

Quantity Surcharge and Consumer Heterogeneity

Hyewon Kim*

The Ohio State University

November 14, 2016

Abstract

This paper studies quantity surcharge and heterogeneity in consumer attention, using scanner data in the peanut butter category. Quantity surcharge exists for packaged goods when a smaller package size is cheaper than its larger size counterpart per unit. Quantity surcharges are frequent in the data, and households have heterogeneous purchasing behavior during quantity surcharge weeks: attentive households purchase multiple smaller jars, but inattentive households choose an inferior option of buying larger jars. However, no existing demand model can accommodate the inattentive type's purchasing behavior. Hence, I develop a dynamic demand model that features heterogeneity in consumer attention and multiple-jar purchases. I assume four household types: attentive non-storer, inattentive non-storer, attentive storer, and inattentive storer. The demand model departs from the standard single discrete choice model in two ways: 1) it separates purchase quantity and package size choice and 2) it considers consumption quantity as a countable variable. I identify household types using the method of *kmeans* and estimate the parameters using maximum likelihood. The estimation results show that storer type households are more price sensitive than non-storer type households, and between the two types of storer households, attentive storers are more price sensitive than inattentive storers.

*I am grateful to Jason Blevins, Maryam Saeedi, Bruce Weinberg, and especially Javier Donna for their guidance and support. I would like to thank JP Dubé and seminar participants at the Ohio State University Micro Brown Bag for helpful comments. I would also like to thank SymphonyIRI Group, Inc. for making the data available. All estimates and analyses in this paper, based on data provided by SymphonyIRI Group, Inc. are by the author and not by SymphonyIRI Group, Inc. Any errors are my own. Contact information: kim.3896@osu.edu

1 Introduction

Quantity surcharges exist when a small size package is cheaper than its larger size counterpart per unit (e.g., roll, ounce). Quantity surcharge is a type of nonlinear pricing for packaged goods, which is the opposite of quantity discount. Though it sounds counterintuitive, quantity surcharge is frequently observed at grocery stores. [Widrick \(1979\)](#) finds in his cross-sectional analysis that 33.8% of the brands in his data set, including canned tuna fish and laundry detergents, had quantity surcharges at store level. Data plans for smartphones that allow customers to use a certain amount of data at a fixed fee and then charge customers a lot more per byte once they pass the limit are also an example of quantity surcharge.

In this paper, I study consumer heterogeneity in attention when quantity surcharge exists, using rich scanner data in the peanut butter category. Consumers' attentiveness to quantity surcharge varies, leading to a heterogeneous purchasing pattern. The two most distinct patterns are purchasing multiple small size jars versus one large size jar in quantity surcharge weeks. I develop a structural demand model that allows heterogeneity in consumer attention, and then estimate the model using the method of maximum likelihood.

Even though consumers face quantity surcharge in their daily lives, minimal research has been done on this topic. Also, most of this research is limited to use store level data, and fails to catch heterogeneity in consumer behavior when quantity surcharge exists. Analyzing heterogeneity in attention is important, especially for packaged goods in the presence of nonlinear pricing. Not considering consumer inattention will overemphasize preference on package size or underestimate the price elasticities of some households when estimating demand.

A substitution opportunity exists during quantity surcharge periods. That is, consumers who demand a large quantity can save money by purchasing multiple small size items instead of one large size item. Assuming no preference on large package size, buying the larger one is an inferior option. However, not all consumers take advantage of this substitution opportunity. [Clerides and Courty \(2015\)](#) find that a significant proportion of the consumers who purchase large packages in quantity discount weeks continue to purchase large packages in quantity surcharge weeks. They call the consumers who are missing this substitution opportunity in quantity surcharge weeks "inattentive".

The literature provides two explanations for quantity surcharge. One explanation is that retailers

price discriminate against consumers with high demand (Agrawal, Grimm, and Srinivasan (1993)). The other argues that quantity surcharge exists as a side effect of temporary price promotions when only the small package is on sale (Sprott, Manning, and Miyazaki (2003) and Clerides and Courty (2015)). The welfare effect of quantity surcharge to consumers is ambiguous in both cases. The key determinant of consumer welfare under the existence of quantity surcharge is identifying the inattentive consumers and understanding their behavior.

The peanut butter market is highly concentrated: the top two national brands together take roughly 60% of market. Within each brand, product differentiation is relatively limited, making it easier to compare apple to apple. Both quantity discount and quantity surcharge are widely observed at grocery stores, with quantity discount being slightly more frequent (55% to 61%) than quantity surcharge.

There are three household purchasing patterns to consider. First, households store peanut butter for future consumption. Peanut butter products are storable, as they have long storage life, and I find evidence of storing behavior from my data. Second, multiple jar purchases are highly frequent, especially small size jars in quantity surcharge weeks. Lastly, some households purchase large size jars in quantity surcharge weeks. The frequency of quantity surcharge and the heterogeneous purchasing patterns in the peanut butter category make it ideal for studying quantity surcharge.

Combining the last two purchasing patterns, we can identify the heterogeneous household type. One type of household takes advantage of quantity surcharge and purchases two jars of small size items. Another type of household is not aware of the existence of quantity surcharge and makes inferior package size choices. I call the former type of household attentive and the latter inattentive.

I develop a structural demand model for packaged goods that considers the three purchasing patterns found. First, I allow four types of households: attentive storer, inattentive storer, attentive non-storer, and inattentive non-storer. Different types of households have different preferences on the product. In each period a household chooses consumption quantity, purchase quantity, and the number of jars of each size to purchase. By separating purchase quantity decision and jar size choice, inattentive type households are allowed to make inferior jar size choices for the given purchase quantity. No existing model considers consumer inattention. ¹

¹Most of the papers that consider purchase quantity assume that package size is equal to the purchase quantity. Allenby, Shively, Yang, and Garratt (2004) do separate jar size choice from purchase quantity decision, but the authors assume all house-

I design a multiple-discrete choice model by considering consumption quantity as a countable variable. More specifically, I assume that consumption quantity follows a negative binomial distribution. This distributional assumption also allows deriving the likelihood function for estimation. I simplify a household's dynamic purchase decision with inventory holding using the concept of effective price by [Hendel and Nevo \(2013\)](#). Under certain assumptions, the storer type household's dynamic problem becomes static where consumption quantity for each period is determined by the effective price.

I estimate the model using the method of maximum likelihood. The rich household level purchase data allow me to estimate the demand parameters. Even though the model is limited to one product, the estimation results still give us helpful insight. The estimation results suggest that attentive storer type households are most price sensitive and attentive non-storer type households are least price sensitive. As price increases by little, attentive storer type households would decrease consumption quantity by 2.21 oz, whereas attentive non-storer type households would decrease it by 0.42 oz only.

1.1 Literature Review

The marketing literature has only a few papers on quantity surcharge. [Widrick \(1979\)](#) is the first paper to measure the frequency of quantity surcharge. The author focused on 10 product categories at 37 grocery stores in upper New York State using cross-sectional data, and found a high percentage of incidents of quantity surcharge across the categories he investigated (e.g., 84.4% for canned tuna fish and 33.3% for laundry detergent).

[Agrawal, Grimm, and Srinivasan \(1993\)](#) and [Sprott, Manning, and Miyazaki \(2003\)](#) study the determinants of quantity surcharge. [Agrawal, Grimm, and Srinivasan \(1993\)](#) interpret quantity surcharge as a practice of price discrimination against households with high demand. On the other hand, [Sprott, Manning, and Miyazaki \(2003\)](#) argue that stores practice quantity surcharge to products whose small packages take a high sale volume in order to build a low store-price image. However, none of the papers mentioned above analyzes household purchasing behavior with the existence of quantity surcharge.

[Clerides and Courty \(2015\)](#) is the first attempt to analyze quantity surcharge on the consumer side, to the best of my knowledge. Using the store level data in the laundry detergents category in the Netherlands are attentive.

lands, the authors found that roughly 45%–75% of households are inattentive in the sense that they miss the substitution opportunity to switch from the large package size to the small size in quantity surcharge weeks. The authors suggest search costs as a rationale for consumer inattention: consumers vary in search costs for finding the best deal, and for those consumers with high search costs, it is rational for them not to spend time on keeping track of quantity surcharge. The authors call it "rational" inattention. However, due to the limitation of store level data, the authors cannot analyze the behavior of "attentive" households.

[Hendel and Nevo \(2003\)](#) suggest methods to test whether households engage in stockpiling at both household level and store level.² [Hendel and Nevo \(2006\)](#) introduce a dynamic demand model for durable goods. The model separates a household's decisions on how much to consume and how much to purchase. However, the model does not consider the substitution among different package sizes of the same product. That is, the package size a household chooses is automatically equal to the quantity the household chooses to purchase. Hence it is not eligible to analyze the effect of quantity surcharge where a substitution opportunity exists.

[Hendel and Nevo \(2013\)](#) show how to simplify a complicated dynamic demand model with inventory holdings using certain assumptions. The model first assumes two types of households, storers and non-storers, and allows the two types to have different preferences. Adding further assumptions on storing technology and households' foresight on future demand and future prices, prices from only a short period of adjacent time window become effective for the current period decisions, instead of the full history of past and future prices. However, again, [Hendel and Nevo \(2013\)](#) do not consider package size choice. Hence I adopt the assumptions from [Hendel and Nevo \(2013\)](#) to simplify the inventory holding part, but include additional components to analyze various jar size choices.

One of the key purchasing patterns in the peanut butter category is frequent multiple jar purchases. Hence allowing the multiple discrete choice is crucial for the analysis. Several papers have tried to handle the multiple discrete choice problem. [Hendel \(1999\)](#) suggests a multiple discrete choice model for personal computers allowing both multiple units and multiple brands, when personal computers are differentiated. [Kim, Allenby, and Rossi \(2002\)](#) and [Dubé \(2005\)](#) also take similar approaches to handle

²I follow the methods the authors suggest to find evidence of consumer inventory holding in section 3.1

the multiple discrete choice. All the three papers mentioned consider that consumers buy complementary products to cater for their various needs.³ Hence the authors cannot handle a household deciding whether to purchase two small package items or one large package item of the same product.

To summarize, my paper is the first analysis of heterogeneous household behavior when quantity surcharge exists. I propose a novel approach to structurally analyze quantity surcharge using a dynamic demand model that allows inventory holdings and multiple unit purchases of the same product.

The paper proceeds as follows. Section 2 describes the data sets used for the analyses and provides summary statistics. Section 3 shows the results of several reduced form analyses, including evidence of household stockpiling behavior and heterogeneous purchasing patterns with the existence of quantity surcharge. Section 4 introduces a structural demand model based on the purchasing patterns identified in section 3. Section 5 describes the estimation procedure, and section 6 shows the estimation results. Section 7 concludes the paper.

2 Data

2.1 Description of Data

I use weekly panel scanner data collected by Information Resources Inc. (IRI) for the analysis. The data set covers a period from January 2001 to December 2011, and 31 categories of products, including beer, carbonated beverages, and laundry detergent. IRI collected data from supermarkets in 50 regional markets defined by IRI, and also from household panels from two of the regional markets. For the analysis, I focus on the peanut butter category in the Eau Claire, Wisconsin, market, from January 2008 to December 2010. Eau Claire is one of the two markets that have both store level and household level data.⁴ The store level data consist of weekly peanut butter product sales observations at store-UPC (Universal Product Code) level in Eau Claire, including the number of jars sold, total sales in terms of dollar amount, and information on promotional activities. The household level data consist of peanut butter purchases, complete history of supermarket visits including those with no peanut butter purchase, and household

³Especially [Dubé \(2005\)](#) considers items with the same characteristics but different package sizes as complements, whereas I consider them as perfect substitutes.

⁴The other household panel market is Pittsfield, Massachusetts.

demographic characteristics. There is also an additional data set available on the attributes of peanut butter products.

I merged the three household level data sets and product attributes data in order to extract the maximum information on households' product purchases. Then I merged the store level data set to get additional information on price levels and promotional activities. I dropped the observations that were not matched during the merging processes. Section A.2 describes how I merged each data set in detail. The final data set contains 2,883 households in total with 601,114 trip observations. Among them, 2,616 households purchased peanut butter products at least once, and these households made 26,024 purchase observations in total. Those purchase observations include 27 brands and 128 UPCs from six grocery stores. From now on, I use "grocery stores" and "stores" interchangeably.

2.2 Peanut Butter Market in Eau Claire

The big three national brands in the peanut butter market are Jif, Skippy, and Peter Pan. Jif and Skippy have especially high market shares in Eau Claire, at 31% and 28%, respectively. The market shares of the top 10 selling brands are listed in Table 1. J.M. Smucker Company, Jif's current parent company, acquired Jif from P&G in 2001. In addition to other Jif brands, such as Simply Jif, Jif to Go, and Jif Natural, J.M. Smucker Company also owns Santa Cruz Organic and its own peanut butter brand, Smucker's. Skippy belongs to Unilever, which also owns Skippy Natural and Skippy Super Chunk. The third top selling brand, Private Label, is actually not a brand. It is also known as a store brand or generic brand, and its market share increased over the period I analyzed (2008–2010).

The panel data include 2,883 households in total. These households visited stores more than once a week on average, and most of them visited multiple stores during the period I studied. Among them, 2,616 households purchased at least one jar of peanut butter products. Fifty percent of the households who purchased peanut butter products purchased more than 10 jars, and 25% of them purchased more than 20 jars. One interesting purchasing pattern is that households frequently purchased multiple jars of peanut butter products on a single trip: 46% of the times when they purchased peanut butter products, these households purchased more than one jar.

Looking into the multiple jar purchases in detail, most of them are of the exactly same product, rather

Table 1: Top 10 Selling Brands

Rank	Brand	Parent Company	Market Share (%)	Cum. Market Share (%)
1	Jif	J.M. Smucker Co.	31.02	31.02
2	Skippy	Unilever	28.25	59.27
3	Private Label	Private Label	16.39	75.66
4	Peter Pan	Conagra Foods, Inc.	7.87	83.53
5	Smucker's	J.M. Smucker Co.	5.17	88.71
6	Skippy Natural	Unilever	3.50	92.21
7	Skippy Super Chunk	Unilever	2.43	94.64
8	Smart Balance	Smart Balance, Inc.	1.13	95.77
9	Simply Jif	J.M. Smucker Co.	1.12	96.88
10	Holsum	Holsum Foods	0.51	97.39

Note: The data used are store level peanut butter product sales data from six grocery stores in the Eau Claire market, 2008–2010. The market share is based on the total number of peanut butter jars sold in the market.

than of different peanut butter products. Table 2 shows the relationship between multiple jar purchases and multiple UPC purchases. The data show that 9,497 times, the households purchased two jars of peanut butter products on a single trip, and 8,956 times, the two jars have the same UPCs. Since the UPC is the finest way to define a product, two jars with the same UPC are identical. Similarly, 1,392 out of 1,955 times when the households purchased three or more jars, the jars have the same UPC. This suggests that these households purchased multiple jars at a time for large quantity, rather than for variety.

