 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 meaningful. Moreover, consumer storage constraints are worthy of study because they represent an enduring and structural feature of markets and are not readily mitigated. It is therefore important to understand their implications for consumer and firm behavior.

To identify instances where storage constraints 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 that purchase and consumption patterns may differ across consumers as follows. Households who are constrained in their ability to store products may exhibit purchase patterns such that the observed purchase volume does not vary much between promotion (low) and regular (high) price points. Conversely, households without storage constraints are able to stockpile should they be confronted with a low price. This stockpiling behavior leads to a longer elapsed time until these consumers return to the market.[2]

Some empirical evidence for this stylized fact – that price responsiveness of consumers is related to their storage constraints – is provided in Bucklin and Gupta (1992). They show price sensitivity in purchase incidence decisions for liquid laundry detergent is higher for house dwellers than for apartment dwellers. They interpret this to imply that house dwellers can manage inventory and stockpile, while apartment dwellers have relatively prohibitive storage constraints. Hendel and Nevo (2003) also analyze detergent purchases (but using data from a different metro area and time period) and find that households that are larger and reside in suburban locations hold larger volumes of inventory. They also conclude that these types of households can benefit from non-linear pricing schedules in which larger package sizes are offered at a per unit discount.

A few empirical papers have attempted to test alternative theories that explain price variation. Recent among these is Lach (2002) who provides an empirical test of Varian's (1980) theoretical proposition that, in the presence of informed and uninformed consumers, retail price variation can be explained as the outcome of mixed strategy equilibria. Using data from both durables (refrigerators) and consumables (chicken, flour, paper products) Lach finds that: (1) cross-sectional variation in store prices remained approximately constant over time for these goods, and (2) consistent with the predictions of Varian (1980), stores constantly rearrange their relative positions within the cross-sectional price distribution in a random manner. Sorenson (2000) examines 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. This is consistent with the classic conjecture of Stigler (1961) on consumer motivations to search for lower prices. Narasimhan (1988) derives a model of loyal and switching customers to show how the behavior of the switching segment affects the frequency and size of price discounts. He proposes an empirical strategy but does not formally test the comparative static predictions (see also Narasimhan 1984 for an empirical test of coupon behavior derived from a theoretical analysis). In a related study, Raju, Srinivasan and Lal (1990) show that a "strong" brand with a larger loyal segment promotes less frequently – empirical support is obtained from a non-parametric test of promotion frequency by national ("strong") brands and private labels ("weak" brands).

In a departure from these previous empirical tests of pricing theories, our research focuses on the extent to which the presence of heterogeneity in storage constraints: (a) influences demand patterns across product categories and (b) affects the optimal pricing behavior of competing firms.[3] Consumer storage constraints are posited to interact with category characteristics – the absence of stockpiling behavior for some households should be more pronounced for product categories that require more space. We not only test the price-setting implications of consumer storage constraints, but also propose and implement an empirical method to first isolate the storage constraints themselves. Measures for storage constraints are developed at the household level using data from the Stanford Market Basket database, the American Housing Survey, and the US Census.

This paper provides three new contributions. First, we offer an empirical approach to assessing consumer storage constraints. Our measure is validated by showing that while it predicts average purchase quantities for stockpiled products (e.g., paper towels, bathroom tissues, and liquid detergents) it has no impact on purchase behavior for pills in capsule form – a product that requires essentially no space. Second, we empirically test the predictions from Salop and Stiglitz (1982). Comparative static predictions for average prices, promotion frequency and promotion depth are supported by the data; moreover price dispersion varies with storage constraints in a manner consistent with the theoretical equilibrium. Finally, we derive quantitative effects of storage constraints and show that consumer movement to smaller (that is, more constrained) housing units has non-trivial implications for shopping behavior and retail price setting and promotion policies. 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 Theoretical Framework

Interest in consumer response to price and the influence of market characteristics on the price-setting behavior of firms has a long history. This is understandable given the primacy of price in decision making and the importance of the price mechanism in markets. While factors such as informational differences or heterogeneity in willingness to search have received a good deal of attention, storage constraints and their role have been studied much less frequently. We briefly review the research findings from the marketing literature on consumer response to price, on storage constraints and retail prices and discuss the positioning of our study with respect to these results. Next we describe the key elements of our underlying theoretical framework—Salop and Stiglitz (1982) Theory of Sales.

1.1 Findings from Marketing: Consumer Response to Price

Empirical research 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) uncover distinct market segments with very different stockpiling profiles.

Others have studied price response using aggregate data (e.g., Pauwels, Hanssens, and Siddarth 2002; van Heerde, Leeflang, and Wittink 2000; van Heerde, Gupta, and Wittink 2003) and 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 evidence suggests that 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, while the largest short term price elasticity is 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 aspect 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. A common implicit assumption in most empirical studies in marketing – including those referenced above – is that the observed purchase response (e.g., the volume of product purchased at a particular price point) is a function of price only.[4] Economic theory suggests that storage constraints may play a significant role in the behavior of consumers, and also influence the equilibrium price-setting behavior of firms. We therefore argue that some consumers operate under real external constraints; specifically, their ability to take advantage of low prices may be hindered or furthered by their capacity to store inventory. Firms are cognizant of this and therefore take this into account when setting prices.

1.2 Findings from Marketing: Storage Constraints and Retail Prices

The marketing literature offers a variety of articles that assess the effect of consumer storage constraints on individual purchase behavior and on the incentive of retailers to offer price discounts. Twenty-five years ago Blattberg, Eppen, and Lieberman (1981) introduced storage as an explanation for retail price promotions. If retailers have opportunity costs of storage 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 holding 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 greater product volumes when prices are low and take longer to replenish subsequent supplies following purchases on promotion.[5] An empirical study by Bell, Chiang, and Padmanabhan (1999) suggests similar effects for similar categories (bathroom tissue, ground coffee, liquid detergent, and paper towels). For these categories average volumes increase during price promotions, as does the average time until a subsequent purchase, suggesting stockpiling but no increase in product consumption.

Analytical models of optimal consumer shopping policies under price uncertainty and consumer inventory costs have been formalized in several studies, including the following. Assuncao and Meyer (1993) show that given some memory for price, consumers will rationally increase consumption when faced with a discount. In Krishna (1992) the optimal response to an increase in price dealing over all alternatives in the market is to purchase less of a given brand at a price discount. Meyer and Assuncao (1990) demonstrate that actual consumer behavior deviates systematically from the normative policy for sequential purchasing. Consumers wait too long to buy when prices are rising and buy too quickly as prices come down.

In a formal analysis of both consumer and retailer behavior based on an economic order quantity (EOQ) setup Eppen and Lieberman (1984) analyze the interaction between consumer and retailer holding costs. They show that the retailer's incentive to deal depends critically on the relative magnitudes of retailer and consumer holding costs. Unless retailer holding costs are much greater than those of consumers, it does not pay the retailer to offer deals. A natural implication is that goods such as paper products that have a high opportunity cost of space for the retailer, but are easy for at least some consumers to store are more likely to be promoted. Jeuland and Narasimhan (1985) introduce an alternative rationale for price cuts when consumers are heterogeneous in their inventory holding costs. When demand and inventory holding costs are positively correlated, the seller can profit from price discrimination put into effect through temporary price cuts.

The preceding studies illustrate that: (a) consumer inventory carrying constraints affect purchasing behavior, and (b) firms may exploit heterogeneity or disparities across consumers in inventory costs when setting prices. These findings motivate our work. Our main contribution to this literature is that we provide a direct and simultaneous test of the impact of household specific storage constraints on consumer and store behavior. More specifically, we empirically test predictions, which we derive from a classic equilibrium model developed by Salop and Stiglitz (1982). In tying our empirical work to Salop and Stiglitz (1982) explicit equilibrium analysis of consumer and firm behavior, we implicitly test and corroborate many of the findings from the marketing literature described above. We introduce and summarize the model and its main predictions below.

1.3 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. 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, $f(p)$. Consumers are homogenous with respect to their reservation price, u, for each unit of the product. Furthermore, price uncertainty in the market implies that consumers select stores at random.

Consumers have unit demand in each period for a total of two units over the consumption cycle. Consumers who are able to stockpile, i.e., buy two units in period 1 and store one for future consumption, 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, $\hat{p}$, is given by $\hat{p}+h \le \overline{p}$, where the variables $\overline{p}$ and h represent the expected price from the stationary price distribution, $f(p)$, and the per unit storage cost, respectively. Consumer risk neutrality ensures the expected price is relevant in the purchase decision.[6] Consumer behavior is rational and the stockpiling strategy is undertaken only when the surplus from doing so outweighs that from a 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).

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.[7] Likewise, retailers are rational and choose pricing strategies to maximize prices, given the behavior of their retail competitors. 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 T/n shoppers. The mixed strategy profile contains exactly two prices; the higher of the two equals the reservation price, u, and is charged with probability λ. The lower of the two prices is exactly equal to the pivot price, $\hat{p}$.[8] 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 $S_h = T/n\left[1+\lambda\right]$ and $S_l = T/n\left[2+\lambda\right]$. The equal profits condition in the mixed strategy equilibrium requires that $p_h S_h = p_l S_l$. With these assumptions in hand, the authors derive Theorem 1 (see Salop and Stiglitz 1982, p. 1124) which encompasses three comparative static predictions that we test empirically. The equilibrium solutions, comparative static predictions and empirical implications are as follows:

(i) The average price is $\overline{p} = (3h+u)/2$, so $\dfrac{\partial \overline{p}}{\partial h} = \dfrac{3}{2}$, which implies that average prices increase with consumer storage constraints (<u>Prediction 1</u>),

(ii) The probability of promotion is $(1-\lambda)=(u-3h)/(u-h)$, so $\frac{\partial(1-\lambda)}{\partial h}=\frac{-2u}{(u-h)^2}$, which implies that promotions are less likely when consumers face greater storage constraints (Prediction 2), and

(iii) Promotion depth defined as the size of the discount (high price minus low price) relative to the regular (high) price $=(u-h)/2u$, so $\frac{\partial u}{\partial h}=-\frac{1}{2u}$, which implies discounts, when offered, will be smaller when consumers face greater storage constraints (Prediction 3).

Finally, the theory also makes predictions about price dispersion in equilibrium. However, in contrast to the above predictions, price dispersion is not monotonically increasing or decreasing in h. To begin with, whether there exists dispersion or a single price equilibrium depends exclusively on the magnitude of h relative to the reservation price u. If $h<u/3$ there exists price dispersion, as measured by the coefficient of variation $\sigma_p^2/\overline{p}^2 = 2(u/h-3)/(u/h+3)^2$. An increase in h first increases and then decreases price dispersion. Moreover, if $h \geq u/3$ there exists only one price in equilibrium.[9]

Next we turn to the empirical analysis. We develop a micro-level empirical approach to identifying and measuring storage constraints, demonstrate how household specific storage constraints interact with product-specific storage requirements, and construct our empirical analysis to test the predictions derived from the theoretical framework

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 constraints. We subsequently elaborate on the estimation procedure for the household-level storage constraint proxy and the shopping frequency and purchase quantity estimation results for this variable. 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 stores are 1521, 1522, 1542, and 1558 (1522 and 1558 are from the same chain). 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 household's travel distance for each household-store pair.[10]

The empirical analysis requires both identification of storage constraints for each household in the sample and sufficient variation in storage constraints within and across the two trade areas. A map of both markets which illustrates the spatial distribution of storage constraints and of consumers by zip codes is provided in Appendix 1. In this map, storage constraints are represented by the following measure: Housing unit size (or more precisely: square footage of living area by zip code). We 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 constraint proxy. 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 submarket 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 constraints: The average housing unit size is 1146 square feet in the downtown market and 1674 square feet in the urban fringe. While this is somewhat evident from the map in Appendix 1, the distributions of consumer storage constraints (and storage cost) 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 to 8 we utilize an estimate of housing unit size as an inverse measure of storage constraints (see next subsection for more details). Using this approach, 996 households remain in the dataset.

[----- 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 constraint 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 submarkets in household storage constraints.

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

2.2 Estimation of Storage Constraint Proxy

Central to the contribution of this research is the creation of a reliable estimate of household storage constraints 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 it in recent work in urban economics (e.g., Glaeser, Gyourko, and Hilber (2001) and Glaser and Gyourko (2001 and 2003).[11]

Our preferred proxy measure for household storage constraints is the panelist's housing unit size. We need to impute this measure as the living area of the panelists' homes is not directly observed. Specifically, we use location and demographic information for the panelists and hedonic living area equations to impute the panelists' housing unit square footage of living area.

The measure is imputed from the following three data sources: the Stanford Market Basket Database, the US Census 1990, and the American Housing Survey (AHS) 1989 (see Appendix 3 for a detailed description of the estimating procedure).[12]

In the interests of parsimony and ease of exposition, in the remainder of the paper we report regressions for one focal measure: Average square footage of living area in the panelist's zip code (subsequently denoted as 'housing unit size'). That is, we assume that all households in a zip code share the same value for this estimate. We also impute the square footage of living area for individual panelists based on their demographic information from the Stanford Market Basket Database (in addition to their location information). The method of imputation is described in detail in Appendix 3.[13] While the underlying hedonic regressions have a somewhat worse fit, it is worth noting that results reported in Tables 3 to 8 are little changed if we use the household specific housing unit size measure instead of the zip code specific measure. Finally, we also estimated models with an estimate of "storage costs" (instead of "storage constraints") calculated as the imputed housing unit value per square foot, both at the zip code and household specific level (again see Appendix 3 for details).[14] This formulation produces qualitatively similar results as well. Overall our findings indicate that our results are robust to alternative measures of storage constraints or costs.

The argument in favor of using the housing unit size as preferred proxy measure relies on the premise that households who consider stockpiling only take into account the size of their housing unit but not the cost per square foot. That is, households in small housing units cannot stockpile because of space limitations, while households in large units will always find some space to stockpile consumables.[15]

It is worth noting that our empirical results are 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 greater storage constraints. Our estimating equation for the purchase frequency of consumer *i* is as follows:

$$(1) \quad \#\,trips_i = \beta_0 + \beta_1 \; housing\ unit\ size_i + \beta_2 \, weighted\ distance\ to\ store_i + \beta_3 \; total\ spending_i + \beta_4 \; demographics_i + \beta_5 \; store\ dummies\ (closest\ store)_i + \varepsilon.$$

We include, in addition to the housing unit size as our (inverse) measure for storage constraints, the distance to stores (weighted by the number of trips to each store), the consumer's total spending (in US dollar) in all stores, and a number of demographic characteristics of the consumers (see Table 1 for a description of these variables). The equation controls for unobservable characteristics related to the consumer's closest store. The dummy variable for a specific store is zero unless it is consumer *i*'s closest store.

Table 3 summarizes the regression results. As noted above, the storage constraint measure is obtained from housing unit size estimates (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.

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

Our coefficient of interest, β_1, has a negative sign and is highly statistically significant. Households with larger housing units shop less frequently. The effect is not only significant in a statistical but also in an economic sense. Quantitative effects are reported in Table 8.[16] The effects are computed under the scenario where a typical household moves from the urban fringe (with an average of 1674 square feet of living space) to the downtown neighborhood (with an average of 1146 square feet of living space). As the first two rows of Table 8 reveal, an urban fringe household moving to the downtown neighborhood would on average increase the percentage of shopping trips taken by approximately 15%, resulting in an average of 22 additional shopping trips over a two year period.