Table 2: Household Multi-Jar and Multi-UPC Purchases at a Trip

Num of Jars	Num of UPCs			Total
	1	2	3	
1	13,442	0	0	13,442
2	8,956	541	0	9,497
3+	1,392	538	25	1,955
Total	23,790	1,079	25	24,894

Note: The data used are household level panel data from the Eau Claire market between 2008 and 2010

3 Reduced Form Analyses

In this section, I share the results of the reduced form analyses. First, I show evidence that households stockpile for future consumption. Second, I show the existence and magnitude of quantity surcharge at stores. Third, I provide evidence of heterogeneity in household behavior when quantity surcharge exists.

3.1 Evidence of Stockpiling

I test for evidence of household stockpiling behavior in three ways. Stockpiling can evolve as an optimal strategy for households, as peanut butter is storable and stores often offer discounts on peanut butter products. When a product is relatively cheaper than usual, households can purchase a large quantity and store some of the product as inventory for future consumption. Then, they also tend to delay their next purchase to consume the inventory.

One way to test the household stockpiling behavior is to find evidence of post-promotion dip. If households store a product for later use, the quantity of that product sold during the week following a price promotion is expected to be lower than the quantity sold during the week not following a price promotion at store level. This phenomenon is called "post-promotion dip", as suggested by [Hendel and Nevo \(2003\)](#). I follow [Hendel and Nevo \(2003\)](#)'s approach and define a sale for a UPC at a store as follows: The regular price is defined as the modal price, and sale is any price at least 5% lower than the regular price. There are 159 UPCs sold at six stores during the time period I analyzed, and for simplification, I focus on the top 30 selling UPCs at store level. The combined market share of the top 30 selling UPCs in the market is 80.83% during the time period.

Table 3 is the contingency table of the average number of jars sold in sale/non-sale weeks followed by sale/non-sale weeks.⁵ The most important comparison is the two numbers in non-sale weeks. During non-sale weeks following non-sale weeks, more peanut butter jars were sold on average than during non-sale weeks following sale weeks (221.35 versus 71.82). This trend suggests a post-promotion dip, which indicates inventory holding. We can also see that more peanut butter jars were sold on average when there was no sale the previous week than when there was a sale (311.91 versus 235.58). Similarly, more

⁵I call "sale weeks" the weeks in which the peanut butter products were on sale and "non-sale weeks" the weeks in which the peanut butter products were not on sale.

peanut butter jars were sold on average during sale weeks than during non-sale weeks (318.66 versus 235.92).

Table 3: Evidence of Stockpiling: Average Number of Jars Sold

	Sale _{t-1} = 0	Sale _{t-1} = 1	
Sale _t = 0	221.35	71.82	235.92
Sale _t = 1	368.74	242.88	318.66
	311.91	235.58	

Note: Sale_t is an indicator of sale, which is equal to 1 if a UPC was on sale during the week *t* at a store. A sale is defined as a price at least 5% less than the modal price at each store for the period analyzed (2008–2010). The table presents the average number of the top 30 selling peanut butter jars sold across 156 weeks and six stores during each week.

Next, I run several linear regressions to find evidence of inventory holding at both store and household levels. I follow [Hendel and Nevo \(2003\)](#) and estimate the following econometric model

$$(3.1) \quad \log(q_{jst}) = \delta_1 \log(p_{jst}) + \delta_2 \text{dur}_{jst} + \text{dumvars} + \epsilon_{jst},$$

where q_{jst} is the quantity of UPC *j* sold at store *s* in week *t*, measured as volume normalized by 16 oz; p_{jst} is price per 16 oz of UPC *j* at store *s* in week *t*; dur_{jst} is the duration from the previous sale, measured as the number of weeks divided by 100 from the previous sale for the brand to which UPC *j* belongs in store *s* for any size in week *t*; dumvars includes feature and display dummy variables and store and UPC-specific intercept terms.⁶ Both display and feature are non-price promotion activities. The key parameter is δ_2 , measuring the effect of the duration since the previous sale on quantity purchased. The consumer inventory-holding hypothesis predicts that consumers holding inventory can delay purchases waiting for the next price discount, but as the number of weeks since the last sale increases, the inventory level decreases and the consumers become more likely to purchase the product again. A positive δ_2 supports the story.

Table 4 shows the estimation results using ordinary least squares (OLS). The first column contains the results using the whole sample, and the second column contains the observations from non-sale

⁶Feature is a dummy variable indicating whether a UPC was advertised, and display indicates whether a UPC was displayed at the lobby or end aisle.

weeks only. All the coefficients are significant and have the expected signs: less quantity is sold if the price increases, but feature or display increases the quantity sold. In particular, the estimate of δ_2 is significantly positive, supporting the consumer inventory-holding hypothesis.

Table 4: Evidence of Stockpiling: Store Level

Variable	(1) All Samples	(2) Non-Sale Periods
log price (δ_1)	-2.197*** (0.0319)	-1.962*** (0.0567)
duration from prev sale (δ_2)	0.0462*** (0.00994)	0.0423*** (0.0105)
featdum	0.199*** (0.0209)	0.428*** (0.0473)
dispdum	1.013*** (0.0179)	0.744*** (0.0242)
Observations	42,368	32,578

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The table reports the estimation results of (3.1) using two different samples: the whole sample and the observations from non-sale periods only. Each observation used in the analyses is a weekly sales event of a top 30 selling UPC at a store. A sale for a UPC is defined as any price lower than or equal to the price level 5% below its modal price. A positive coefficient of duration from the previous sale (δ_2) indicates that consumers hold inventory.

I also estimate an econometric model designed to find evidence of consumer inventory holding at household level. The household level data allow us to measure the impact of a sale on the number of jars purchased or the timing of purchase, in addition to the total quantity purchased. The model is

$$(3.2) \quad y_{it} = \alpha + \beta S_{it} + \gamma_i + \epsilon_{it},$$

where y_{it} represents five different measures of quantity purchased by household i at purchase instance t ; S_{it} is an indicator whether the product was on sale; and γ_i is a household-specific effect. The model is adopted from [Hendel and Nevo \(2003\)](#). Among the five different dependent variables, the first three are quantity variables: quantity normalized to 16 oz, number of jars, and size of each jar purchased. A positive coefficient on the sale dummy is expected for normalized quantity and number of jars to sup-

port household stockpiling behavior. As for size of jar, I expect it to have a negative coefficient, as households in the panel data frequently purchased multiple jars of relatively small size (16–18 oz) rather than buying one large size jar.⁷ The last two dependent variables are related to the timing of purchases: days from previous purchase and days to next purchase. A negative coefficient for the days from previous purchase and a positive coefficient for the days to next purchase will support the stockpiling hypothesis.

The within estimator allows individual household intercepts, and it estimates the effects of household stockpiling. Table 5 shows the obtained estimates of β for the five different dependent variables. All the models have significant estimates and the signs are as expected. Column (1) indicates that when there is a sale on an item, households are more likely to purchase a larger volume of that product. Columns (2) and (3) indicate that households are more likely to purchase a higher number of jars and smaller size jars during sale weeks. Columns (4) and (5) suggest that households tend to advance a purchase and wait longer for the next purchase when there is a sale.

Table 5: Evidence of Stockpiling: Household Level

Dependent Variable	(1) Quantity (16 oz)	(2) Num of Jars	(3) Jar Size (16 oz)	(4) Days from Prev Pur	(5) Days to Next Pur
$S_{it} \quad (\beta)$	0.187*** (0.0144)	0.290*** (0.0111)	-0.132*** (0.00632)	-5.711*** (1.248)	3.851** (1.243)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The five columns in this table report the estimation results of the model (3.2) with five different dependent variables. Each observation used in this analysis is a purchase observation at UPC level that a household made on a single trip to a store. A sale for a UPC is defined as any price lower than or equal to the price level 5% below its modal price; and S_{it} is an indicator variable of sale (obtained estimates of the variables other than S_{it} are omitted). Quantity is the amount of a UPC purchased, normalized to 16 oz, and num of jars is the number of jars of a UPC purchased. Jar size is the volume of each jar purchased normalized to 16 oz. Days from prev is the number of days from the last peanut butter product purchase, and days to next is the number of days to the next purchase.

3.2 Evidence of Quantity Surcharge

In this section, I show the frequency and magnitude of quantity surcharge at the stores in my data set. In order to do that, I first define a product and quantity surcharge, and then identify products from the

⁷See Table 8.

data using the definition.

Suppose a product has two different package sizes, small (S) and large (L). Here the two size items are identical except for the package size. The product is quantity surcharged if

$$(3.3) \quad p_S < p_L,$$

where p_S and p_L are prices of the small and large size packages, respectively, normalized to their package sizes. Quantity discount exists if the opposite holds: $p_S > p_L$. It is important to control the product characteristics except for the package size within a product. Otherwise, we cannot differentiate quantity surcharge from price difference due to product differentiation.

I identify the products in the peanut butter category using the definition of a product. Among the 159 UPCs shown in the data, I focus on the top 30 selling UPCs at store level for simplification. I group UPCs with the same observable characteristics but package size. The characteristics variables available in the data are brand, product type, sugar content, processing type, texture, flavor, and salt contents. I identify three products, and each product has three different package sizes. The list of products is presented in Table 6.

The first product identified is a group of three UPCs from the brand Skippy with creamy texture. The item with the smallest package size (16.3 oz) among the three is the most popular peanut butter in the market. I call this product "Skippy Creamy PB". The second and the third products are both from the brand Jif and have creamy texture and chunky texture, respectively. I call these products "Jif Creamy PB" and "Jif Chunky PB", respectively.

We are interested to find out the price dynamics across the three sizes within a product. Figure 1 shows the weekly volume price (price per 16 oz) fluctuations of Skippy Creamy PB for the time period analyzed. The small size item (16.3 oz) has the most volatile price among the three items, as it has more frequent price discounts and these are deeper compared to the discounts for the two larger size items (28 oz and 40 oz). Next, quantity surcharge is highly frequent in this product: the volume price of the small size item is lower than the price levels of the two larger size items during most of the weeks.

It is somewhat complicated to determine quantity surcharge and quantity discount for a product

Table 6: Three Products Identified

Product	UPC	Brand	oz	Texture	Num of Jars
Skippy Creamy PB	14800100643	Skippy	16.3	Creamy	64,601
	14800127068	Skippy	28.0	Creamy	13,940
	14800127072	Skippy	40.0	Creamy	5,010
Jif Creamy PB	25150024128	Jif	18.0	Creamy	59,353
	15150024177	Jif	28.0	Creamy	23,991
	15150072001	Jif	40.0	Creamy	11,483
Jif Chunky PB	15150024135	Jif	18.0	Chunky	14,659
	15150024163	Jif	28.0	Chunky	5,666
	15150072002	Jif	40.0	Chunky	4,278

Note: The table contains the three peanut butter products identified. Three UPCs with different package sizes are identified as a product based on the observable characteristics, such as brand, texture, and processing type. Num of jars is the total number of jars sold for the period analyzed (2008–2010) from six grocery stores combined.

when it has three different sizes, since price comparison is bilateral. Hence I combine the medium size and the large size of each product in Table 6 and call it a large size item. I use the average weekly price of the two sizes for this "combined" large size item. For each product, I count the frequency of quantity surcharge and quantity discount as follows: 1) at each store, I determine whether the product was quantity surcharged or quantity discounted every week during the time period analyzed; 2) I count the number of quantity surcharge weeks and quantity discount weeks at a store; 3) I sum up the number of quantity surcharge weeks across the six stores and do the same thing with the number of quantity discount weeks. The result is presented on the first row of Table 7. It varies across products, but one common pattern is that quantity discount is still more frequent than quantity surcharge, but there are significantly many weeks of quantity surcharge. For example, for Skippy Creamy PB, roughly 38% of the weeks were quantity surcharge weeks on average across the six stores.

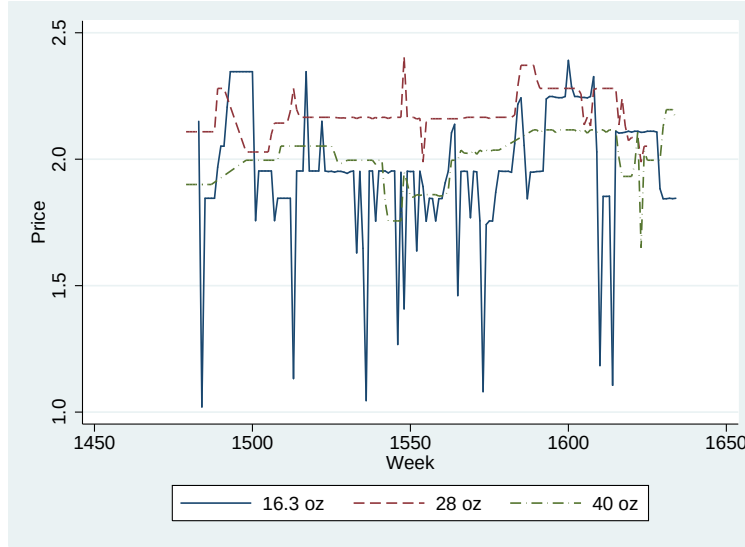
The third row of the first block in Table 7 shows the average magnitudes of quantity discount and quantity surcharge of each product. On average, the price difference is larger during quantity surcharge weeks than during quantity discount weeks. For example, for Jif Creamy PB, during quantity discount weeks, the large size item was cheaper than the small size item by \$0.13 per 16 oz on average. On the other hand, during quantity surcharge weeks, the large size item was more expensive than the small size item by \$0.21 per 16 oz on average.

Table 7: Quantity Discount and Quantity Surcharge at Grocery Stores

	Skippy Creamy PB		Jif Creamy PB		Jif Chunky PB	
	QD ($p_S > p_L$)	QS ($p_S < p_L$)	QD ($p_S > p_L$)	QS ($p_S < p_L$)	QD ($p_S > p_L$)	QS ($p_S < p_L$)
Num of Weeks	536	330	516	420	511	398
% of Weeks	61.89	38.11	55.13	44.87	56.22	43.78
Ave $p_L - p_S$ (\$)	-0.18	0.39	-0.13	0.21	-0.14	0.21
Num of Jars Small	13,979	50,170	23,871	35,482	5,153	9,379
Num of Jars Large	12,486	5,980	21,995	13,479	5,540	4,346
Num of Jars Share Small (%)	64.03	83.52	60.56	74.32	60.53	69.54
Num of Jars Share Large (%)	35.97	16.48	39.44	25.68	39.47	30.46
Ave Weekly Num of Jars Small	26.08	152.03	46.26	84.48	10.08	23.57
Ave Weekly Num of Jars Large	17.57	12.74	28.74	21.82	6.42	5.88
% of Weeks Sale on Small Only	1.49	56.06	1.94	24.52	0.39	20.60
% of Weeks Sale on Large Only	25.75	4.55	24.42	15.48	26.61	28.39
% of Weeks Sale on Both Sizes	19.59	20.91	9.69	24.52	14.68	27.89
% of Weeks No Sale	53.17	18.48	63.95	35.48	58.32	23.12

Note: A product is defined as a group of UPCs with the same observable characteristics but package size. The three products are identified in Table 6 with two different sizes, small and large; p_S and p_L are the volume price (price per 16 oz) of the small and large size item, respectively. A quantity discount (QD) exists for a product at a certain week at a store if $p_S > p_L$ holds, and a quantity surcharge (QS) exists if the opposite ($p_S < p_L$) holds. The statistics in this table are obtained using the weekly price and quantity data during 2008–2010 from six grocery stores in the Eau Claire market. Num of weeks is the number of weeks at the six grocery stores combined when a certain price order existed; ave $p_L - p_S$ is the average volume price difference between the small and the large items across weeks and stores, and it measures the average magnitude of quantity discount and quantity surcharge. The difference is negative in quantity discount weeks and positive in quantity surcharge weeks. Num of jars small (large) is the total number of jars of small (large) size sold for each price order; ave weekly num of jars small (large) is the average number of jars sold of the small (large) across weeks and size. A sale is defined as a price at least 5% less than the modal price at each grocery store for the period analyzed.