The other coefficients have plausible signs. For example, higher income households are visiting stores less often, likely because of higher opportunity costs of time (which in turn lead to higher shopping trip and search costs). Household size shows the same effect, again likely due to the opportunity costs of time. 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 interest is the weighted distance to stores. The estimated coefficient is negative, implying the further households must travel, the less frequently they shop. The further the travel distance to the store, the greater the fixed cost of shopping. Rational cost-minimizing shoppers should therefore buy larger baskets per visit to amortize this fixed cost, and therefore all else equal, shop less frequently (Bell, Ho and Tang 1998).

Overall, the results in Table 3 offer strong support for the assumption that greater storage constraints (i.e., a smaller square footage of living area) will lead consumers to take more shopping trips.

(b) Purchase Quantity per Trip. Our second testable assumption is that a consumer *i* will purchase smaller quantities per trip if he or she faces restrictive storage constraints. Here we report separate regression results for all five product categories described in Table 2: Paper towels, bathroom tissue, liquid detergents, ground coffee and pills in capsules. Our estimating equation for each category parallels equation (1) as follows:

$$\begin{aligned} (2) \qquad \textit{purchase quantity}_i = {} & \beta_0 + \beta_1 \textit{ housing unit size}_i + \beta_2 \textit{ av. unit price of product category}_i + \\ & \beta_3 \textit{ weighted distance to store}_i + \beta_4 \textit{ total spending}_i + \\ & \beta_5 \textit{ demographics}_i + \beta_6 \textit{ store dummies (closest store)}_i + \varepsilon. \end{aligned}$$

The results given previously in Table 3 show that storage constraints have the expected macro-level effect on consumer shopping behavior. The coefficient β_1 is negative and significant, indicating that more constrained households (with a small square footage of living space) 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 constraints. Storage requirements for paper towels, bathroom tissue, and liquid detergents are relatively high, whereas ground coffee and, particularly, pills in capsules consume very little space. Even households with 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 six rolls of paper towels or 32oz of laundry detergent.

In reporting the results for equation (2) in Table 4 we separate the findings into "high storage use items" (paper towels, bathroom tissue, and liquid detergents) and "low storage use items" (ground coffee and pills in capsules). Again, our coefficient of interest is β_1 and the prediction is that – at least for high storage use items – the coefficient will be positive: Less constrained consumers with larger housing units purchase larger quantities on average, per shopping trip. As before, our equations contain a number of control variables (including the average unit price faced by the panelist).

Table 4 reveals that storage constraints have statistically significant effects on purchase quantities in all high-storage requirement categories (paper towels, bathroom tissue, and liquid detergent). For the products that use much less storage space (coffee and pills in capsules) there is no statistically significant effect as expected.

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

Thus, the findings pertaining to the key model assumption – storage constraints influence consumer behavior – are unequivocal. Greater storage constraints 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 proxy is further enhanced by the fact that quantitative effects of our storage constraint proxy on purchase quantities is quite meaningful for high storage use items (paper towels, bathroom tissue, and liquid detergent) but rather insignificant economically for low storage use items (ground coffee and pills in capsules). Table 8 reports the percentage change effect of a customer's move from the urban fringe to the downtown neighborhood on average purchase quantities for all five product categories. 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. The increase in the storage constraint due to the move to the downtown leads the household to purchase significantly smaller volumes on average, for each shopping trip, *for product categories that have non-trivial storage requirements*. In the case of paper towels, we see that a household moving from a low to a high storage constraint environment (i.e., from the urban fringe to a downtown neighborhood) would reduce the average purchase quantity by about 16%. Negative and quantitatively meaningful percentage changes in average purchase quantities are also observed for bathroom tissues and liquid detergents, although, the effects are somewhat smaller (-14% for bathroom tissues and -8% for liquid detergents). The changes for ground coffee and pills in capsules are not statistically significant and the calculated quantitative effects are much smaller (and have the opposite sign in the case of pills in capsules).

As indicated at the bottom of Table 4, as an alternative to the reported quantity regression specifications, Heckman selection equations were estimated to account for sample selection effects and differences in buyers and non-buyers of each product category. Selectivity is rejected for paper towels, bathroom tissue and liquid detergents, but not for coffee and pills. In all cases, the qualitative results of the selection model are identical to those reported in Table 4. Finally, it is worth noting that results are similar when we use the housing unit value per square foot as a proxy for storage costs rather than housing unit size as a proxy for storage constraints.

Having established the validity of the consumer behavior assumption and our ability to measure storage constraints, we now turn our attention to the predictions of Salop and Stiglitz' (1982) 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 inter-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). We therefore compute SKU summary measures over time, within store and SKU. The measures are average price, empirical promotion frequency, average promotional depth, and price dispersion (measured as the coefficient of variation). Thus, for a given SKU (say Tide Detergent, 32oz) sold in a particular store for 104 weeks, we compute the average price, empirical proportion of times the product is promoted, and the average promotional depth, given that a promotion is offered.

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

$$
\begin{aligned}
\text{(3)} \qquad \textit{offer price}_{p,s} = {} & \beta_0 + \beta_1 \textit{ av. housing unit size of customers faced by store and SKU}_{p,s} + \\
& \beta_2 \textit{ demographics of customers}_{p,s} + \beta_3 \textit{ weeks on shelf}_{p,s} \\
& + \beta_4 \textit{ SKU fixed effects}_p + \varepsilon.
\end{aligned}
$$

The theoretical model presented in Section 1 predicts that the coefficient β_1 will be negative (holding everything else constant) as the comparative static result shows that $\overline{p}$ is increasing in h (or decreasing in the square footage of living area, as this measure is inversely related to storage cost h). In our empirical analysis, we would therefore expect that stores set higher average prices if they face more storage constrained households that live in smaller housing units. The set of control variables mirrors that used in the consumer behavior regressions. In particular, we utilize proxies for consumer search costs to control for the possibility that stores are also cognizant of consumer mobility and their ability to take advantage of lower prices. This is especially important as search costs and storage constraints have theoretically confounding effects on firm behavior. Stores that face less storage constrained consumers will need to set lower prices; stores that face consumers with lower opportunity costs of time and greater mobility will also need to set lower prices.

We characterize the store's customer base for specific SKUs according to four search and opportunity cost proxies: their income, age, household size, and employment status. Our independent measures are store *and* SKU specific to account for the fact that stores may attract a different mix of customers for different products. For example, the customer mix that shops for Tide in a particular store may be different from the customer mix that purchases pills in the same store. Bucklin and Lattin (1992) document product specific effects for different stores; Montgomery (1997) shows how stores practice mirco-marketing by tailoring prices for specific products to the customer mix for that product. Similar to the regressions on the consumer side, we calculate the percentage of Blacks and Hispanics in the customer base faced by each store. 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.

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—with relatively higher opportunity costs of time—should be less sensitive to prices compared to lower income customers. The same may be true for elderly consumers, as searching for better deals may be more cumbersome for them.

As shown in Table 5, Prediction 1 of the theory has strong support: Average prices are lower in stores that face a distribution of less storage constrained customers who live in larger housing units. This substantive finding is robust to different specifications as outlined in the endnote of Table 5. Moreover, the effect is not only highly statistically significant but also reasonably meaningful in economic terms. As Table 8 reveals, a store that faces storage constrained downtown customers sets prices *for the same item* on average 1.7% higher than a store that faces less constrained customers from the urban fringe, all else equal. In dollar terms this is equivalent to an increase in the price of an average volume of bathroom tissue from \$1.64 to approximately \$1.67.

The control variables are significant and have expected and plausible signs. For example, prices are higher if stores face households that have higher income or are older. However, prices are lower if stores face a large share of households with unemployed members or large families.[17] All these results are consistent with search and opportunity cost theory.[18] As expected, products that are available more often also have lower average prices, everything else constant. We also include SKU-level fixed effects to control for all variation in prices that are due to idiosyncratic differences in the product characteristics themselves.

[----- Table 5 About Here -----]

(b) Comparative Static Prediction 2: Probability of Promotion. The empirical estimation strategy is consistent with that employed for the average price prediction, however in this case we must account for selection effects. Specifically, the probability of promotion and promotion depth equations are linked; promotional depth exceeding zero is only observed for SKU-store combinations where promotions take place. The selection equation in the selectivity model is as follows:

$$\begin{aligned} (4) \quad \Pr(\mathit{SKU\ is\ promoted} = \mathit{yes})_{p,s} = \ & \beta_0 + \beta_1\ \mathit{average\ housing\ unit\ size\ of\ customers}_{p,s}\ + \\ & \beta_2\ \mathit{demographics\ of\ customers}_{p,s}\ + \\ & \beta_3\ \mathit{weeks\ on\ shelf}_{p,s}\ + \varepsilon. \end{aligned}$$

The theoretical model presented in Section 1 predicts that the coefficient β_1 will be positive (holding all else constant) as the comparative static result shows that $(1-\lambda)$ is decreasing in h. The probability that promotions are offered goes down when households have greater storage constraints, or, in our empirical analysis the likelihood that a product is promoted increases when stores face households that live in larger (less storage constrained) housing units. The dependent variable in equation (4) is a dummy variable that equals 1 if promotion for SKU p in store s takes place and 0 otherwise.

Again, we find strong support for the Salop and Stiglitz (1982) model: The bottom section of Table 6 reports estimates for the selection equation and shows that when stores face less storage constrained customers (i.e. customers who consume more square footage of living area), they are more likely to promote. While these results also hold if the promotion probability and promotion depth equations are estimated separately, there is strong justification for the Heckman selectivity estimator as indicated by test statistics (see notes at bottom of Table 6).[19] The effect of storage constraints on the probability of promotion is not only statistically highly significant but also quantitatively meaningful. As Table 8 shows, a store that faces storage constrained downtown customers is about 9% less likely to promote any given product than a store that faces customers from the urban fringe, all else equal.

Where statistically significant, control variables have plausible signs. In particular, stores are more likely to promote a particular product if many of their customers are members of large families and if the product itself is on the shelf for a longer period of time.

[----- Table 6 About Here -----]

(c) Comparative Static Prediction 3: Promotional Depth. Promotional depth is defined as the relative size of the price cut, expressed as a percentage of the regular price, averaged over all discounted occasions when the product is available on the shelf. The estimating equation for the promotional depth of product p in store s, conditional on the probability of promotion is as follows:

$$
(5) \qquad \begin{aligned}
\textit{promotional depth}_{p,s} = \ & \beta_0 + \beta_1 \textit{ average housing unit size of customers}_{p,s} + \\
& \beta_2 \textit{ proportion of times SKU is displayed}_{p,s} + \\
& \beta_3 \textit{ demographics of customers}_{p,s} + \beta_4 \textit{ weeks on shelf}_{p,s} + \varepsilon.
\end{aligned}
$$

Note that the proportion of times a product is displayed in a particular store is included as an additional control variable in equation (5). Offering deeper promotions may be a more sensible strategy if the product can also be displayed.

The comparative static prediction from theory is that, all else equal, the relative size of the discount $\left(p_h - p_l\right)/p_h$ is decreasing in the storage cost variable h. In empirical terms, promotions should be more pronounced when stores face customers who live in larger (i.e., less storage constrained) housing units.

Results given in the top section of Table 6 are again consistent with theory. Stores that face customers who are less storage constrained must offer deeper discounts. The effect is both statistically highly significant and quantitatively meaningful. As Table 8 reveals, differences in customers' storage constraints lead stores to offer almost 7% deeper promotions when they face less constrained consumers from the urban fringe compared to the more constrained downtown customers. Among the control variables, it is worth noting that the additional control for display frequency has the predicted positive impact on promotional depth.[20]

Thus, taken together the coefficients for housing unit size in the bottom and top portion of Table 6 show that stores facing less constrained households promote more often, and conditional on offering promotions give proportionally greater price cuts, respectively.

(d) Discussion of Possible Tests of the Non-Monotonic Equilibrium Relationship between Storage Constraints and Price Dispersion. An important equilibrium result arising from the theory is that price dispersion, as measured by the coefficient of variation $\sigma_p^2 / \overline{p}^2$ first increases then decreases in relative storage costs (see Salop and Stiglitz 1982, p. 1125 and Figure 1). Moreover, for very high magnitudes of storage costs $h \geq u/3$, theory suggests that there exists a single price equilibrium.

On the empirical side, testing the non-monotonic relationship between storage constraints and price dispersion is evidently difficult. At first, while it is reasonable to assume that our measure of storage constraints (the housing unit size) and the true storage costs of households are inversely and monotonically related, the precise relationship is not known. Moreover, theory only makes a qualitative – but not a quantitative – statement about the threshold of storage costs, at which, price dispersion is no longer the equilibrium solution. Given these issues, it is difficult to accurately test the equilibrium prediction with respect to price dispersion.

In an attempt to at least assess whether our data are consistent with the equilibrium prediction, we analyzed them as follows. To begin with, price dispersion is almost four times smaller (0.0015 compared to 0.0053) for stores that face severely storage constrained customers (with a living area below 1000 square feet) compared to stores that face customers who likely are not storage constrained at all (panelists with a living area greater than 2000 square feet). This empirical finding is broadly consistent with the assertion that very high storage costs lead to a single price equilibrium. Moreover, our finding is quite robust with respect to other thresholds of 'severely constrained' and 'little constrained'. Next, we limit our sample to product-store-combinations that face less constrained customers (with a living area greater than 1300 square feet). Theory predicts that we should find an inverted U-shaped relationship between storage constraints and price dispersion for such a sample of 'less constrained' customers. More precisely, we estimate a Heckman selection model of price dispersion (similar to the Heckman selection model for the probability of promotion and promotional depth described above) in order to account for the fact that (selected) observations with observed price dispersion may have different characteristics from (non-selected) observations with no price dispersion. We include the average housing unit size and the average housing unit size squared as explanatory variables in order to test for the predicted non-linear relationship between price dispersion and storage constraints. Results are reported in Table 7.

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

As Table 7 illustrates the coefficient on the linear term is positive and the coefficient on the quadratic term is negative. Hence, price dispersion indeed first increases then decreases in housing unit size, as predicted by theory.[21] The estimated relationship is statistically significant and test statistics suggest that the selection model is justified. However, it should be noted that our results are somewhat sensitive to the choice of the threshold value for 'less constrained customers' (see Table 7 for more details). Nevertheless, overall, our results are broadly consistent with the theoretical equilibrium prediction suggested by Salop and Stiglitz' (1982) Theory of Sales.

3 Summary and Conclusion

While Sorensen (2000) and Lach (2002) provide empirical tests of the more visible 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 constraint explanation posited by Salop and Stiglitz (1982) in their Theory of Sales. The key assumption and predictions of the theory are tested using and combining data from two remotely adjacent fields; marketing (market basket data) and urban and real estate economics (housing constraint and cost data). Our empirical analysis provides evidence for both the validity of the assumptions and the predictions of the Theory of Sales.

A critical underlying assumption of the 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 constraints causes consumers to purchase goods more often and to buy smaller quantities on a store visit. 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 constraints have no appreciable effect on consumer purchase behavior.