Figure 1: Volume Price of Three Skippy Creamy PB UPCs at a Store



Note: The figure shows weekly volume prices (price per 16 oz) of the three items in the product Skippy Creamy PB for the period 2008–2010 at store 1. The three items share the same observable characteristics but package size. Store 1 is the grocery store in the market selling the highest number of peanut butter jars for the three years.

The second and third block of Table 7 compare the number of small and large size jars sold for each product. When we compare the quantity discount weeks and quantity surcharge weeks, we notice households purchase more large size jars during quantity discount weeks than during quantity surcharge weeks, and more small size jars during quantity surcharge weeks than during quantity discount weeks. It is rational for households to purchase large size jars during quantity discount weeks. However, there are positive sales in small size jars during quantity discount weeks. For example, 13,979 jars of small package size Skippy Creamy PB were sold during the period analyzed. This can be due to the low demand for peanut butter products, strong preference for small package sizes, or both. A smaller jar can keep peanut butter more fresh than a larger jar does, as a smaller quantity of product is consumed quicker than a larger quantity. A smaller jar is also more convenient to carry around.

During quantity surcharge weeks, the stores sold a significantly larger number of small size jars than large size jars. For example, the stores sold almost 10 times more smaller jars of Skippy Creamy PB than larger ones, and at least two times more smaller jars of the other two products than larger jars of the same products. In fact, the number of small size jars sold during quantity surcharge weeks is larger than the number of large size jars sold during quantity surcharge weeks, small size jars sold during quantity

discount weeks, and large size jars sold during quantity discount weeks.

There are several possible explanations for this observation. First, some households who do not usually purchase peanut butter products may have purchased small size jars when they were cheap to test the products. In other words, new demand was created during quantity surcharge weeks. If we convert the number of jars sold to the total volume (normalized to 16 oz) sold, a higher volume was sold during quantity surcharge weeks than during quantity discount weeks for all the three products analyzed. This could be because quantity surcharges are sometimes triggered by a temporary price reduction on small size items. For Skippy Creamy PB, during 56% of the quantity surcharge weeks, only the small item was on sale. For the other two products, the small item was on sale between 20% and 25% of the quantity surcharge weeks. The information on quantity surcharge and price reductions on each size is reported in the last block of Table 7.

Second, among those households who regularly purchase peanut butter products, some of them would purchase small size items during quantity surcharge weeks just because that is the usual size they are used to purchase. Last, some of them would strategically purchase small size items because small size items were cheaper per ounce than large size items during quantity surcharge weeks.

On the other hand, there are still positive sales of large size items during quantity surcharge weeks. For example, 5,980 large jars of Skippy Creamy PB and 4,346 large jars of Jif Chunky PB were sold during quantity surcharge weeks over the period analyzed. It is clearly an inferior choice to purchase a large size item during quantity surcharge weeks, since there exists a substitution opportunity: a household can purchase two (or three) jars of a small size item instead of one large size item and get (roughly) the same quantity while paying less. The positive sales of large size items during quantity surcharge weeks imply that some households were unaware of the existence of quantity surcharges and missed the substitution opportunity.

There could be alternative explanations for the positive sales of large size items during quantity surcharge weeks. One possible explanation is that some households prefer large package sizes. However, it is hard to reason that when buying small package sizes has relative advantages to buying large package sizes, as mentioned earlier. Another alternative explanation could be due to stock-out. That is, those households who purchased large size items did so not because they wanted large size items but because

small size items were not available upon their visit to the store. [Conlon and Mortimer \(2013\)](#) argue that ignoring incomplete product availability may bias demand estimates. I cannot verify this argument, as the store level data are recorded at weekly level.⁸ However, the products I am looking at consist of top 30 selling UPCs, and therefore, they are more likely to be stocked compared to the other less popular UPCs. Hence, even if some of the large size items were purchased during quantity surcharge weeks due to households' strong preference for larger size items or stock-outs of small size items, it is hard to tell that those are the major cases.

Quantity surcharges are related to temporary price reductions on small size items, but that is not the only cause of such a high frequency of quantity surcharges. The last block in Table 7 shows the relationship between quantity discount or quantity surcharge and sale. In 50% to 60% of the quantity discount weeks, there was no sale on any size item. A sale on the large size item only is the second most frequent case. A sale on the small size item only was very rare during the quantity discount weeks. Overall, quantity discount seems closely correlated with sale.

However, the relationship between sale and quantity surcharge varies over the products analyzed. For example, for Skippy Creamy PB, quantity surcharge was most frequently caused by a sale on the small size item only, and it was very rarely caused by a sale on the large size item only. For the other two Jif products, however, the relationship between quantity surcharge and sale seems loose, as the four cases (i.e., sale on the small size item only, sale on the large size item only, sale on both size items, and no sale on either size item) are somewhat evenly likely. Hence quantity surcharge is more of a consistent phenomenon than being mainly caused by a sale on the small size item only.

3.3 Quantity Surcharge and Households' Heterogeneous Behavior

As mentioned previously, a substitution opportunity exists during quantity surcharge weeks. If all the households were aware of the substitution opportunity, the two following purchasing patterns would be expected during quantity surcharge weeks: 1) no (or minimal, in case of a strong preference for large size item or stock-out of small size items) large size jar purchase during quantity surcharge weeks and 2) more frequent multiple small size jar purchases than during quantity discount weeks. In other words,

⁸The household level data have information on checkout time. However, this information allows me to observe product availability only when there exists a household who purchased the item of interest.

all the households who would demand a large quantity should have purchased multiple small size jars. In order to test households' awareness, I analyze how multiple jar purchases of each size vary during quantity discount and quantity surcharge weeks at household level. The results are presented in Table 8. A purchase is "single" when a household purchases a single jar on a trip, and "multi" when the household purchases more than one jar on a trip.

Table 8: Quantity Surcharges and Multiple Jar Purchases

	Skippy Creamy PB					
	Small			Large		
	Single	Multi	Total	Single	Multi	Total
QD ($p_S > p_L$)	354	507	861	310	293	603
QS ($p_S < p_L$)	1,340	1,902	3,242	155	53	208
Total	1,694	2,409	4,103	465	346	811
	Jif Creamy PB					
	Single	Multi	Total	Single	Multi	Total
	Single	Multi	Total	Single	Multi	Total
QD ($p_S > p_L$)	659	529	1,188	556	430	986
QS ($p_S < p_L$)	744	1,272	2,016	240	168	408
Total	1,403	1,801	3,204	796	598	1,394
	Jif Chunky PB					
	Single	Multi	Total	Single	Multi	Total
	Single	Multi	Total	Single	Multi	Total
QD ($p_S > p_L$)	159	125	284	119	74	193
QS ($p_S < p_L$)	235	401	636	67	21	88
Total	394	526	920	186	95	281

Note: A product is defined as a group of UPCs with the same observable characteristics but size. The three products are identified in Table 6 with two different sizes: small and large. Variables p_S and p_L are the volume price (price per 16 oz) of the small size and large size items, respectively. A quantity discount (QD) exists for a product during a certain week at a store if $p_S > p_L$ holds, and a quantity surcharge (QS) exists if the opposite ($p_S < p_L$) holds. The statistics in this table are the total number of purchase observations obtained using the three-year (2008–2010) data from 2,616 households in the Eau Claire market. A single (multi) jar purchase observation is defined as a purchase observation where a household visited a store during a week and purchased a single jar (more than one jar) of a peanut butter product.

There is heterogeneity in households' awareness of quantity surcharge. Comparing quantity discount weeks and quantity surcharge weeks for the small size items, there are a lot more multiple jar purchases during quantity surcharge weeks than during quantity discount weeks. This implies that some households were aware of the existence of quantity surcharge and chose to buy multiple small size jars

rather than a large size jar. I classify these households as attentive. On the other hand, there is a positive number of purchase observations of large size items during quantity surcharge weeks, both single and multiple. This implies that there were some other households who were unaware of the existence of quantity surcharge and missed the opportunity to substitute. I classify these households as inattentive.

There could be several reasons why some households are inattentive. Some households might have a limited ability to calculate price per ounce at grocery stores. High search costs could hurry some households and keep them from comparing prices of different size items. [Clerides and Courty \(2015\)](#) call this case rational inattention. Some households might have a strong belief that quantity discount always exists. Those reasons are not mutually exclusive.

In order to identify the characteristics of the inattentive type households, I identify those households from the data and characterize their purchasing patterns. There are 2,046 households who purchased at least one jar of the three products during the period analyzed. Among them, 1,669 households (81.57%) did not miss any chance of substitution. However, 377 households (18.43%) missed the chance once or multiple times: 225 households (11%) missed this chance once and 152 households (7.43%) missed it more than once.

Table 9 shows the purchasing patterns of the different groups of households. Compared to the households with zero misses, the households with one or multiple misses tend to purchase a higher number of jars of the peanut butter products, including large size jars. The households with zero misses purchased 7.54 jars on average during the three-year period analyzed, but the households with one and multiple misses purchased 11.17 jars and 18.59 jars on average, respectively. Considering the households with one or multiple misses tend to purchase more large size jars, the households with one or multiple misses purchased a higher volume than the households with zero misses. This purchasing pattern rejects the hypothesis that the households who purchase the product more frequently have better information on prices and, as a result, are more aware of the existence of quantity surcharge.

The households with one or more misses also tend to purchase more large size jars during quantity discount weeks. Hence, overall, the households who purchased a higher number of jars and purchased more large size jars during quantity discount weeks tend to miss the substitution opportunities more. Purchasing multiple small size jars during quantity surcharge weeks is a strategic behavior to take ad-

Table 9: Missing Substitution Opportunities in QS weeks: Purchasing Patterns

Num of Misses	Num of HHs	Variable	Mean	SD	Min	p25	p50	p75	Max
0	1,669	Num Of Jars Pur	7.54	8.71	1.00	2.00	5.00	10.00	80.00
		Num Of L Jars Pur	0.76	1.93	0.00	0.00	0.00	1.00	23.00
		Num Of L Jars Pur In QD	0.76	1.93	0.00	0.00	0.00	1.00	23.00
		Num Of Multi S Jar Pur In QS	1.76	2.39	0.00	0.00	1.00	2.00	30.00
		Loss (\$)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	225	Num Of Jars Pur	11.17	12.12	1.00	4.00	7.00	15.00	89.00
		Num Of L Jars Pur	3.81	3.99	1.00	1.00	2.00	5.00	26.00
		Num Of L Jars Pur In QD	2.41	3.88	0.00	0.00	1.00	3.00	25.00
		Num Of Multi S Jar Pur In QS	1.75	3.12	0.00	0.00	1.00	3.00	24.00
		Loss (\$)	0.49	0.57	0.00	0.10	0.31	0.65	3.16
2 +	152	Num Of Jars Pur	18.59	14.92	2.00	9.00	15.00	24.00	109.00
		Num Of L Jars Pur	10.93	10.90	2.00	4.00	7.50	14.00	83.00
		Num Of L Jars Pur In QD	6.55	8.83	0.00	1.00	3.50	9.00	58.00
		Num Of Multi S Jar Pur In QS	1.57	1.99	0.00	0.00	1.00	3.00	9.00
		Loss (\$)	1.63	1.98	0.02	0.57	1.06	1.97	15.62
Total	2,046	Num Of Jars Pur	8.76	10.17	1.00	2.00	5.00	11.00	109.00
		Num Of L Jars Pur	1.85	4.59	0.00	0.00	0.00	2.00	83.00
		Num Of L Jars Pur In QD	1.38	3.58	0.00	0.00	0.00	1.00	58.00
		Num Of Multi S Jar Pur In QS	1.75	2.45	0.00	0.00	1.00	2.00	30.00
		Loss (\$)	0.18	0.72	0.00	0.00	0.00	0.00	15.62

Note: The sample is 2,046 households among the household panels who purchased at least one jar of the three products defined in Table 6. A miss is defined as an incident of purchasing a large size item during quantity surcharge weeks. Loss is the total amount of money each household could have saved in a hypothetical scenario that the household had never missed a substitution opportunity during quantity surcharge weeks.

vantage of quantity surcharge, and indicates the households' awareness of the existence of quantity surcharge. The group of households with one miss is similar to the group of households with zero misses in terms of the number of multiple small size jar purchases during quantity surcharge weeks. However, the group of households with multiple misses made fewer multiple small size jar purchases on average than the two other groups.

I calculate the amount of money those households with one or multiple misses could have saved if they had not missed the substitution opportunities and switched to small size jars instead of purchasing large size jars during quantity discount weeks. Whenever a household missed the substitution opportunity, I calculate the price difference between the small size and large size items per ounce and multiply

by the volume (in ounces) of the large size jar that household purchased. Here I hypothetically assume the household would have purchased the same volume of product the household purchased as the cumulative volume of the smaller size jars. Then I sum up the potential savings within a household for the three-year period analyzed. The households who made one miss could have saved \$0.49 on average through substitution, and the households with multiple misses could have saved \$1.63 on average.

To find out what makes the households with missing substitutions different from the households with zero misses, I compare the characteristics of the demographics of the different household groups in Table 10. The income of the households with one or more misses tends to be higher than the income of the households with zero misses. Income is a categorical variable. The median income of the households with zero misses is 7, which means it lies in the range of \$35,000 to \$44,999. The median income of the households with one miss and more than one miss is 9 and 8, respectively (8 is in the range of \$45,000 to \$54,999 and 9 is in the range of \$55,000 to \$64,999). The households in the latter two groups also tend to have a larger family size and are more likely to have children.

In this section I defined quantity discount and quantity surcharge, and found that quantity surcharge is almost as frequently observed as quantity discount. Households show a heterogeneous behavior when quantity surcharge exists, and I identify two types of households based on this heterogeneity. Attentive type households are aware of the existence of quantity surcharge and purchase multiple small size jars instead of one large size jar when they want to purchase a large quantity. On the other hand, inattentive type households are unaware of the existence of quantity surcharge and purchase large size jars. The households who miss this substitution opportunity during quantity surcharge weeks tend to purchase more peanut butter products, both in terms of the number of jars and volume purchased. They also tend to have a higher income and bigger family size.

4 Model

In this section, I develop a demand model that is simple but still captures the key patterns that I find from the data, such as inventory holding and multiple jar purchases.