Three empirically testable comparative static predictions for store behavior can be derived from Salop and Stiglitz' (1982) equilibrium framework: An increase in consumer storage constraints will lead to: (1) an increase in the average price, (2) a reduction in the probability of promotion, and (3) a decrease in promotional depth. Each result is supported by the data. We also discuss the framework's equilibrium prediction with respect to price dispersion. The predicted relationship between storage constraints and price dispersion is non-monotonic. Storage constraints determine whether there exists a price dispersion or a single price equilibrium. While the functional form is difficult to estimate empirically, our empirical findings are consistent with the equilibrium model prediction; price dispersion first increases then decreases in housing unit size.

Furthermore, we show that the reported effects are not only statistically significant but also quantitatively (i.e., economically) important. Quantitative effects for key measures of consumer and store behavior are summarized in Table 8 and reported along with statistical significance levels throughout the empirical section of the paper.

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

The effects of consumer storage constraints on consumer behavior are quantitatively meaningful. Table 8 shows, for example, an urban fringe household moving to a downtown neighborhood would on average increase the percentage of shopping trips taken by approximately 15% (an extra 22 trips over a two year period). Similarly, Table 8 reveals that consumer storage constraints affect the store behavior in an economically significant way. For example, a store that faces storage constrained downtown customers sets prices on average 1.7% higher than an identical store that faces less constrained customers from the urban fringe. Corresponding negative effects are observed for the probability of promotion and the conditional depth of promotion. That is, stores that face more storage constrained consumers not only promote less often, but also offer shallower discounts when they do.

Moreover, Table 8 validates our empirical measure of consumer storage constraints by demonstrating that the effects are only quantitatively (and statistically) significant for product categories that have storage requirements (i.e., bathroom tissue, paper towels, and liquid detergents) but not for products that essentially require very limited or no storage space (i.e., coffee and pills). A nice feature of the results is that the rank ordering of quantitative magnitudes is highly correlated with the volume of space required by the product category. That is, a household moving from a relatively unconstrained to a storage-constrained housing unit shows the largest change in average purchase volumes for paper towels followed by bathroom tissue and liquid detergents (ground coffee and pills show no significant effect).

In conclusion, the estimated effects have the theoretically correct signs, are statistically significant, are robust to a number of alternative storage constraint (or storage cost) proxies, and are economically significant and of a plausible magnitude. The analysis also controls for potentially confounding effects due to opportunity costs of time. Empirical support for the Theory of Sales is very encouraging and points to its validity as a characterization of consumer and firm behavior in real markets.

Notes

[1] Salop and Stiglitz (1982) use the term "storage costs". We use this term interchangeably with "storage constraints" which better describes our empirical measure, housing unit size. Note that housing unit size is an inverse measure of storage constraints. That is, households with a larger square footage of living area are less storage constrained. Clearly the two terms "storage costs" and "storage constraints" are complementary. We also impute alternative measures for "storage constraints" and "storage costs" (see Appendix 3). Sensitivity tests reveal that the results are quite robust (in a qualitative and statistical sense) with respect to the choice of alternative proxy measures. We thank an anonymous reviewer for suggesting we focus on housing unit size as the appropriate measure.

[2] While storage constraints are primarily a *consumer* specific phenomenon related to the size 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.

[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 constraints on consumer or firm behavior.

[4] 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, Mazumbdar, 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 constraints in affecting purchase quantities.

[5] Actual storage constraints (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 Hendel and Nevo (2003).

[6] 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).

[7] Salop and Stiglitz initially 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.

[8] 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.

[9] Obviously, such a complex relationship between h and price dispersion is very difficult to estimate empirically. We elaborate on this issue in Section 2 below.

[10] 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).

[11] Glaeser, Gyourko, and Hilber (2001) and Glaeser and Gyourko (2001 and 2003) 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 housing unit size and house value per square foot, our storage constraint and storage cost proxy measures of interest. The most disaggregated location information available is the MSA level.

[13] 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 2003).

[14] Summary statistics of these alternative two storage cost measures are provided in Table 1.

[15] We are grateful to an anonymous reviewer for this conceptualization.

[16] The calculations are based on the estimated coefficients from the models presented in Tables 3, 4, 5 and 6.

[17] Unemployed customers have lower opportunity cost of time and consequently lower search costs. Members of large families likely have greater opportunity costs of time but at the same time they buy greater quantities per trip resulting in lower search cost per purchased unit. That is, members of large families may make fewer shopping trips but they may, during any given shopping trip, compare prices more intensely between nearby stores.

[18] We also estimated the empirical model including the average distance of customers to the store (by store and SKU) as an additional proxy measure for search cost. One might expect that greater travel distances and longer shopping trips result in higher search costs and consequently in a higher price. However, the shopping trip distance is negatively related to average prices. This is likely due to the fact that consumers care about the total delivered price of the shopping basket. Hence, consumers are only willing to undertake long shopping trips if they are compensated with lower prices. That is, the causality is reversed; lower prices lead to longer shopping trips and not vice versa. We decided to drop the variable from the final specification due to this potential endogeneity issue. It is worth noting, however, that the results reported in Tables 5 and 6 are virtually unchanged if the distance-to-store measure is included as an additional control variable.

[19] We are very grateful to an anonymous reviewer for the suggestion that selectivity be accounted for in the model.

[20] The results remain qualitatively unchanged and our measure for storage constraints remains statistically significant with the predicted sign if the display frequency is dropped from the estimating equation.

[21] Figure 1 in Salop and Stiglitz (1982, p. 1125) or straightforward differentiation shows that price dispersion $\sigma_p^2 / \overline{p}^2$ is first increasing then decreasing in $v = u / h$. An increase in storage cost h is equivalent to a move from the right to the left on the x-axis (variable v). Hence, dispersion is first increasing then decreasing in h. Assuming that storage costs and storage constraints are positively and monotonically related, it follows that an increase in storage constraints also means an increase in storage costs (move from the right to the left on the x-axis of Figure 1 in Salop and Stiglitz). That is, dispersion must first increase then decrease in storage *constraints.* Finally, housing unit size is our *inverse* measure of storage constraints (assuming a negative and monotonic relationship). An increase in storage constraints (move from the right to the left on the x-axis) is equivalent to a decrease in housing unit size. Hence, an increase in the housing unit size (move from the left to the right on the x-axis) is qualitatively equivalent to an increase in v in Figure 1. It follows that dispersion first increases then decreases in the housing unit size. In empirical terms we would therefore expect a positive sign on the housing unit size-variable and a negative sign on the housing unit size squared-variable. The fact that dispersion first increases and then decreases in all three measures (storage cost h, storage constraints, *and* housing unit size) may not be immediately intuitive. The result is due to the inverted U-shape of the dispersion function. Moving from the left to the right on the x-axis or, vice versa, from the right to the left, both implies first an increase then a decrease in dispersion.

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.

Assuncao, Joao and J. Robert Meyer. (1993). "The Rational Effect of Price Promotions on Sales and Consumption", Management Science 5, 517–535.

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.

Bucklin, Randolph E. and James M. Lattin. (1992), "A Model of Product Category Competition Among Grocery Retailers", Journal of Retailing 68, 271-293.

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.

Eppen, Gary D., and Yehoshua Lieberman (1984). "Why Do Retailers Deal? An Inventory Explanation", Journal of Business 57, 519-530.

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 (forthcoming Journal of Political Economy).

Glaeser, Edward L. and Joseph Gyourko. (2003). "The Impact of Zoning on Housing Affordability", Economic Policy Review 9, 21-30, Federal Reserve Bank of New York.

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.

Jeuland, Abel P, and Chakravarthi Narasimhan. (1985). "Dealing – Temporary Price Cuts – by Seller as a Buyer Discrimination Mechanism", Journal of Business 58, 295-308.

Kahneman, Daniel and Amos Tversky. (1979). "Prospect Theory: An Analysis of Decision under Risk", Econometrica 47, 263-291.