Table 10: Missing Substitution Opportunities in QS Weeks at Grocery Stores: Demographic Info

Num of Misses	Num of HHs	Variable	Mean	SD	Min	p25	p50	p75	Max
0	1669	Income	7.26	3.15	1	5	7	10	12
		Family Size	2.34	1.22	1	2	2	3	8
		Num Of Children	0.25	0.57	0	0	0	0	3
1	225	Income	8.09	3.06	1	6	9	11	12
		Family Size	2.75	1.28	1	2	2	4	6
		Num Of Children	0.35	0.64	0	0	0	1	3
2+	152	Income	7.83	3.24	1	5.5	8	11	12
		Family Size	2.68	1.16	1	2	2	3.5	6
		Num Of Children	0.35	0.63	0	0	0	1	2
Total	2046	Income	7.39	3.16	1	5	7	10	12
		Family Size	2.41	1.23	1	2	2	3	8
		Num Of Children	0.27	0.59	0	0	0	0	3

Note: The sample is 2,046 households among the household panels who purchased at least one jar of the three products defined in Table 6. A miss is defined as an incident of purchasing a large size item during quantity surcharge weeks. Income is a categorical variable representing the combined pre-tax income of the head of the household. Income is equal to 1 if the combined pre-tax income of the head of the household is in the range of \$00,000 to \$9,999 per year, 2 if in the range of \$10,000 to \$11,999, 3 if in the range of \$12,000 to \$14,999, 4 if in the range of \$15,000 to \$19,999, 5 if in the range of \$20,000 to \$24,999, 6 if in the range of \$25,000 to \$34,999, 7 if in the range of \$35,000 to \$44,999, 8 if in the range of \$45,000 to \$54,999, 9 if in the range of \$55,000 to \$64,999, 10 if in the range of \$65,000 to \$74,999, 11 if in the range of \$75,000 to \$99,999, and 12 if greater than or equal to \$100,000. Family size is the number of members in the household.

4.1 A Simple Demand Model

I propose a demand model when a single good is offered at two package sizes and prices between the two package sizes are not necessarily linear. In each period, households choose not only the quantity to purchase but also the bundle of different package sizes conditional on the given purchase quantity.

4.1.1 The Setup

First, I assume that there is heterogeneity in households' willingness or ability to store.

Assumption 1. *A proportion of households do not store.*

This could arise endogenously if a fraction of the households have storage costs that make it unprofitable to store. Therefore, assumption 1 can be interpreted as an assumption on the distribution of storage costs.

Assumption 2. *A proportion of households are inattentive.*

Household inattention could arise for various reasons: Some households may be unable to calculate the per unit price for an item and compare it to the other package size items. Other households could have high search costs or high opportunity costs to the time spent at a grocery store. Combining Assumption 1 and 2 yields four types of households: attentive non-storer (A-NS), inattentive non-storer (IA-NS), attentive storer (A-S), and inattentive storer (IA-S).

I suppose there are a finite number of periods, R . There is one product and it is offered in two different package sizes, small (S) and large (L). A large size item is twice as big as a small size item. I normalize quantity to small size so that quantity is equal to 1 for the small size item and 2 for the large size item. I define $p_t = (p_{tS}, p_{tL})$ as a vector of the normalized prices (per quantity) for the small and large size items in period t . I allow the four types of households to have different preferences to the product, but none of them has preference to package size. A type h household has indirect utility as follows:

$$(4.1) \quad U_t^h = \beta_0^h + \beta_p^h \tilde{p}_t + \epsilon_t^h,$$

where $\tilde{p}_t$ is the representative price of the product in period t , which I will discuss later; ϵ_t^h represents a stochastic preference shock in period t ; $\beta^h = (\beta_0^h, \beta_p^h)$ is a vector of preference parameters of type h households; β_0^h represents type h households' baseline preference to the product; and β_p^h captures type h households' sensitivity to the price of the product.

In each period, households decide how much to consume, c_t , and how much to purchase, q_t , in terms of quantity, and package size to meet the purchase quantity. That is, both consumption quantity and purchase quantity are discrete variables: $c_t, q_t \in \{0, 1, 2, \dots\}$. If households do not store, quantity to consume and quantity to purchase are the same, but if households store, those two quantities are not necessarily the same. I structure the problem a household faces in a two-stage setting. In stage 1, a household makes quantity decisions: how much to consume and how much to purchase. In stage 2, for the given quantity to purchase, the household decides the number of small package size items, y_{tS} , and the number of large package size items, y_{tL} , to purchase such that $y_{tS} + 2y_{tL} = q_t$. Let $y_t = (y_{tS}, y_{tL})$.

4.1.2 Negative Binomial Distribution

Note that all the three decision variables households choose are discrete, more precisely, nonnegative integers. I introduce the negative binomial distribution to accommodate the discreteness.⁹ Poisson distribution could be an alternative, but negative binomial distribution allows more flexibility, as it has an additional parameter. First, suppose consumption quantity c follows the Poisson distribution conditional on the parameter λ with density function

$$(4.2) \quad f(c|\lambda) = \frac{\lambda^c}{c!} e^{-\lambda}, \quad \text{where } c = 0, 1, 2, \dots$$

Now allow the parameter λ to be random. Let $\lambda = \mu\nu$, where μ is a deterministic function and $\nu > 0$ is iid with gamma density $g(\nu)$,

$$(4.3) \quad g(\nu) = \frac{\delta^\delta}{\Gamma(\delta)} \nu^{\delta-1} e^{-\nu\delta}.$$

Then consumption quantity c follows the negative binomial distribution with a mixture density, as follows:

$$(4.4) \quad \begin{aligned} h(c|\mu, \alpha) &= \int_0^\infty \frac{e^{-\mu\nu(\mu\nu)^c}}{c!} \frac{\nu^{\delta-1} e^{-\nu\delta} \delta^\delta}{\Gamma(\delta)} d\nu \\ &= \frac{\mu^c \delta^\delta \Gamma(c + \delta)}{\Gamma(\delta) c! (\mu + \delta)^{c+\delta}} \\ &= \frac{\Gamma(\alpha^{-1} + c)}{\Gamma(\alpha^{-1}) \Gamma(c + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}} \right)^c, \end{aligned}$$

where $\alpha = 1/\delta$. Using the standard notation of the negative binomial distribution, δ represents the number of failures until the experiment stopped and $\mu/(\mu + \delta)$ is the success probability in each experiment. Lastly, let $\mu = \exp(\beta_0 + \beta_p \bar{p})$, exponential of the mean utility, and $\nu = \exp(\epsilon)$, exponential of the stochastic preference shock. Note that the mean and the variance of ν are 1 and α , respectively. Then consumption quantity c is determined by the binomial distribution with density $h(c|\mu, \alpha)$. The two moments of

⁹Saeedi (2014) also uses the negative binomial distribution to model the sellers' discrete choice of quantity to sell.

consumption quantity are

$$(4.5) \quad \begin{aligned} E[c|\mu, \alpha] &= \mu \\ V[c|\mu, \alpha] &= \mu(1 + \mu\alpha) \end{aligned}$$

As we can see, α determines the variance of v , which is the exponential of the stochastic preference shock, and hence, it also affects the variance of consumption quantity. Thus α can be interpreted as a dispersion parameter.

In order to test whether the consumption quantity can be well represented by the negative binomial distribution, I plot the quantity purchased on a given trip for each household. Figure 2 shows the quantity purchased, instead of the quantity consumed, since the quantity consumed is unobserved. We can see that the negative binomial distribution is a good approximation of the quantity purchased. Conditioning that the quantity consumed is not significantly different from the quantity purchased, the negative binomial distribution is also a good approximation of the quantity consumed.

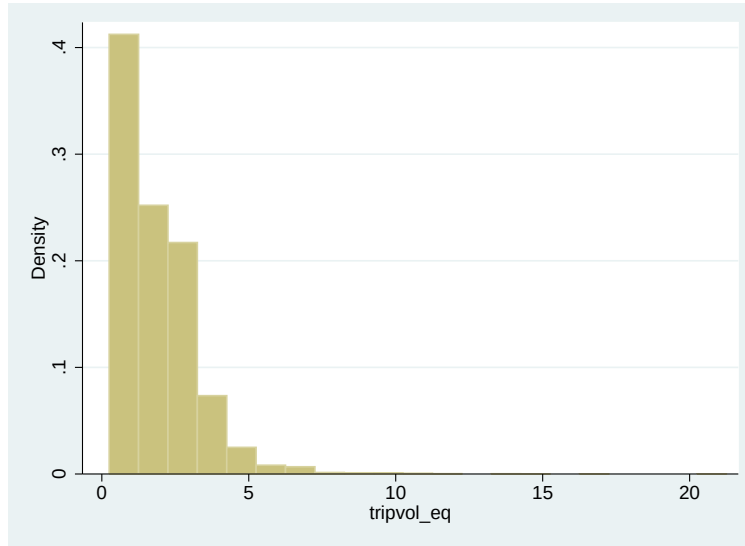


Figure 2: Histogram of Quantity Purchased at Household-Trip Level

Note: The figure shows the histogram of quantity (normalized to 16 oz) each household purchased on a trip, conditional on purchase. The data used are the household level weekly purchase data from 2008 to 2010 in the Eau Claire market.

4.1.3 The Non-storing consumer's problem

In this section, I look at the non-storing household's static problem, while in the next section, I look at the storing household's dynamic problem. First, however, there is one problem to discuss: what price to use for the stage 1 quantity decision? In the previous section, I showed that the consumption quantity is drawn from the negative binomial distribution conditional on two parameters, the exponential of mean utility μ^h and the dispersion parameter α^h . The mean utility of a type h household is determined by the household's preference to the product and its price. Here we face the "one product-two prices" problem. That is, there is one product but two prices exist for the product, one for the small size package and another for the large size package.

In order to determine the mean utility a type h household enjoys from consuming the product, we need to pick a single price $\tilde{p}_t$ for the product. There are several candidates for this single price of the product. We can use either the small size price or the large size price. The lower price between the two sizes or the most frequently charged price across the sizes (regular price) can also be an option. Intuitively, attentive households and inattentive households could use different prices to determine their consumption quantities. However, for simplicity, I assume that both types of households use the minimum price between the two sizes in each period. Let $\underline{p}_t = \min(p_{tS}, p_{tL})$.

In the absence of storing, the purchasing quantity and consumption quantity are the same. How much to consume in each period is determined by the binomial distribution with density $h(\cdot)$ in equation (4.4). Let $\mu^h(\underline{p}_t) = \exp(\beta_0^h + \beta_p^h \underline{p}_t)$, exponential of the mean utility of a type h household when the price the household faces for the quantity decision is $\underline{p}_t$. Also, for simplicity, I limit the consumption quantity to 0, 1, or 2 in this analysis. Then the probability of a non-storer household to choose consumption quantity c is

$$(4.6) \quad \text{Prob}(c_t^h = c) = h(c | \mu^h(\underline{p}_t), \alpha^h), \quad \text{where } c = 0, 1, 2, \text{ and } h = A - NS, IA - NS.$$

For the given quantity decision c_t^h , the two types of non-storing household choose a bundle of two package sizes, small and large, y_t^h . When $c_t^h = 0$ for both types of households, $y_t^h = (0, 0)$ automatically. When $c_t^h = 1$, there is only one option available: purchasing one small size jar. Hence, $y_t^h = (1, 0)$ for both

types of households. However, when $c_t^h = 2$, there are two alternatives available: purchasing one large size jar or two small size jars. Here I take a behavioral approach to attentive and inattentive type households' jar size choice. I assume attentive type households minimize their expenditure and inattentive type households minimize the total number of jars to purchase.

Define quantity discount in period t as $p_{tS} > p_{tL}$ and quantity surcharge as $p_{tS} < p_{tL}$, assuming no tie. Then when $c_t^h = 2$, the attentive non-storer household purchases one large size jar during quantity discount periods and two small size jars during quantity surcharge periods. However, the inattentive non-storer household purchases one large size jar regardless of the period. The two types of non-storing households' jar size choice for a given purchase quantity is presented in Table 11.

4.1.4 The Storing Consumer's Problem

If households store, the quantity purchased and the quantity consumed are not necessarily the same. In order to predict storers' purchases, I make the following assumptions:

Assumption 3. *Storage is free, but inventory lasts for only T periods (fully depreciates afterwards).*

One way to interpret the assumption is the product's perishability. For example, $T = 1$ means the product goes bad in two weeks. An alternative interpretation of the assumption is households' capacity to consider their future consumption at purchase. Even though the product may last longer, households only consider purchasing for T periods ahead. The role of this assumption is to dramatically simplify the state space of the model to consider. That is, I do not need to keep track of how much is left in storage in different states, which is what a typical dynamic problem with inventory holdings would do.

Assumption 4. *(perfect foresight): Households have perfect foresight regarding prices T periods ahead.*

The perfect foresight assumption further simplifies the storing household's problem. One may worry that perfect foresight is restrictive, and thus, invalidate demand estimates. However, [Hendel and Nevo \(2013\)](#) show that this perfect foresight assumption fits the data as well as an alternative assumption with households having a rational expectation of future prices.

I assume $T = 1$ for simplicity in the analysis. The storing household's problem is as follows. The type h storing household decides how much to purchase in each period, $\{q_t^h\}_{t=1}^R$, and how much to con-

sume, $\{c_t^h\}_{t=1}^R$, maximizing the sum of utilities subject to the household's budget constraint and storage technology. Here I ignore discounting for utilities from future consumption. As the non-storing type household's problem, both variables are non-negative integers. Define the effective price in period t as $p_t^{ef} = \min\{\underline{p}_{t-1}, \underline{p}_t\}$. [Hendel and Nevo \(2013\)](#) show that, under Assumptions 3-4, it is equivalent to solve for a series of static optimal consumption quantity with respect to p_t^{ef} . Let $c_t^h(p_t^{ef})$ be the optimal consumption quantity of the type h household. The variable $c_t^h(p_t^{ef})$ is drawn from the negative binomial distribution with density $h(\cdot|\mu^h(p_t^{ef}), \alpha^h)$, as shown in the previous section.