Krishna, Aradhna. (1992). "The Normative Impact of Consumer Price Expectations for Multiple Purchase Brands on Consumer Purchase Behavior", Marketing Science 11, 266-286.

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.

Meyer, Robert J. and Joao Assuncao. (1990). "The Optimality of Consumer Stockpiling Strategies", Marketing Science 9, 18-41.

Montgomery, Alan L. (1997). "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data", Marketing Science 16, 315-337.

Narasimhan, Chakravarthi. (1984). "A Price Discrimination Theory of Coupons", Marketing Science 3, 128-147.

Narasimhan, Chakravarthi. (1988). "Competitive Promotional Strategies", The Journal of Business 61, 427-449.

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., Srinivasan, V. and Rajiv Lal. (1990). "The Effects of Brand Loyalty on Competitive Price Promotional Strategies", Management Science 36, 427-449.

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 unit size in square foot, based on zip code characteristics (Storage *constraint* proxy no. 1)	1399.2	390.0	125.7	2454.2
Estimated housing unit size in SF, based on individual consumer and zip code characteristics (Proxy no. 2), *N=991*	1414.8	399.6	159.7	2643.4
Estimated housing costs per square foot (in $), based on zip code characteristics (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), *N=991*	194.2	227.1	21.8	1058.4
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 United States 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 Detergent	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 (Unconstrained) Consumers with Large Housing Units Buy Goods Less Often? —Purchase Frequency Regression Results

Dependent Variable: Total Number of Trips to All Stores

Explanatory Variable	Coefficient (Robust Standard Error)	
Housing unit size of consumer (zip code average), 1990	**-.042**	***
	(.014)	
Weighted distance to stores (in miles, weight=trips to each store)	-1.1	**
	(.53)	
Total spending in all stores	.0078	***
	(.0016)	
Household income	-.00059	***
	(.00017)	
Household size	-9.1	***
	(2.1)	
Average household age > 65	9.3	
	(7.1)	
At least one member of household is unemployed (dummy equals 1 if true)	4.7	
	(6.6)	
Race of household is Black	21.4	***
	(8.2)	
Race of household is Hispanic	-15.6	*
	(9.5)	
Store dummies (equal 1 if consumer residence is closest to store)	Yes	
Constant	242.1	***
	(37.1)	
Number of observations (consumers)	996	
Adjusted R^2	.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 qualitatively similar if alternative proxy measures for storage constraints and storage costs are used (i.e., imputed individual housing unit size, zip code level house value per square foot, imputed individual house value per square foot).

Table 4

Do (Unconstrained) Consumers with Low Storage Constraints Buy Larger Quantities?

Dependent Variable: Average purchase quantity per customer and trip

	High Storage Use Items			Low Storage Use Items	
Explanatory Variable	Paper Towels	Bathroom Tissue	Liquid Detergent	Ground Coffee	Pills in Capsule Form
Average housing unit size of consumers (based on zip code level information)	**.00050** *	**.0014** ***	**.00085** **	**.0029**	**-.0059**
	(.00027)	**(.00056)**	**(.0044)**	**(.0020)**	**(.0086)**
Weighted distance to stores (in miles, weight= trips to each store)	-.0074	.020	-.0026	-.083	-.41
	(.0076)	(.024)	(.015)	(.076)	(.51)
Unit price (average)	-1.1 ***	-10.4 ***	-2.9 ***	-72.0 ***	-814.8 ***
	(.20)	(.63)	(.15)	(3.2)	(69.9)
Total spending in all stores (x 10^{-3})	.014 **	.029 **	.026	.016	-.40
	(.0060)	(.014)	(.017)	(.063)	(.30)
Household income (x 10^{-3})	.0028 **	.0082 *	.0074 **	-.0072	.14
	(.0017)	(.0046)	(.0037)	(.017)	(.096)
Household size	.020	.15 ***	.18 ***	.75 **	-3.6 ***
	(.026)	(.058)	(.069)	(.33)	(1.0)
Average household age >65	-.0014	-.033	-.076	.78	.80
	(.083)	(.19)	(.18)	(1.0)	(3.5)
At least one member of household is unemployed (dummy equals 1 if true)	-.080	.093	.076	-1.1	-.48
	(.060)	(.18)	(.19)	(.78)	(3.6)
Race of household is Black	-.085	-.016	-.096	-.44	-1.7
	(.077)	(.28)	(.20)	(1.7)	(5.5)
Race of household is Hispanic	-.18	.28	-.011	-2.5	4.8
	(.15)	(.49)	(.48)	(1.9)	(9.7)
Store dummies (for closest store)	Yes	Yes	Yes	Yes	Yes
Constant	1.3 **	6.7 ***	5.5 ***	33.2 ***	160.6
	(.63)	(.73)	(1.1)	(5.1)	(26.7)
Number of observations (consumers)	954	981	861	760	573
Adjusted R^2	.082	.24	.38	.42	.46

Notes: Numbers in parentheses are robust standard errors. ***/**/* Significantly different from zero with 99/95/90 percent confidence. Regressions with alternative storage cost proxies give qualitatively similar results. Heckman selection models using Maximum Likelihood were estimated in order to determine whether standard OLS yields biased results. The *z*-values for ρ and the χ^2 statistic are completely stat. insignificant for the selection models for paper towels, bathroom tissue, and liquid detergent indicating that no selection bias is present. The hypothesis that $\rho \neq 0$ cannot be rejected for coffee and pills. However, results with respect to our storage constraint measure remain qualitatively completely unchanged when Heckman selection models are applied (with the *z*-values of the housing size variable in the regression equations being 1.0 for coffee and 0.41 for pills.). Results are available upon request.

Table 5

Do Storage Constraints of Consumers Affect Average Prices Charged?

Dependent Variable: Average purchase price of product by store

	Average Purchase Price	
	(1)	
Average (zip code specific) housing unit size of consumers faced by store, by product (x 10^{-3})	-.13 (.018)	***
Average household income faced by store, by product (x 10^{-3})	.0026 (.00052)	***
Percentage of households with average age above 65 faced by store, by product	.12 (.25)	***
Average household size faced by store, by product	-.077 (.0084)	***
Percentage of households with at least one unemployed member faced by store, by product	-.080 (.024)	***
Percentage of Black households faced by store, by product	.067 (.33)	**
Percentage of Hispanic households faced by store, by product	-.45 (.061)	***
Number of weeks product is on shelf (by store and product)	-.0022 (.00028)	***
SKU fixed effects	Yes	
Constant	4.3 (.047)	***
Number of observations (SKUs x stores)	4760	
Adjusted R^2	.98	

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 constraint measure are little changed if an alternative proxy based on the household specific information from the AHS is used. Moreover, results are qualitatively very similar if house value per square foot is used as alternative proxy. Finally, results are qualitatively very similar if the average distance of consumers to the store (by store and SKU) is included as a proxy measure for search cost. The proxy was dropped from the final specification because it is arguably endogenous.

Table 6

Do Storage Constraints of Consumers Affect Promotion Strategies of Stores?