Now I consider the purchase quantity. Since inventory lasts for one period, we only need to consider the two adjacent periods for the purchase quantity decision in period t . The storing household's purchase quantity in period t , q_t^h , is a sum over the consumption in period t if $\underline{p}_t$ is the effective price in period t , and the consumption in period $t+1$ if $\underline{p}_t$ is the effective price in period $t+1$. By definition, $\underline{p}_t$ is the effective price in period t if it is lower than $\underline{p}_{t-1}$, and the effective price in period $t+1$ if it is lower than $\underline{p}_{t+1}$. Define a sale period S as $\underline{p}_t < \underline{p}_{t+1}$ and a non-sale period N otherwise. Then there are four possible events in period t : a sale in period t followed by a sale in period $t-1$ (SS), a sale in period t followed by a non-sale in period $t-1$ (NS), a non-sale in period t followed by a sale in period $t-1$ (SN), and lastly, a non-sale in the two periods in a row (NN). The purchase quantity in period t of the type h storing household in each event can be written as

$$(4.7) \quad q_t^h(\underline{p}_{t-1}, \underline{p}_t, \underline{p}_{t+1}) = \begin{cases} c_t^h(\underline{p}_t) + 0 & \text{in NN;} \\ 0 + 0 & \text{in SN;} \\ c_t^h(\underline{p}_t) + c_{t+1}^h(\underline{p}_t) & \text{in NS;} \\ 0 + c_{t+1}^h(\underline{p}_t) & \text{in SS} \end{cases}, \text{ where } h = A - S, IA - S.$$

Hence, the purchase quantity in period t for the two storing type households is determined by the minimum price between the two sizes in period t , $\underline{p}_t$, and the sale states in the previous period and the current period. Again, I restrict the consumption quantity to no more than 2 for simplicity. Then the purchase quantity can have an integer value from 0 to 4, depending on the sale event. As described earlier, both $c_t^h(\underline{p}_t)$ and $c_{t+1}^h(\underline{p}_t)$ follow the negative binomial distribution with density $h(\cdot|\mu^h(\underline{p}_t), \alpha^h)$. Thus, the probability of a storing type household to choose the purchase quantity q in the sale event NN

or SS is

$$(4.8) \quad q_t^h(\underline{p}_t|NN) = q_t^h(\underline{p}_t|SS) = q \quad \text{with prob. } h(q|\mu^h(\underline{p}_t), \alpha^h), \quad \text{where } q = 0, 1, 2 \text{ and } h = A - S, IA - S.$$

$q_t^h(\underline{p}_t|SN)$ is simply zero with probability 1.

It is little bit more complicated in the event of NS since the storing household purchases the consumption quantity for both period t and period $t + 1$. For example, there are two possible scenarios why the household purchases quantity 1 in period t in the event of NS: 1) the household's consumption quantity is 0 in period t and 1 in period $t + 1$; 2) the household's consumption quantity is 1 in period t and 0 in period $t + 1$. Assuming independence between the consumption quantity in period t and period $t + 1$, the probability of having each purchase quantity is as follows:

$$(4.9) \quad q_t^h(\underline{p}_t|NS) = \begin{cases} 0 & \text{with prob. } h(0|\mu^h(\underline{p}_t), \alpha^h)h(0|\mu^h(\underline{p}_t), \alpha^h) \\ 1 & \text{with prob. } h(0|\mu^h(\underline{p}_t), \alpha^h)h(1|\mu^h(\underline{p}_t), \alpha^h) + h(1|\mu^h(\underline{p}_t), \alpha^h)h(0|\mu^h(\underline{p}_t), \alpha^h) \\ 2 & \text{with prob. } h(0|\mu^h(\underline{p}_t), \alpha^h)h(2|\mu^h(\underline{p}_t), \alpha^h) + h(1|\mu^h(\underline{p}_t), \alpha^h)h(1|\mu^h(\underline{p}_t), \alpha^h) \\ & \quad + h(2|\mu^h(\underline{p}_t), \alpha^h)h(0|\mu^h(\underline{p}_t), \alpha^h) \\ 3 & \text{with prob. } h(1|\mu^h(\underline{p}_t), \alpha^h)h(2|\mu^h(\underline{p}_t), \alpha^h) + h(2|\mu^h(\underline{p}_t), \alpha^h)h(1|\mu^h(\underline{p}_t), \alpha^h) \\ 4 & \text{with prob. } h(2|\mu^h(\underline{p}_t), \alpha^h)h(2|\mu^h(\underline{p}_t), \alpha^h) \end{cases}$$

Conditional on the purchase quantity determined, the storing household decides the number of jars to purchase for each size as the non-storing household does. The attentive storer households minimize their expenditure and the inattentive storer households minimize the total number of jars to purchase. In this section, I discuss the cases where the purchase quantity is larger than 2. When the purchase quantity in period t is 3, the attentive storer chooses one small size jar and one large size jar if quantity discount exists in period t , and three small size jars in the case of quantity surcharge. The inattentive storer, on the other hand, purchases one small size jar and one large size jar all the time, for the given purchase quantity 3. When the purchase quantity is equal to 4, the attentive storer purchases two large size jars if quantity discount exists, and four small size jars if quantity surcharge exists. Again, the inattentive storer's jar size choice does not depend on quantity discount or quantity surcharge, and the household

purchases two large size jars in both cases. Table 11 describes each type of households' jar size choice for a given purchase quantity.

Table 11: Jar Size Choices of Four Household Types, for Given Purchase Quantity

HH type q	A-NS		IA-NS	A-S		IA-S
	QD	QS		QD	QS	
0	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
1	(1,0)	(1,0)	(1,0)	(1,0)	(1,0)	(1,0)
2	(0,1)	(2,0)	(0,1)	(0,1)	(2,0)	(0,1)
3				(1,1)	(3,0)	(1,1)
4				(0,2)	(4,0)	(0,2)

Note: There are four types of households: attentive non-storer (A-NS), inattentive non-storer (IA-NS), attentive storer (A-S), and inattentive storer (IA-S). Variable q represents the given purchase quantity. For a given quantity, the attentive type households choose the number of jars to purchase for each size, (y_S, y_L) , depending on quantity discount or quantity surcharge state. However, non-attentive households' jar size choice does not depend on quantity discount or quantity surcharge.

5 Estimation

5.1 Empirical Model

In this section, I modify the simple demand model from section 4 for estimation using household level purchase data in the peanut butter category. There is one peanut butter product with two package sizes, small and large. There are S stores carrying the product and N households in the market. Assume there are four types of households: attentive non-storer (A-NS), inattentive non-storer (IA-NS), attentive storer (A-S), and inattentive storer (IA-S). Household types are not observable to researchers, but can be identified based on the households' purchasing patterns described in section 5.3. Let N_1 and N_2 be the number of A-NS and IA-NS type households, and N_3 and N_4 the number of A-S and IA-S type households, respectively.

There is a finite number of weeks, R . I consider a household's visit to a store s in week t as given. This assumes that households do not plan to visit a store just to purchase the peanut butter product. Upon a visit, a type h household i observes the volume prices of the small and large size items, p_{tSs} and p_{tLs} , respectively, and decides how much to consume, c_{it}^h , how much to purchase, q_{it}^h , and how many jars of each size to purchase, $y_{it}^h = (y_{itS}^h, y_{itL}^h)$, such that $y_{itS}^h + 2y_{itL}^h = q_{it}^h$. From the data we observe the

following: week t and store s each household visited, $y_{it}^h = (y_{itS}^h, y_{itL}^h)$, and $p_{ts} = (p_{tSs}, p_{tLs})$.

The probability to observe a purchase incident of $y_{it}^h = (y_{itS}^h, y_{itL}^h)$ is the probability for the type h household i of having the purchase quantity $q_{it}^h = y_{itS}^h + 2y_{itL}^h$ multiplied by the probability of making such a jar choice (y_{itS}^h, y_{itL}^h) for the given purchase quantity. This can be written as

$$(5.1) \quad \text{Prob}(y_{it}^h | \alpha^h, \beta^h) = \text{Prob}(y_{it}^h | q_{it}^h) \times \text{Prob}(q_{it}^h | \mu_{it}^h, \alpha^h),$$

where α^h is the dispersion parameter and $\beta^h = (\beta_0^h, \beta_p^h)$ is a vector of the preference parameters. The first term on the right hand side, $\text{Prob}(y_{it}^h | q_{it}^h)$, represents the probability of making such a jar size choice y_{it}^h for the given purchase quantity q_{it}^h . The second term on the right hand side, $\text{Prob}(q_{it}^h | \mu_{it}^h, \alpha^h)$, represents the probability of purchasing quantity choice q_{it}^h , where $\mu_{it}^h = \exp(\beta_0^h + \beta_p^h \underline{p}_{ts})$ when s is the store household i visited in week t .

The quantity that a non-storer type household ($h = A - NS, IA - NS$) purchases is equal to the consumption quantity drawn from the negative binomial distribution. Hence the household's purchase quantity follows the same negative binomial distribution as the household's consumption quantity. Therefore, $\text{Prob}(q_{it}^h | \mu_{it}^h, \alpha^h)$ is equal to $h(q_{it}^h | \mu_{it}^h, \alpha^h)$, where $h(\cdot)$ is the density function of the negative binomial distribution, as shown in equation (4.4).

For the two storer types ($h = A - S, IA - S$), the purchase quantity depends on the consumption quantity and the sale event, as shown in equation (4.7). When the sale event is NS or SS, storer type households purchase in week t the future consumption quantity (for week $t + 1$). I would like to emphasize that storer type households also draw their consumption quantity in week $t + 1$ from the same negative binomial distribution with density $h(\cdot | \mu_{it}^h, \alpha^h)$ as the households draw their consumption quantity in week t , since these households face the same price $\underline{p}_{ts}$ for both. Thus the probability of a purchase quantity choice of a storer type household is

$$(5.2) \quad \begin{aligned} \text{Prob}(q_{it}^h | \mu_{it}^h, \alpha^h) &= h(q_{it}^h | \mu_{it}^h, \alpha^h)^{1 - \mathbf{1}(SN) - \mathbf{1}(NS)} \times \mathbf{1}(q_{it}^h = 0)^{\mathbf{1}(SN)} \\ &\times \left(\sum_{l=0}^{q_{it}^h} h(l | \mu_{it}^h, \alpha^h) h(q_{it}^h - l | \mu_{it}^h, \alpha^h) \right)^{\mathbf{1}(NS)}, \quad h = A - S, IA - S. \end{aligned}$$

For a given visit to a store s in week t and the purchase quantity q_{it}^h , the type h household i decides how many jars to purchase for each package size, y_{itS} and y_{itL} . Attentive type households ($h = A - NS, A - S$) choose y_{it}^h to minimize their expenditure. Hence the jar size choice depends not only on the purchase quantity but also on the quantity discount or quantity surcharge status. During quantity discount weeks, an attentive household purchases as many large size jars as possible for the given purchase quantity. During quantity surcharge weeks, on the other hand, an attentive household purchases small size jars only. Thus, attentive type households' jar size choice probability for the given purchase quantity, depending on the quantity discount or quantity surcharge status, is

$$(5.3) \quad \begin{aligned} \text{Prob}(y_{it}^h | q_{it}^h, \text{QD}_t) &= \begin{cases} 1 & \text{if } y_{it}^h = (q_{it}^h - 2\lfloor \frac{1}{2}q_{it}^h \rfloor, \lfloor \frac{1}{2}q_{it}^h \rfloor) \\ 0 & \text{otherwise} \end{cases} \\ \text{Prob}(y_{it}^h | q_{it}^h, \text{QS}_t) &= \begin{cases} 1 & \text{if } y_{it}^h = (q_{it}^h, 0) \\ 0 & \text{otherwise if } h = A - NS, A - S. \end{cases} \end{aligned}$$

However, inattentive type households ($h = IA - NS, IA - S$) choose y_{it}^h to minimize the total number of jars to purchase. Therefore, inattentive households' jar size choice depends on the purchase quantity only, and they purchase as many large size jars as possible all the time. Inattentive type households' jar choice probability for the given purchase quantity is

$$(5.4) \quad \text{Prob}(y_{it}^h | q_{it}^h) = \begin{cases} 1 & \text{if } y_{it}^h = (q_{it}^h - 2\lfloor \frac{1}{2}q_{it}^h \rfloor, \lfloor \frac{1}{2}q_{it}^h \rfloor) \\ 0 & \text{otherwise if } h = IA - NS, IA - S. \end{cases}$$

Let ϕ_{it}^h be the probability of jar choice of a type h household i in week t for simplification. Then the

log likelihood function can be written as

$$\begin{aligned}
(5.5) \quad \log L(\alpha, \beta) &= \sum_h \log L^h(\alpha^h, \beta^h), \quad \text{where} \\
\log L^h(\alpha^h, \beta^h) &= \sum_{t=1}^R \sum_{i=1}^{N^h} \left\{ \log \phi_{it}^h + \log h(q_{it}^h | \mu_{it}^h, \alpha^h) \right\} \quad \text{for } h = A - NS, IA - NS, \\
\log L^h(\alpha^h, \beta^h) &= \sum_{t=1}^R \sum_{i=1}^{N^h} \left\{ \log \phi_{it}^h + (1 - \mathbf{1}(SN) - \mathbf{1}(NS)) \log h(q_{it}^h | \mu_{it}^h, \alpha^h) + \mathbf{1}(SN) \log(\mathbf{1}(q_{it}^h = 0)) \right. \\
&\quad \left. + \mathbf{1}(NS) \log \left(\sum_{l=0}^{q_{it}^h} h(l | \mu_{it}^h, \alpha^h) h(q_{it}^h - l | \mu_{it}^h, \alpha^h) \right) \right\} \quad \text{for } h = A - S, IA - S.
\end{aligned}$$

The parameters' identification is discussed in Appendix B.

5.2 Data Cleaning

I clean the data in the following way to fit them to the model presented in section 5.1. The data set I use for the estimation includes household level peanut butter product purchase data combined with trip data from the Eau Claire, Wisconsin, market from 2008 to 2010. I also use the store level peanut butter sales data for price information. I defined the three products earlier (see Table 6): Skippy Creamy PB, Jif Creamy PB, and Jif Chunky PB. As the model assumes a single product, I combine the three products into one. Any purchase of the three products from the data is considered a purchase, and a purchase of any other products or no purchase is considered a no purchase.

To define the package size, I consider the 16.3 oz jar of Skippy Creamy PB and the 18 oz jar of Jif Creamy PB and Jif Chunky PB a small size item, and I consider the 28 oz jar of any of the three products mentioned a large size item. The large size item is not exactly twice as big as the small size item in the data. In order to handle this issue, I set the quantity to 1 if the jar size is between 16 oz and 18 oz, and to 2 if the jar size is between 28 oz and 36 oz.

To define the price of each size item, I calculate the average price of the three items of each size at week-store level. Let p_{tks} be the average price per 16 oz of the three size k items in week t at a store s . Then I define sale and quantity discount or quantity surcharge at week-store level using the minimum price between the two sizes, $\underline{p}_{ts}$.

There are cases of multiple trips to the same or different stores in a week, and some households

purchased peanut butter multiple times on those trips. However, the model assumes maximum one visit per week. Hence I drop or merge some of those observations in the following way. First, I identify the trip observations with no purchase. If there are no other trip observations the same week, I keep the observation. If there are other trip observations the same week but there was no purchase, I keep only one representative trip observation.¹⁰ If there is any purchase observation the same week, I drop all the other trip observations with no purchase that week.

Some households made multiple purchases at multiple stores the same week. In that case, I drop all the observations.¹¹ In the case of multiple purchases from the same stores in a week, I merge those purchase observations of the same size. In addition, I drop a few observations that have no matching price information in the store level data. Lastly, I drop the purchase observations with a total quantity purchased greater than 4. Those purchase observations take only 3.1% of the total purchase observations. Lastly, I drop 10 households with no purchase throughout the period analyzed. As a result, the clean data set includes 1,990 households with 9,079 purchase observations and 233,243 trip observations.

5.3 Estimation Procedure

I estimate the parameters of each household type h , (α^h, β^h) , using the method of maximum likelihood. The first step of the estimation is to assign a type to each household. I identify the unobservable household type using the method of *kmeans*. Let h_i be the individual specific moments which are informative about the unobservable individual effects. The *kmeans* method estimates a partition of individual units by finding the best grouped approximation to the moments $\{h_i\}$ based on K groups

$$(5.6) \quad (\hat{\varphi}, \hat{k}_1, \dots, \hat{k}_N) = \operatorname{argmin}_{(\varphi, k_1, \dots, k_N)} \sum_{i=1}^N \|h_i - \varphi(k_i)\|^2,$$

where $\|\cdot\|$ denotes the Euclidean norm; $\{k_i\} \in \{1, \dots, K\}^N$ are partitions of $\{1, \dots, N\}$ into at most K groups; and $\varphi = (\varphi(1)', \dots, \varphi(K)')'$ are $K \times 1$ vectors.

Since I assume four types of households, I set K to 4. I choose one moment for each dimension of household heterogeneity: storing and attentiveness. The first moment is the percentage of quantity pur-

¹⁰I choose the observation recorded earliest in the week.

¹¹An alternative way to handle this case is keeping one purchase observation out of the multiple purchase observations and drop all the rest.

chased in sale weeks (PSALE) for each household. This moment represents the individual household's storing behavior. The second moment is the percentage of quantity purchased as a miss (PMISS). Here a miss is defined as a large size jar purchase in quantity surcharge weeks. The second moment measures the household's attentiveness. For example, a household with low PSALE and low PMISS is likely to be the attentive non-store type.

Using the two moments, I assign an initial household type according to the following criteria: 1) a household is assigned to the storer type if PSALE is larger than 0.5, and to the non-storer type otherwise; 2) a household is assigned to the attentive type if PMISS is 0, and to the inattentive type otherwise. With this initial type assignment, I perform *kmeans* and obtain the estimation results as presented in Table 12. There are significantly more attentive type households than inattentive ones. The attentive storer type has the largest number of households, and the inattentive non-storer type has the smallest number of households.

Table 12: Results of Household Type Assignments

Type	Num of HHs	Num of Obs	Ave PSALE (%)	Ave PMISS (%)
A-NS	659	76,166	0.25	0.02
IA-NS	81	8,734	0.16	0.73
A-S	1,096	131,306	0.85	0.01
IA-S	154	17,037	0.71	0.57
Total	1,990	233,243	0.61	0.09

Note: The four household types are attentive non-storer (A-NS), inattentive non-storer (IA-NS), attentive storer (A-S), and inattentive storer (IA-S). I assign a type to each household using the method of *kmeans*. Two household moments are used to perform *kmeans*: PSALE and PMISS, where PSALE represents the percentage of quantity purchased on sale weeks and PMISS represents the percentage of quantity purchased as a miss.

In addition, I adjust the probability of observing certain jar size choices. When the purchase quantity is larger than 1, multiple alternatives exist in terms of jar size choice. For the attentive types ($h = A - NS, A - S$), the "right" choice is determined by quantity discount or quantity surcharge, as shown in equation (5.3). For example, when the purchase quantity is 2 and quantity surcharge exists, the right choice is to purchase two small size jars. The model predicts the probability of observing $y_{it}^h = (2, 0)$ in that case to be 1, and the probability of observing the "wrong" choice, $y_{it}^h = (0, 1)$, to be 0. However, log of 0 goes to negative infinity and it makes the log likelihood function explode to negative infinity. The same

problem arises for the inattentive type households in some cases. Hence, for technical reasons, I adjust the probability of making the right choice to 0.9 and the probability of making the wrong choice to 0.1.

In addition, for the storing type households, the model predicts the probability of having the purchase quantity equal to 0 to be 1 in the sale event SN, as shown in equation (4.7). Again, the same technical issue arises, so I adjust the probability to have purchase quantity 0 in the SN event to 0.9 instead of 1, and the probability of having a positive purchase quantity to 0.1.

6 Results

I first present the estimation results when all the households are assumed to be non-storers (see Table 13). Column (1) presents the demand estimates when all the households are assumed to be attentive non-storer types. The obtained estimates of the two preference parameters $\beta^{A-NS} = (\beta_0^{A-NS}, \beta_p^{A-NS})$ are 1.3754 and -1.9868. The interpretation of the coefficients is as follows. As shown in equation (4.5), the mean of consumption quantity c_{it}^h is equal to the exponential of the mean utility μ_{it}^h . Thus, one unit change in $\underline{p}_{ts}$ increases the expectation of c_{it}^h by

$$(6.1) \quad \frac{\partial E[c_{it}^h | \underline{p}_{ts}]}{\partial \underline{p}_{ts}} = \beta_p^h \exp(\beta_0^h + \beta_p^h \underline{p}_{ts}).$$

We can calculate the average response by taking the average across individual households' response in each week as follows:

$$(6.2) \quad \frac{1}{RN^h} \sum_{t=1}^R \sum_{i=i}^{N^h} \frac{\partial E[c_{it}^h | \underline{p}_{ts}]}{\partial \underline{p}_{ts}} = \frac{1}{RN^h} \sum_{t=1}^R \sum_{i=i}^{N^h} \beta_p^h \exp(\beta_0^h + \beta_p^h \underline{p}_{ts}).$$

The formula yields the average response -0.1465. That is, when all the households are assumed to be the attentive non-storer types, the average consumption quantity decreases by 0.1465 units when the minimum price between the two sizes of the product increases by one unit.

Column (2) in Table 13 shows the estimation results when the households are divided into two groups: attentive non-storer type and inattentive non-storer type. I calculate the average response of consumption quantity to the change in price for each type. When the minimum price increases by

Table 13: ML Estimation Results: Non-Storer Types Only

Parameter	Description	(1)		(2)	
		Estimate	Std	Estimate	Std
β_0^{A-NS}	preference to the PB product of A-NS	1.3754	0.1229	1.9844	0.1335
β_p^{A-NS}	sensitivity to price of A-NS	-1.9868	0.0604	-2.4117	0.0662
α^{A-NS}	dispersion parameter of A-NS	25.1093	0.4378	24.1308	0.5352
β_0^{IA-NS}	preference to the PB product of IA-NS			0.2191	0.2545
β_p^{IA-NS}	sensitivity to price of IA-NS			-1.1146	0.1240
α^{IA-NS}	dispersion parameter of IA-NS			20.7929	0.6109
-logL	negative value of the log likelihood function	584548		584394	

Note: Column (1) shows the estimation results when every household is assumed to be attentive non-storer types. Column (2) shows the estimation results when households are divided to attentive non-storer group and inattentive non-storer group. In total, 1,592 households who never purchased a large size jar in quantity surcharge weeks are assigned to the attentive non-storer group. The rest of 398 households are assigned to the inattentive non-storer group. Standard errors are obtained using numerical Hessian.

one unit, attentive non-storer type households decrease their consumption quantity by 0.1440 on average. On the other hand, inattentive non-storer type households decrease their consumption quantity by 0.1445 on average. When all the households are assumed to be non-storer types, the average response to price change is very similar between the attentive and inattentive types.

Next I estimate the model assuming all the households are storer type. Column (1) in Table 14 shows the estimation results when all the households are assumed to be attentive storer type. The average response is -0.0926. That is, when all the households are assumed to be attentive storer type, they decrease consumption quantity by 0.0926 on average when the minimum price increases by one unit.

Column (2) in Table 14 shows the estimation results when the households are divided into attentive storer and inattentive storer types. The obtained estimates are $\hat{\beta}^{A-S} = (1.4615, -2.3336)$ and $\hat{\beta}^{IA-S} = (-1.0046, -0.6734)$, and the average responses are -0.0960 and -0.0624, respectively. That is, when all the households are assumed to be storer types, when the minimum price increases by one unit, attentive and inattentive type households decrease their consumption quantities by 0.0960 and 0.0624 units on average, respectively. Unlike the earlier case when all the households are assumed to be non-storer types, inattentive households are less sensitive to price change than attentive households are when they are assumed to be storer types.

Table 14: ML Estimation Results: Storer Types Only

Parameter	Description	(1)		(2)	
		Estimate	Std	Estimate	Std
β_0^{A-S}	preference to the PB product of A-S	0.6562	0.1412	1.4615	0.1402
β_p^{A-S}	sensitivity to price of A-S	-1.8046	0.0715	-2.3336	0.0717
α^{A-S}	dispersion parameter of A-S	32.7915	0.6443	31.0887	0.7696
β_0^{IA-S}	preference to the PB product of IA-S			-1.0046	0.2584
β_p^{IA-S}	sensitivity to price of IA-S			-0.6734	0.1294
α^{IA-S}	dispersion parameter of IA-S			28.1567	0.9552
-logL	negative value of the log likelihood function	586648		586622	

Note: Column (1) shows the estimation results when every household is assumed to be attentive storer types. Column (2) shows the estimation results when households are divided to attentive storer group and inattentive storer group. In total, 1,592 households who never purchased a large size jar in quantity surcharge weeks are assigned to the attentive storer group. The rest of 398 households are assigned to the inattentive storer group. Standard errors are obtained using numerical Hessian.

Lastly, I estimate the model assuming the four groups of households as assigned in Table 12. The estimation results are presented in Table 15. I calculate the average response for each type, which is -0.0264, -0.0387, -0.1383, and -0.0611 for attentive non-storer type, inattentive non-storer type, attentive storer type, and inattentive store type, respectively. Storer types are more sensitive to price change than non-storer types, which is intuitive. Between the two non-storer types, attentive non-storer type households are less price sensitive than inattentive non-storer type households. However, attentive storer type households are more price sensitive than inattentive storer type households.

7 Concluding Comments

I study quantity surcharge at grocery stores and consumers' heterogeneous behavior. The paper has three main contributions. The first is to reconfirm the existence of quantity surcharge, which is often forgotten when packaged goods and nonlinear pricing are of interest. Using peanut butter category scanner data, I find strong evidence of quantity surcharge. It existed in 38% to 45% of the weeks between 2008 and 2010 at grocery stores in Eau Claire, Wisconsin. During those quantity surcharge weeks, small size peanut butter jars were cheaper than the corresponding large size jars by \$0.20 to \$0.40 per 16 oz on average.

Table 15: ML Estimation Results: All Four Types

Parameter	Description	Estimate	Std Error
β_0^{A-NS}	preference to the PB product of A-NS	-1.8776	0.2182
β_p^{A-NS}	sensitivity to price of A-NS	-0.3687	0.1055
α^{A-NS}	dispersion parameter of A-NS	27.0915	0.8151
β_0^{IA-NS}	preference to the PB product of IA-NS	-1.5325	0.8728
β_p^{IA-NS}	sensitivity to price of IA-NS	-0.5943	0.4294
α^{IA-NS}	dispersion parameter of IA-NS	38.3992	3.5634
β_0^{A-S}	preference to the PB product of A-S	2.3093	0.1532
β_p^{A-S}	sensitivity to price of A-S	-2.6624	0.0787
α^{A-S}	dispersion parameter of A-S	25.9842	0.6863
β_0^{IA-S}	preference to the PB product of IA-S	-0.9909	0.4671
β_p^{IA-S}	sensitivity to price of IA-S	-0.7346	0.2340
α^{IA-S}	dispersion parameter of IA-S	36.0139	2.1180
-logL	negative value of the log likelihood function	586570	

Note: Standard errors are obtained using numerical Hessian.

The second contribution of this paper is to identify heterogeneity in consumer attention in regard to quantity surcharge. The current literature focuses on the existence of inattentive consumers who are unaware of quantity surcharge and purchase large size items during quantity surcharge weeks. However, taking advantage of the rich household panel data available for my analysis, I also identify attentive consumers, who actively respond to quantity surcharge and purchase multiple small size jars instead of a large one.

The last contribution of this paper is to develop a simple dynamic demand model that can capture the heterogeneity in consumer attention. There are four components of the model. First, the model is dynamic, as peanut butter is storable. Second, the model considers a multiple discrete choice in terms of quantity. Third, the model separates a purchase quantity decision and a package size choice. The fourth component is two-dimensional heterogeneity in consumers: attention and storability. In each period, consumers decide consumption quantity, purchase quantity, and the number of jars of each size to purchase. The second and the third components together allow to generate multiple jar purchases that are frequently observed in the data. Also, the third and fourth components leave room for inattentive type consumers to make inferior package size choices.

I impose each component to the model in the following way. First, I assume four types of consumers: attentive non-storer, inattentive non-storer, attentive storer, and inattentive storer. In order to handle multiple discreteness, I consider consumption quantity as a countable variable that follows a negative binomial distribution. Consumption quantity and purchase quantity are the same for non-storing type consumers, but not necessarily the same for storing type consumers. As for a storing type consumers' dynamic problem, I introduce the concept of effective price and few additional assumptions to simplify the state space. For the given purchase quantity, consumers decide how many jars of each size to purchase. I take a behavioral approach and assume that an attentive type consumer chooses a package bundle that minimizes expenditure, and an inattentive type consumer chooses a package that minimizes the total number of jars to purchase.

I drive a log likelihood function using the distributional assumption on consumption quantity, and then estimate the parameters using maximum likelihood. As household types are unobservable from the data, I estimate the partition of households minimizing the difference within a group in terms of the two characteristics of purchasing patterns. The estimation results suggest that storer type households are more sensitive to price than non-storer type households. Among the two storer types, attentive storer type households are more price sensitive than inattentive storer types.

The interpretation of the estimation results is limited, as the model assumes a single price. In order to estimate demand more accurately and discuss the consumer welfare effect of quantity surcharge, the model needs to be extended to multiple products.

A Appendix: Data

A.1 Data Set Descriptions

In this paper, I use five different data sets from IRI scanner data from 2008 to 2010 for the analysis. The first data set is product attributes data recorded at UPC level. The product attributes data originally contain 1,018 UPCs of peanut butter and peanut butter related products. For each UPC listed, information on its parent company, brand, product type, texture, flavor, and few other product characteristics is provided. As for product type, approximately 85% of UPCs belong to peanut butter and the rest belong to five other peanut butter related product types, such as peanut butter combo and peanut butter spread. There are 17 different textures and, ranging from super chunky and chunky to smooth and creamy.

The second data set is store level data containing information on weekly sales at UPC level at each store in Eau Claire. The available variables are store ID assigned by IRI, week, UPC, number of jars sold, total sales in dollars, and promotional activities such as feature and display. There are three types of stores in the market: two mass, three drug, and six grocery stores.¹² I focus on grocery stores only for the following reasons: 1) household level store visit observations from mass stores are missing; 2) drug stores have minimal sales volume compared to grocery stores¹³; 3) grocery stores also have somewhat different pricing strategies and promotional activities than grocery stores. Dropping observations from mass stores and drug stores, the data contain six stores, 31 brands, 159 UPCs, and 42,634 sales observations. In the rest of paper, I call grocery stores as stores.

The rest of three data sets are household level: household trip data, household demographic characteristics data, and household peanut butter purchase data. Household trip data contain records of trips to stores each household made for the period 2008-2010. For each household I can tell the store visited, the checkout time, and the total dollar amounts spent on the trip. This trip data originally contain 5,727 households and 719,711 trip observations. There are two types of stores, grocery stores and drug stores. Only 34,784 out of 719,711 trip observations are from drug stores, so I focus on the observations from grocery stores only. There are 5,702 households and 684,927 trip observations left after dropping the ob-

¹²One of the three drug stores opened in 2009

¹³The sum of sales in dollar amount of all the drug stores during the three year time period in Eau Claire takes roughly 1% of the total market sales.

servations from drug stores. Not every household has complete three-year trip records: some have trip observations from one year and no record for the next year. Those households are not appropriate for the analysis to understand household purchase behaviors with dynamic inventory holdings, so I keep the households with at least one trip observation each year. As a result, 4,076 households and 680,851 trip observations remain.