Dependent Variables: (1) Promotional depth, (2) Promotion (selection) dummy
Estimator: Heckman Selection Model Using Maximum Likelihood

	Promotional Depth		
Regression Equation	Coefficient		Robust Standard Error
Average (zip code specific) housing unit size of consumers faced by store, by product (x 10^{-3})	.022	***	.0091
Proportion of times product is in in-store display (by store and product)	.36	***	.027
Average household income faced by store, by product (x 10^{-3})	.00020		.00022
Percentage of households with average age above 65 faced by store, by product	.030	***	.012
Average household size faced by store, by product	.000099		.0039
Percentage of households with at least one unemployed member faced by store, by product	-.035	***	.010
Percentage of Black households faced by store, by product	-.0028		.013
Percentage of Hispanic households faced by store, by product	-.056	**	.024
Number of weeks product is on shelf (by store and product)	.0011	***	.000083
Constant	.14	***	.020
Selection Equation	Promotion (Selection) = yes		
Average (zip code specific) housing unit size of consumers faced by store, by product (x 10^{-3})	.56	***	.069
Average household income faced by store, by product (x 10^{-3})	.00016		.00018
Percentage of households with average age above 65 faced by store, by product	.052		.090
Average household size faced by store, by product	.072	**	.031
Percentage of households with at least one unemployed member faced by store, by product	-.026		.093
Percentage of Black households faced by store, by product	.46	***	.13
Percentage of Hispanic households faced by store, by product	-.62	**	.19
Number of weeks product is on shelf (by store and product)	.0011	***	.00012
Constant	-1.1	***	.14
Inverse hyperbolic tangent of ρ	-.28	***	.048
Ln σ	-1.8	***	.013
ρ	-.28		.044
σ	.17		.0021
$\lambda = \rho\sigma$	-.046		.0075

N=4760 (Uncensored N=3669). Log pseudolikelihood = -640.2; Wald $\chi^2(9) = 401.0$; $\Pr > \chi^2 = 0.0000$.

Wald test of independent equations (ρ = 0): $\chi^2(1) = 6.34$, $\Pr > \chi^2 = 0.01$.

Notes: *** / ** / * Significantly different from zero with 99 percent / 95 percent / 90 percent confidence. The z =-6.0 of the inverse hyperbolic of ρ and the χ^2 of 6.34, both significantly different from zero, justify the Heckman selection equation. Results are qualitatively similar if alternative storage cost proxies are used or if the average distance of consumers to the store (by store and SKU) is included as a proxy measure for search cost. The proxy was dropped from the final specification because it is arguably endogenous.

Table 7

Do Storage Constraints of Consumers Affect Price Dispersion?

Dependent Variables: (1) Standard deviation of price divided by average price, (2) Existence of price dispersion (selection) dummy

Estimator: Heckman Selection Model Using Maximum Likelihood

	Price Dispersion		
Regression Equation	Coefficient		Robust Standard Error
Average (zip code specific) housing unit size of consumers faced by store, by product (x 10^{-3})	.18	***	.058
Average (zip code specific) housing unit size of consumers faced by store squared, by product (x 10^{-6})	-.035	**	.017
Average household income faced by store, by product (x 10^{-3})	-.00013	**	.000067
Percentage of households with average age above 65 faced by store, by product	.0022		.0026
Average household size faced by store, by product	.00073		.00082
Percentage of households with at least one unemployed member faced by store, by product	-.0040	**	.0016
Percentage of Black households faced by store, by product	-.00041		.0022
Percentage of Hispanic households faced by store, by product	-.0012		.0025
Number of weeks product is on shelf (by store and product)	.00062	***	.000036
Constant	-.17	***	.051
Selection Equation	Price Dispersion (Selection) = yes		
Average (zip code specific) housing unit size of consumers faced by store, by product (x 10^{-3})	4.1	***	.96
Average (zip code specific) housing unit size of consumers faced by store squared, by product (x 10^{-6})	-1.0	***	.28
Number of weeks product is on shelf (by store and product)	.010	***	.00068
Constant	-3.7	***	.83
Inverse hyperbolic tangent of ρ	4.1	***	.17
Ln σ	-2.7	***	.029
ρ	1.0		.00019
σ	.066		.0019
$\lambda = \rho\sigma$	.066		.0019

N=2969 (Uncensored N=2345). Log pseudolikelihood = -2484.0;
Wald test of independent equations (ρ = 0): $\chi^2(1) = 591.5$, $\Pr > \chi^2 = 0.0000$.

Notes: *** / ** / * Significantly different from zero with 99 percent / 95 percent / 90 percent confidence. The z =24.3 of the inverse hyperbolic of ρ and the χ^2 of 591.5, both significantly different from zero, justify the Heckman selection equation. The sample is restricted to product-store-combinations that face 'less constrained' customers. Customers are considered to be storage constrained if their living area is smaller than 1300 square feet (roughly the 40th percentile). The standard errors of our coefficients of interest are somewhat sensitive to the choice of the threshold value for 'less constrained' customers. However, within a reasonable range of threshold values (+/- 100 square feet), the coefficients are typically significantly different from zero with 90 percent confidence. In qualitative terms (signs of coefficients) results are robust for a wide range of threshold values.

Table 8

Quantitative Effects: Effect of Consumer Relocation from Urban Fringe to Downtown

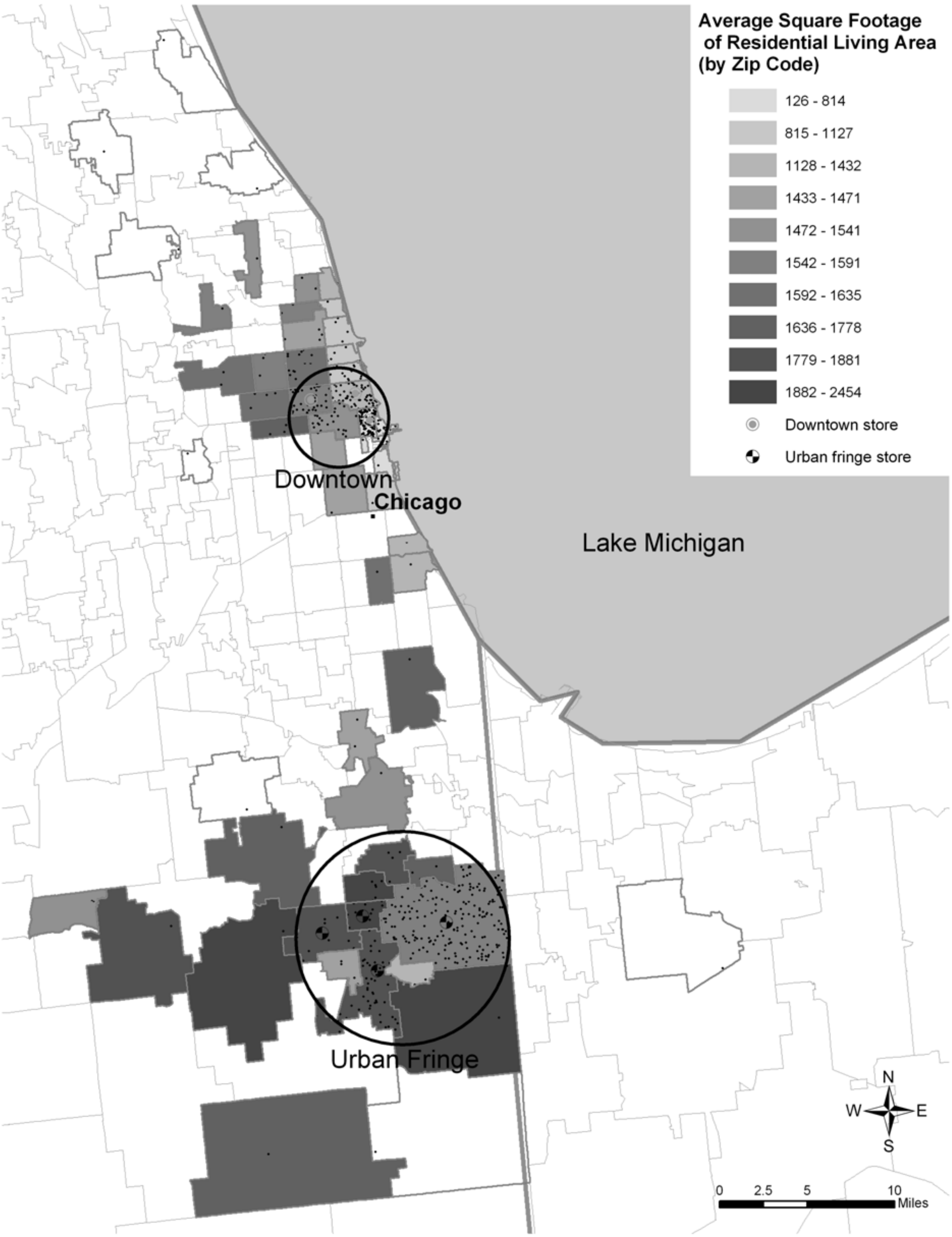
Assumptions: Estimated housing unit size decreases from 1674 to 1146 square feet