Next, household demographic characteristics data provide information on household income, family size, age and education level of household head, and few other characteristics. For the period 2008-2010, the demographic information was collected in Summer 2012, so the demographic characteristics for each household stay the same over the three years I analyze. There are 2,994 households listed on the data with at least one year observation. Using the fact that the demographic characteristics do not vary over the three years, I filled up the missing observations.

The last and the most important household level data are household purchase data. The data include the complete peanut butter product purchase records of the household panels during the time period analyzed: UPC of the peanut butter product and the number of jars purchased, dollar amount paid, and the store, week, and minute where and when the purchase occurred. Each observation is at household-store-week and minute-UPC level. Thus, if a household visited a store at a certain time and purchased two peanut butter jars with different UPCs, this purchase event is recorded as two separate purchase observations. However, if a household purchased two jars of the same peanut butter product (same UPC), then this purchase event is recorded as a single purchase observation. There are initially 2,713 households with 27,838 purchase observations in the data. Those purchases occurred at three different types of stores: mass, drug, and grocery stores. However, the purchase observations from mass stores and drug stores are very minimal.¹⁴ Due to this negligible number of observations, and also in order to maintain the coherence with the other data sets, I keep the purchase observations from grocery stores only. This leaves 2,713 households and 27,661 purchase observations.

¹⁴Only 149 and 28 purchase observations occurred at drug stores and mass stores, respectively.

A.2 Merging Data Sets

I merged the five different data sets into one data set in the following order: 1) household trip data and household demographic characteristics data; 2) household peanut butter purchase data and peanut butter product attributes data; 3) results of merge 1 and merge 2; 4) result of merge 3 and store peanut butter sales data. In the process of merging, I dropped some observations that were not matched from one to another. Here are some details of each merge.

Merge 1: First, I merged the household demographic characteristics data to the household trip data. I dropped 88 households from the demographic characteristics data who do not have any matching trip observation in the trip data. Also 1,170 households (with 53,625 corresponding trip observations) from the household trip data have no demographic characteristics information available, so I dropped all of their trip observations. In addition, one household (with 45 trip observations) has most of demographic characteristics information available but no income information. Household income is one of the most important characteristics, so I dropped all of the household's trip observations. As a result 2,905 households and 605,688 trip observations are left.

Merge 2: The second step is merging the household product purchase data and the product attributes data. All the purchase observations from the household purchase data are successfully matched to the UPCs listed in the product attributes data. Households purchased four different kinds of product type: peanut butter, peanut butter combo, peanut butter spread, and peanut spread. I kept purchase observations of peanut butter and peanut butter spread types only. The reason is as follows: peanut butter combo is a type of products that has peanut butter and jelly or peanut butter and chocolate spread together in one jar. Peanut spread is usually flavored with honey or chocolate. However, peanut butter spread has no difference in observable characteristics than peanut butter. Also, the number of peanut butter jars that the household purchased during the time period analyzed, peanut butter combo and peanut spread product types together take a small share.¹⁵ After dropping the purchase observations of those two product types, the same number of households remain, but the number of purchase observations are reduced to 27,527.

¹⁵Households purchased 42,425 jars in total, and among those, 125 jars are peanut butter combo type and 20 jars belong to peanut spread type.

Merge 3: The third step is to combine the two merged data sets, the household trip and demographic characteristics data set merged in step 1, and the household peanut butter purchase and peanut butter product attributes data set merged in step 2. I dropped 1,229 purchase observations with no matching trip observations. This yields 2,905 households in total with 606,833 trip observations. Among 2,905 households, 2,638 households bought at least one peanut butter jar during the time period analyzed. The total number of purchase observations is 26,298, and they include 28 brands and 129 UPCs. However, there is an additional grocery store in this merged data set other than the six grocery stores listed in the store level peanut butter sales data. 22 households visited the additional store during the time period analyzed, and the households made 54 trips and 2 peanut butter product purchases. In order to make it compatible to the store level peanut butter sales data, I dropped those 22 households. 2,883 households with 601,114 trip observations remain. Among them 2,616 households purchased at least one peanut butter product, and these households made 26,024 purchase observations in total.

Merge 4: The last step is to merge the store level peanut butter sales data and the combined household level data obtained by the merge 3. The final data set contains 42,634 sales observations at 6 different stores in the Eau Claire market for the period 2008-2010, including 31 brands and 159 UPCs. There are 2,883 households in total with 601,114 trip observations. Among them, 2,616 households purchased peanut butter products at least once, and those households made 26,024 purchase observations in total. Those purchase observations include 27 brands and 128 UPCs.

A.3 Data Summary Statistics

There are six stores in the Eau Claire market, and those stores carry 31 brands and 159 UPCs in total. Table 16 shows the top 30 selling UPCs. The best selling UPC for the period 2008-2010 is Skippy's 16.3 oz peanut butter with creamy texture, and Jif's 18 oz peanut butter with creamy texture is a close runner-up. Overall we can see that households prefer regular size jars to large or super size jars, and creamy texture to the chunky.¹⁶ Another interesting pattern is that regular size UPCs tend to be less expensive per ounce than large size UPCs on average. For example, the rank 2 UPC, Jif's 18 oz creamy peanut butter, has average

¹⁶Here I group peanut butter UPCs to four sizes, small, regular, large, and super, for convenience. The small size jar ranges from 12 oz. to 15 oz, and the regular size jar is between 16 oz. and 18 oz. The large size jar varies from 26 oz. to 28 oz. and anything larger than 36 oz. is classified as the super size.

Table 16: Top 30 Selling UPCs at Store Level

Rank	Brand	oz	Texture	Sales (\$)	Num of Jars	Market Share (%)	Ave Jar Price (\$)	Ave Volume Price (\$)
1	Skippy	16.3	CR	115,737	64,601	12.59	1.79	1.76
2	Jif	18.0	CR	124,288	59,353	11.56	2.09	1.86
3	Private Label	18.0	CR	43,418	28,130	5.48	1.54	1.37
4	Jif	28.0	CR	86,029	23,991	4.67	3.59	2.05
5	Skippy	16.3	CH	39,636	22,869	4.46	1.73	1.70
6	Peter Pan	16.3	CR	29,750	21,986	4.28	1.35	1.33
7	Skippy	16.3	CR	32,655	16,249	3.17	2.01	1.97
8	Jif	18.0	CR	33,505	15,364	2.99	2.18	1.94
9	Jif	18.0	CH	29,498	14,659	2.86	2.01	1.79
10	Skippy	28.0	CR	53,184	13,940	2.72	3.82	2.18
11	Smuckers	16.0	CR	32,631	12,915	2.52	2.53	2.53
12	Jif	40.0	CR	58,019	11,483	2.24	5.05	2.02
13	Smuckers	16.0	CH	25,063	9,923	1.93	2.53	2.53
14	Peter Pan	18.0	CR	17,349	9,775	1.90	1.77	1.58
15	Private Label	18.0	CH	12,934	8,340	1.63	1.55	1.38
16	Skippy Natural	15.0	CR	16,259	7,370	1.44	2.21	2.35
17	Private Label	18.0	CR	9,914	6,215	1.21	1.60	1.42
18	Private Label	40.0	CR	24,690	6,009	1.17	4.11	1.64
19	Skippy	16.3	CH	12,332	5,939	1.16	2.08	2.04
20	Simply Jif	17.3	CR	13,465	5,727	1.12	2.35	2.17
21	Jif	28.0	CH	20,447	5,666	1.10	3.61	2.06
22	Jif	40.0	CR	24,291	5,655	1.10	4.30	1.72
23	Private Label	28.0	CR	15,253	5,519	1.08	2.76	1.58
24	Skippy Super Chunk	28.0	CH	20,918	5,428	1.06	3.85	2.20
25	Skippy Natural	15.0	CH	11,554	5,185	1.01	2.23	2.38
26	Skippy	40.0	CR	26,028	5,010	0.98	5.20	2.08
27	Private Label	18.0	CR	8,545	4,607	0.90	1.85	1.65
28	Jif	18.0	CH	10,272	4,532	0.88	2.27	2.01
29	Jif	40.0	CH	21,368	4,278	0.83	4.99	2.00
30	Skippy Natural	26.5	CR	16,876	4,133	0.81	4.08	2.47

Note: The data used are store level peanut butter product sales data from six grocery stores in the Eau Claire market, 2008-2010. Rank and market share are based on the total number of jars sold for the three years. Texture is either creamy (CR) or chunky (CH). Sales is the sum of dollar values of peanut butter jars sold in the market. Ave jar price is the average price per jar and ave volume price is average price per 16 oz. across weeks and stores.

volume price \$1.86, which is lower than \$2.05, average volume price of Jif's 28 oz creamy peanut butter. This pattern indirectly suggests quantity surcharges.

Table 17 shows the summary statistics of demographic characteristics of households and trips to stores those households made for the period 2008-2010 in the Eau Claire market. The upper panel shows that the median income household earned annual income between \$35,000 and \$44,999. The family sizes are relatively small, as the median household consists of two family members and more than 75% of the households have no children. The lower panel shows the summary statistics of trips households to stores. The median number of trips is 179, which is more than once a week on average, considering that there are 156 weeks in three years. The majority of households visited multiple stores during the time period analyzed, that 75% of households visited greater than equal to three different stores and 50% of households visited greater than equal to four different stores. The median average number of trips in a week is above 1, and the median of the average number of weeks since last trip is less than a week. The two numbers together suggest that roughly 50% of households made regular weekly trips to stores. The median of the average expenditure is approximately \$37 per trip and \$62 per week.

Among the total number of 601,114 trip observations, 24,894 observations include peanut butter product purchases. The top panel in the table 18 shows the summary statistics of households' peanut butter purchases on a single trip. The average number of jars purchased is 1.61, which is significantly larger than 1, while the average number of UPCs and the average number of brands purchased are close to 1. These suggest that the households frequently purchased multiple jars at a single trip, and most of those cases were exactly the same UPCs, rather than purchasing multiple UPCs.

The middle panel in the table 18 shows the summary statistics when the purchases are combined in a weekly level. In only few cases the households purchased peanut butter products through multiple trips in a week. Hence, the summary statistics at a weekly level are not far from those of a trip level presented in the top level. During the time period analyzed, 2,616 households purchased peanut butter products at least once, and the summary statistics of those households' purchases are presented in the bottom panel. Roughly half of households purchased more than 10 jars, and more than 190 oz. in terms of volume, conditional on purchase. Even though households usually purchased a single UPC on a trip, they sought for varieties in the long term and purchased multiple UPCs and brands during the time

Table 17: Summary Statistics of Demographic Characteristics and Trips

Variable	Obs	Mean	SD	Min	P25	P50	P75	Max
Income	2,883	7.22	3.23	1	5	7	10	12
Fam Size	2,883	2.35	1.23	1	2	2	3	8
Num of Children	2,883	0.26	0.58	0	0	0	0	3
HH Age	2,870	4.78	1.17	1	4	5	6	6
HH Educ	2,870	6.79	15.57	1	3	4	5	99
Race	2,883	1.50	6.57	1	1	1	1	99
Num of Trips	2,883	208.50	133.08	5	122	179	261	1,927
Num of Stores	2,883	4.14	1.40	1	3	4	5	6
Ave Num of Trips	2,883	1.78	0.72	1.00	1.34	1.59	1.99	12.35
Ave Num of Weeks								
Since Last Trip	2,883	0.99	0.80	0.08	0.59	0.84	1.21	15.71
Ave Expend (\$)	2,883	41.48	22.81	5.34	24.99	36.97	53.00	233.25
Ave Weekly Expend (\$)	2,883	67.53	32.14	10.34	44.19	62.40	85.06	387.37

Note: The data used are household level panel data from the Eau Claire market, 2008-2010. The upper panel shows the summary of demographic information of households and the lower panel shows the summary information of the trips households made from six grocery stores in the market. Income is a categorical variable representing the combined pre-tax income of the head of household (HH). Income is equal to 1 if the combined pre-tax income of HH is in the range of \$00,000 to \$9,999 per year, 2 if in the range of \$10,000 to \$11,999, 3 if in the range of \$12,000 to \$14,999, 4 if in the range of \$15,000 to \$19,999, 5 if in the range of \$20,000 to \$24,999, 6 if in the range of \$25,000 to \$34,999, 7 if in the range of \$35,000 to \$44,999, 8 if in the range of \$45,000 to \$54,999, 9 if in the range of \$55,000 to \$64,999, 10 if in the range of \$65,000 to \$74,999, 11 if in the range of \$75,000 to \$99,999, and 12 if greater than or equal to \$100,000. Fam size and num of children represent the number of family members and the number of children in the household, respectively. HH age is a categorical variable representing the age of HH. HH age is equal to 1 if the HH's age lies in the range of 18 to 24, 2 if 25 to 34, 3 if 35 to 44, 4 if 45 to 54, 5 if 55 to 64, and 6 if greater than equal to 65. Race is a categorical variable representing the HH's ethnicity. Race is equal to 1 if the HH is white and 96.94% of 2,883 households in the data have white HHs.

period analyzed.

B Appendix: Identification

I first introduce the identification of the standard negative binomial maximum likelihood estimator, and then show how the estimator maximizing the log likelihood function in equation (5.5) is related to the standard negative binomial maximum likelihood estimator.

Suppose q follows a negative binomial distribution with density

$$(B.1) \quad h(q|\mu, \alpha) = \frac{\Gamma(q + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(q + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}} \right)^q,$$

Table 18: Summary Statistics of Household Peanut Butter Purchases

Variable	Obs	Mean	SD	Min	P25	P50	P75	Max
On a Trip								
Num of Jars	24,894	1.61	0.89	1	1	1	2	16
Volume (oz)	24,894	31.76	19.12	12	18	33	36	336
Num of UPCs	24,894	1.05	0.21	1	1	1	1	3
Num of Brands	24,894	1.01	0.12	1	1	1	1	3
In a Week								
Num of Jars	23,502	1.70	1.08	1	1	1	2	18
Volume (oz)	23,502	33.65	21.74	12	18	33	36	336
Num of UPCs	23,502	1.07	0.28	1	1	1	1	6
Num of Brands	23,502	1.03	0.17	1	1	1	1	4
For Three Years								
Num of Jars	2,616	15.29	17.09	1	4	10	20	166
Volume (oz)	2,616	302.27	343.42	12	81	189	388	2,977
Num of UPCs	2,616	4.44	3.18	1	2	4	6	25
Num of Brands	2,616	2.62	1.39	1	1	2	4	10

Note: The data used are household level panel data from the Eau Claire market, 2008-2010. There are three groups of summary statistics of peanut butter purchases the household panels made, conditional on purchase: 1) on a single trip; 2) in a week (possibly multiple trips combined); 3) for the three years combined.

where $\alpha \geq 0$, $q = 0, 1, 2, \dots$, and $\mu = \exp(x'\beta)$. [Cameron and Trivedi \(2013\)](#) show that $\Gamma(q + \alpha^{-1})/\Gamma(\alpha^{-1}) = \prod_{j=0}^{q-1} (j + \alpha^{-1})$ when q is an integer, which means

$$(B.2) \quad \log\left(\frac{\Gamma(q + \alpha^{-1})}{\Gamma(\alpha^{-1})}\right) = \sum_{j=0}^{q-1} \log(j + \alpha^{-1}).$$

Then the log likelihood function can be written as

$$(B.3) \quad \begin{aligned} \log L(\alpha, \beta) &= \sum_{i=1}^n \log h(q_i | \mu_i, \alpha) \\ &= \sum_{i=1}^n \left\{ \left(\sum_{j=0}^{q_i-1} \log(j + \alpha^{-1}) \right) - \log q_i! \right. \\ &\quad \left. - (q_i + \alpha^{-1}) \log(1 + \alpha \exp(x_i'\beta)) + q_i \log(\alpha) + q_i x_i'\beta \right\}. \end{aligned}$$

The negative binomial maximum likelihood estimator $(\hat{\alpha}, \hat{\beta})$ is the solution to the first-order conditions

$$(B.4) \quad \sum_{i=1}^n \frac{q_i - \mu_i}{1 + \alpha \mu_i} x_i = 0,$$

$$\sum_{i=1}^n \left\{ \frac{1}{\alpha^2} \left(\log(1 + \alpha \mu_i) - \sum_{j=0}^{q_i-1} \frac{1}{(j + \alpha^{-1})} \right) + \frac{q_i - \mu_i}{\alpha(1 + \alpha \mu_i)} \right\} = 0.$$

Now I show the identification of the parameters (α^h, β^h) for the two non-storer types ($h = A-NS, IA-NS$). The log likelihood function is given in equation (5.5). Note that the probability of jar size choice, ϕ_{it}^h , does not depend on the parameters (α^h, β^h) . Let $x'_{it} = (1, \underline{p}_{ts})'$. The maximum likelihood estimator $(\hat{\alpha}^h, \hat{\beta}^h)$ for the two non-storer types is the solution to the first order conditions

$$(B.5) \quad \sum_{t=1}^R \sum_{i=1}^{N^h} \frac{q_{it}^h - \mu_{it}^h}{1 + \alpha^h \mu_{it}^h} x_{it} = 0,$$

$$\sum_{t=1}^R \sum_{i=1}^{N^h} \left\{ \frac{1}{(\alpha^h)^2} \left(\log(1 + \alpha^h \mu_{it}^h) - \sum_{j=0}^{q_{it}^h-1} \frac{1}{(j + (\alpha^h)^{-1})} \right) + \frac{q_{it}^h - \mu_{it}^h}{\alpha^h(1 + \alpha^h \mu_{it}^h)} \right\} = 0,$$

which is the negative binomial maximum likelihood estimator.

Next I show the identification of the parameters (α^h, β^h) for the two storer types ($h = A-S, IA-S$). The log likelihood function is given in equation (5.5). The first order condition with respect to β^h is

$$(B.6) \quad \frac{\partial \log L^h(\alpha^h, \beta^h)}{\partial \beta^h}$$

$$= \sum_{t=1}^R \sum_{i=1}^{N^h} \left\{ (1 - \mathbf{1}(SN) - \mathbf{1}(NS)) \frac{\partial \log h(q_{it}^h | \mu_{it}^h, \alpha^h)}{\partial \beta^h} + \mathbf{1}(SN) \frac{\partial \log(\mathbf{1}(q_{it}^h = 0))}{\partial \beta^h} + \mathbf{1}(NS) \frac{\partial \log D}{\partial \beta^h} \right\} = 0,$$

where $D = \sum_{j=0}^{q_{it}^h} h(j | \mu_{it}^h, \alpha^h) h(q_{it}^h - j | \mu_{it}^h, \alpha^h).$

The three partial derivatives in the equation (B.6) are

$$\begin{aligned}
(B.7) \quad \frac{\partial \log h(q_{it}^h | \mu_{it}^h, \alpha^h)}{\partial \beta^h} &= \frac{q_{it}^h - \mu_{it}^h}{1 + \alpha^h \mu_{it}^h} x_{it}, \\
\frac{\partial \log(\mathbf{1}(q_{it}^h = 0))}{\partial \beta^h} &= 0, \\
\frac{\partial \log D}{\partial \beta^h} &= \frac{1}{D} \frac{\partial D}{\partial \beta^h} = \frac{q_{it}^h - 2\mu_{it}^h}{1 + \alpha^h \mu_{it}^h} x_{it}.
\end{aligned}$$

Plugging the equation (B.7) to equation (B.6) yields

$$\begin{aligned}
(B.8) \quad \sum_{t=1}^R \sum_{i=1}^{N^h} \left\{ (1 - \mathbf{1}(SN) - \mathbf{1}(NS)) \left(\frac{q_{it}^h - \mu_{it}^h}{1 + \alpha^h \mu_{it}^h} x_{it} \right) + \mathbf{1}(NS) \left(\frac{q_{it}^h - 2\mu_{it}^h}{1 + \alpha^h \mu_{it}^h} x_{it} \right) \right\} &= 0 \\
\Rightarrow \sum_{t=1}^R \sum_{i=1}^{N^h} \left(\frac{(1 - \mathbf{1}(SN)) q_{it}^h - (1 - \mathbf{1}(SN)) + \mathbf{1}(NS) \mu_{it}^h}{1 + \alpha^h \mu_{it}^h} \right) x_{it} &= 0.
\end{aligned}$$

The first order condition with respect to α^h is

$$\begin{aligned}
(B.9) \quad \frac{\partial \log L^h(\alpha^h, \beta^h)}{\partial \alpha^h} &= \sum_{t=1}^R \sum_{i=1}^{N^h} \left\{ (1 - \mathbf{1}(SN) - \mathbf{1}(NS)) \frac{\partial \log h(q_{it}^h | \mu_{it}^h, \alpha^h)}{\partial \alpha^h} + \mathbf{1}(SN) \frac{\partial \log(\mathbf{1}(q_{it}^h = 0))}{\partial \alpha^h} + \mathbf{1}(NS) \frac{\partial \log D}{\partial \alpha^h} \right\} = 0, \\
\text{where } D &= \sum_{l=0}^{q_{it}^h} h(l | \mu_{it}^h, \alpha^h) h(q_{it}^h - l | \mu_{it}^h, \alpha^h).
\end{aligned}$$

The three partial derivatives in equation (B.9) are

$$\begin{aligned}
(B.10) \quad \frac{\partial \log h(q_{it}^h | \mu_{it}^h, \alpha^h)}{\partial \alpha^h} &= \left\{ \frac{1}{(\alpha^h)^2} \left(\log(1 + \alpha^h \mu_{it}^h) - \sum_{j=0}^{q_{it}^h-1} \frac{1}{(j + (\alpha^h)^{-1})} \right) + \frac{q_{it}^h - \mu_{it}^h}{\alpha^h (1 + \alpha^h \mu_{it}^h)} \right\}, \\
\frac{\partial \log(\mathbf{1}(q_{it}^h = 0))}{\partial \alpha^h} &= 0, \\
\frac{\partial \log D}{\partial \alpha^h} &= \frac{1}{D} \frac{\partial D}{\partial \alpha^h},
\end{aligned}$$

where the last term can be written as

$$\begin{aligned}
\text{(B.11)} \quad \frac{\partial D}{\partial \alpha^h} &= \sum_{l=0}^{q_{it}^h} \left\{ h(l|\mu_{it}^h, \alpha^h) \frac{\partial h(q_{it}^h - l|\mu_{it}^h, \alpha^h)}{\partial \alpha^h} + \frac{\partial h(l|\mu_{it}^h, \alpha^h)}{\partial \alpha^h} h(q_{it}^h - l|\mu_{it}^h, \alpha^h) \right\} \\
&= \left(\sum_{l=0}^{q_{it}^h} h(l|\mu_{it}^h, \alpha^h) h(q_{it}^h - l|\mu_{it}^h, \alpha^h) \right) \times \left(\frac{1}{(\alpha^h)^2} 2\log(1 + \alpha^h \mu_{it}^h) + \frac{q_{it}^h - 2\mu_{it}^h}{\alpha^h(1 + \alpha^h \mu_{it}^h)} \right) \\
&\quad - \sum_{l=0}^{q_{it}^h} \left\{ h(l|\mu_{it}^h, \alpha^h) h(q_{it}^h - l|\mu_{it}^h, \alpha^h) \frac{1}{(\alpha^h)^2} \left(\sum_{j=0}^{q_{it}^h - l - 1} \frac{1}{(j + (\alpha^h)^{-1})} + \sum_{j=0}^{l-1} \frac{1}{(j + (\alpha^h)^{-1})} \right) \right\} \\
\Rightarrow \frac{\partial \log D}{\partial \alpha^h} &= \left(\frac{1}{(\alpha^h)^2} 2\log(1 + \alpha^h \mu_{it}^h) + \frac{q_{it}^h - 2\mu_{it}^h}{\alpha^h(1 + \alpha^h \mu_{it}^h)} \right) \\
&\quad - \frac{\sum_{l=0}^{q_{it}^h} \left\{ h(l|\mu_{it}^h, \alpha^h) h(q_{it}^h - l|\mu_{it}^h, \alpha^h) \frac{1}{(\alpha^h)^2} \left(\sum_{j=0}^{q_{it}^h - l - 1} \frac{1}{(j + (\alpha^h)^{-1})} + \sum_{j=0}^{l-1} \frac{1}{(j + (\alpha^h)^{-1})} \right) \right\}}{\sum_{l=0}^{q_{it}^h} h(l|\mu_{it}^h, \alpha^h) h(q_{it}^h - l|\mu_{it}^h, \alpha^h)}.
\end{aligned}$$

Thus

$$\begin{aligned}
\text{(B.12)} \quad \sum_{t=1}^R \sum_{i=1}^{N^h} &\left\{ (1 - \mathbf{1}(SN) - \mathbf{1}(NS)) \left[\frac{1}{(\alpha^h)^2} \left(\log(1 + \alpha^h \mu_{it}^h) - \sum_{j=0}^{q_{it}^h - 1} \frac{1}{(j + (\alpha^h)^{-1})} \right) + \frac{q_{it}^h - \mu_{it}^h}{\alpha^h(1 + \alpha^h \mu_{it}^h)} \right] \right. \\
&\quad + \mathbf{1}(NS) \left[\left(\frac{1}{(\alpha^h)^2} 2\log(1 + \alpha^h \mu_{it}^h) + \frac{q_{it}^h - 2\mu_{it}^h}{\alpha^h(1 + \alpha^h \mu_{it}^h)} \right) \right. \\
&\quad \left. \left. - \frac{\sum_{l=0}^{q_{it}^h} \left\{ h(l|\mu_{it}^h, \alpha^h) h(q_{it}^h - l|\mu_{it}^h, \alpha^h) \frac{1}{(\alpha^h)^2} \left(\sum_{j=0}^{q_{it}^h - l - 1} \frac{1}{(j + (\alpha^h)^{-1})} + \sum_{j=0}^{l-1} \frac{1}{(j + (\alpha^h)^{-1})} \right) \right\}}{\sum_{l=0}^{q_{it}^h} h(l|\mu_{it}^h, \alpha^h) h(q_{it}^h - l|\mu_{it}^h, \alpha^h)} \right] \right\} = 0.
\end{aligned}$$

The maximum likelihood estimator $(\hat{\alpha}^h, \hat{\beta}^h)$ for the two storer types is the solution to the first order conditions in equation (B.8) and equation (B.12).

References

- Jagdish Agrawal, Pamela E Grimm, and Narasimhan Srinivasan. Quantity surcharges on groceries. *Journal of consumer affairs*, 27(2):335–356, 1993.
- Greg M Allenby, Thomas S Shively, Sha Yang, and Mark J Garratt. A choice model for packaged goods: Dealing with discrete quantities and quantity discounts. *Marketing Science*, 23(1):95–108, 2004.
- A Colin Cameron and Pravin K Trivedi. *Regression analysis of count data*, volume 53. Cambridge university press, 2013.
- Sofronis Clerides and Pascal Courty. Sales, quantity surcharge, and consumer inattention. 2015.
- Christopher T Conlon and Julie Holland Mortimer. An experimental approach to merger evaluation. Technical report, National Bureau of Economic Research, 2013.
- Jean-Pierre Dubé. Product differentiation and mergers in the carbonated soft drink industry. *Journal of Economics & Management Strategy*, 14(4):879–904, 2005.
- Igal Hendel. Estimating multiple-discrete choice models: An application to computerization returns. *The Review of Economic Studies*, 66(2):423–446, 1999.
- Igal Hendel and Aviv Nevo. The post-promotion dip puzzle: What do the data have to say? *Quantitative Marketing and Economics*, 1(4):409–424, 2003.
- Igal Hendel and Aviv Nevo. Measuring the implications of sales and consumer inventory behavior. *Econometrica*, 74(6):1637–1673, 2006.
- Igal Hendel and Aviv Nevo. Intertemporal price discrimination in storable goods markets. *American Economic Review*, 103(7):2722–51, 2013.
- Jaehwan Kim, Greg M Allenby, and Peter E Rossi. Modeling consumer demand for variety. *Marketing Science*, 21(3):229–250, 2002.
- Maryam Saeedi. Reputation and adverse selection, theory and evidence from ebay. *Theory and Evidence from eBay*, 2014.

David E Sprott, Kenneth C Manning, and Anthony D Miyazaki. Grocery price setting and quantity surcharges. *Journal of Marketing*, 67(3):34–46, 2003.

Stanley M Widrick. Measurement of incidents of quantity surcharge among selected grocery products. *Journal of Consumer Affairs*, 13(1):99–107, 1979.

List of Tables

1	Top 10 Selling Brands	8
2	Household Multi-Jar and Multi-UPC Purchases at a Trip	8
3	Evidence of Stockpiling: Average Number of Jars Sold	10
4	Evidence of Stockpiling: Store Level	11
5	Evidence of Stockpiling: Household Level	12
6	Three Products Identified	14
7	Quantity Discount and Quantity Surcharge at Grocery Stores	15
8	Quantity Surcharges and Multiple Jar Purchases	19
9	Missing Substitution Opportunities in QS weeks: Purchasing Patterns	21
10	Missing Substitution Opportunities in QS Weeks at Grocery Stores: Demographic Info	23
11	Jar Size Choices of Four Household Types, for Given Purchase Quantity	31
12	Results of Household Type Assignments	36
13	ML Estimation Results: Non-Storer Types Only	38
14	ML Estimation Results: Storer Types Only	39
15	ML Estimation Results: All Four Types	40
16	Top 30 Selling UPCs at Store Level	46
17	Summary Statistics of Demographic Characteristics and Trips	48
18	Summary Statistics of Household Peanut Butter Purchases	49

List of Figures

1	Volume Price of Three Skippy Creamy PB UPCs at a Store	16
2	Histogram of Quantity Purchased at Household-Trip Level	26