Change in Dependent Variable	*Effect of Relocation from Urban Fringe to Downtown Due to Greater Consumer Storage Constraints*	
% Change in purchase frequency *(based on regression reported in Table 3)*	+15.0%	***
Change in no. of trips to store over a two year period *(based on Table 3)*	+22.2 trips	***
% Change in average purchase quantity *(based on Table 4)*		
• Paper Towels *(Table 4, Column 1)*	-15.6%	*
• Bathroom Tissue *(Table 4, Column 2)*	-13.9%	***
• Liquid Detergent *(Table 4, Column 3)*	-7.9%	**
• Ground Coffee *(Table 4, Column 4)*	Not stat. sign.	[-4.8%]
• Pills in Capsules *(Table 4, Column 5)*	Not stat. sign.	[+4.4%]
% Change in offered product price *(based on Table 5, Column 1)*	+1.7%	***
% Change in promotional depth *(based on Table 6—regression equation)*	-6.6%	***
Change in % points	-1.7% points	***
% Change in probability that product is promoted at least once (i.e., selected) *(based on Table 6—selection equation)*	-9.2%	***
Change in % points	-7.6% points	***

Notes: 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 a product is ever promoted within a particular urban fringe store is 81.7 percent, and the typical depth of promoted products is 26.0 percent. Quantitative effects in square brackets are not statistically significantly different from zero with 90 percent confidence.

Appendix

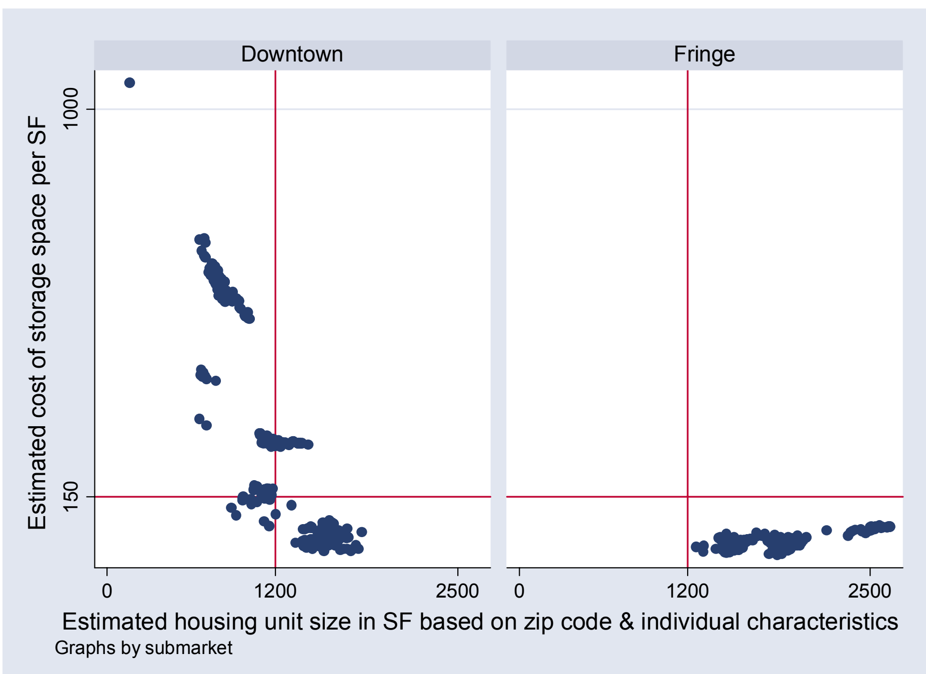
Appendix 1: Map of Chicago Downtown and Urban Fringe



Appendix 2: Distribution of Consumer Storage Constraints and Storage Costs

Figure A3-1: Distribution of Zip Code Specific Imputed Housing Unit Sizes and Consumer Storage Costs, by Submarket

Figure A3-2: Distribution of Household Specific Imputed Housing Unit Sizes and Consumer Storage Costs, by Submarket



Appendix 3: Imputation Method for Storage Constraint and Alternative Storage Cost Measures

This Appendix provides details related to the creation of our preferred storage constraint variable and a number of alternative proxies that we use to test the sensitivity of our results with respect to the choice of the storage constraint/storage cost variable.

We first describe the creation of our preferred proxy: the estimated average *housing unit size in the zip code* in which the panelist is located. Next, we explain how we generate an alternative proxy measure that captures storage *(opportunity) costs* rather than the magnitude of constraints. Lastly, we describe a method of how to impute *panelist specific* storage cost and constraint measures based on the demographic characteristics of the households. We begin by describing our preferred proxy measure.

Zip Code Level Storage Constraint and Storage Cost Measures

The average square footage of living area must be imputed for each zip code, as the US Census does not collect it. Because the American Housing Survey (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 1990 US Census. Specifically, 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:

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

The adjusted R^2 is 56.5%. Next we impute the average housing unit size for all US zip codes with available data using the coefficients from equation (A1) and zip code level housing unit characteristics from the 1990 US Census. Finally, we allocate the imputed measures to the panelists that are included in the regression samples based on their residential locations (71 zip codes).

While our preferred measure provides a good approximation of the true *storage constraints* that consumers are facing, one could also imagine that the *(opportunity) cost of storage*—the house value or rent per square foot—is an important factor in determining consumer's shopping patterns and individual purchase decisions. The measure is derived as follows: The median house value per square foot in a zip code can be computed as the median value of a housing unit in a particular zip code *j* (derived from the 1990 US Census) divided by the predicted average square footage of living area of zip code *j* (our preferred storage cost proxy).

Panelist Specific Storage Constraint and Storage Cost Measures

One limitation of the above two measures is that they are zip code specific rather than panelist specific. This is due to the fact that, so far, we only used zip code specific housing unit characteristics to impute our measures of interest. However, the size and value of a household's housing unit may also be partly inferred from individual demographic characteristics, which are available from the AHS and the Stanford Market Database. In order to derive individual storage constraint and storage cost measures we use in a first step the National AHS 1989 in order to impute the individual housing unit size 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:

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

The adjusted R^2 of the regression is 42.3%. This value is relatively low and raises some concerns with respect to the imputation of *individual* storage constraint and storage cost measures. The relatively low R^2 is also the reason for why we use the *zip code level* specific storage constraint measure as our preferred proxy variable.

Next, we use the estimated coefficients from equation (A2) to impute the individual housing unit sizes of the panelists. The predicted unit size for panelist *i* in zip code *j* can be computed as follows:

(A3) $$\textit{unit size of panelist}_i = \phi \textit{ unit size in zip code}_j\left(i\right) + \hat{\beta} \times \left(X_i \text{-} X_j \right).$$

The average unit size in zip code *j* is our preferred (imputed) storage constraint proxy measure described above, $\hat{\beta}$ denotes the vector of predicted coefficients from equation (A2), and the vectors of variables X_j and X_i denote the individual demographic characteristics of panelist *i* and the average demographic characteristics of zip code *j*. 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 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).

The individual house value of panelist *i* is computed using the identical methodology as described above except that we derive the median house value in zip code *j* directly from the 1990 US Census and thereby do not have to rely on an imputed measure. Finally, the individual house value per square foot can be computed as the imputed individual house value divided by the imputed individual housing unit size.