Bundling to save: Analyzing package size choices in South African grocery stores

Andrea Szabó
Department of Economics
University of Houston
aszabo2@uh.edu

October 22, 2024

Abstract

Storable goods such as laundry detergent come in different package sizes with different associated unit prices. Buying larger packages is an opportunity to save, but low-income consumers in African countries often appear to forego this opportunity and buy small packages instead. I investigate the determinants of these choices by estimating a model of dynamic consumer demand using scanner data from all stores of South Africa's leading grocery chain. The estimation accounts for "bundling:" due to temporary sales and non-linear pricing of the product, consumers sometimes find it less expensive to purchase multiple small packages instead of a large package. The results show that this phenomenon is quantitatively important in explaining observed patterns in the data. Counterfactual simulations use the model's findings to evaluate the impact of offering different package sizes, which is a relevant consideration for the current expansion of small-format chain stores to low-income areas.

1 Introduction

Storable goods like laundry detergent, bathroom tissue, or non-perishable food items are often sold in the same store in widely different package sizes. The unit price for these goods is typically lower if the consumer buys a larger package. Thus, buying larger packages is a form of saving. In developing countries, low income consumers often appear to forego this opportunity to save and instead buy smaller packages. While buying small packages can have some benefits, if smaller packages have a larger unit price, then households overspend on these storable goods. In this paper, I use a unique dataset on laundry detergents from South Africa to estimate a dynamic demand model that allows quantifying the determinants of these choices. Assuming that a store can carry only one size of laundry detergent, what size would a social planner choose? This question has current relevance in the African context, where grocery store chains are expanding in low-income informal neighborhoods. Traditionally, these areas were served only by a collection of small informal stores, and high travel costs meant that consumers had limited access to formal supermarkets. Now, formal chains are entering these neighborhoods by opening small square-footage stores that do not carry the full selection of package sizes available in their larger stores. Since laundry detergent is a one of the most expensive items a household typically buys in a grocery store, understanding the impact of which size such a store chooses to carry can have large welfare implications.

Studying these questions requires high-frequency purchase data, including market shares and prices, for an entire product category, linked to demographic characteristics of the consumers. While readily available scanner datasets with these features exist for the US, I am not aware of another such dataset from a developing country.

Even with such data, estimating consumers' dynamic decisions about a storable good faces a number of challenges. First, the product inherently features non-linear pricing and, in addition, temporary promotions that often target only specific sizes can substantially change unit prices. These two facts combined can result in situations where consumers are better off buying two smaller packages instead of a larger one. As I show below, ignoring the presence of these "bundling opportunities" can lead to very misleading conclusions regarding consumers' preferences for small package sizes in developing countries.

Second, the identification of dynamic demand for storable goods faces the typical difficulty that even if purchases are observed, individual households' consumption and inventory are not. In this paper I alleviate this issue by fielding an original survey that directly collects information on consumption and inventory (as well as other aspects of detergent use) from a sample of households in my study area.

Cooperating with Unilever South Africa, I have obtained a unique dataset of supermarket scanner data with country-wide coverage for a 16-month period, from July 2011 to October 2012. The data contains monthly information on laundry detergents at the store level from all 330 stores of South Africa's leading grocery chain, Shoprite. In this scanner data, I separately observe all types and brands of Unilever detergents sold. I focus on 3 brands of powdered hand-wash detergent (by far the most common category of laundry detergent sold): Omo, Sunlight, and Sunlight Tropical. These are nationally important brands, comprising 70-85 percent of all detergent sales in Shoprite stores (with a similar market share nationally), and all three are produced by Unilever. These three distinct products are sold in various package sizes, ranging from 250g to 5kg, for a total of 14 brand-size combinations over the period of study.

Using the geo-coded location of the stores, I match the scanner data to household demographics from the South African census. I also use this combined dataset to define local markets for each store. I complement these data with an original survey on detergent use fielded to a sample of consumers in the market area of three selected stores.

I use this dataset to estimate a model of dynamic consumer choice building on Hendel and Nevo (2006). In the model, the consumer chooses which quantity (package size) and brand to purchase, how much to consume, and hence what inventory to hold in a given period.

Applying the estimation strategy in Hendel and Nevo (2006) to my data is not straightforward. First, since I analyze market level data, I estimate the consumer's brand choice conditional on size using BLP. Second, I explicitly take into account the bundling opportunities that arise in some markets when, due to temporary sales, consumers may find it cheaper to purchase multiple small packages instead of a large package. As a result, the econometric model becomes more complex, as these situations violate the discrete choice assumption. I present a method to estimate the share of small packages that were purchased as part of a bundle. This is based on the idea that estimates from markets with no bundling opportunities can be used to infer the counterfactual consumer demand on markets with bundling opportunities if such opportunities were absent.

The estimation of dynamic models with unobserved state variables raises inherent difficulties for the identification of the parameters. Typical applications must address the issue that only households' purchases are observed, while consumption and inventory are unobserved. For example, Hendel and Nevo (2006) simulate an initial distribution of inventories and compute optimal consumption based on the model (see also Erdem, Imai and Keane (2003)).

Here, I use direct information on households' consumption, inventory, and purchase from the survey to help identify the dynamic parameters of the model. Since the solution of the dynamic programming problem requires identifying the dependence between purchased package size, inventory, and consumption, I directly ask households about each of these elements and use this information combined with the scanner data to identify the dynamic parameters.

My first finding is that accounting for bundling opportunities has important effects for the interpretation of observed market shares. I estimate that, on the median market, there are over 3 times as many households who purchase 5 kg of detergent as households who purchase a single 5 kg package. Correspondingly, many of the households buying the smallest package sizes are in fact creating bundles and purchasing a larger total quantity. These findings run against the common belief that due to severe financial constraints, households often purchase products with the lowest package price rather than the lowest unit price. As the case studied here illustrates, looking purely at market shares can be misleading. A consumer might optimally purchase multiple small packages because this bundle is less expensive than a larger package.

Second, I study a counterfactual scenario in which consumers face a decreased fixed cost of purchase (such as lower transportation costs). The results show that the fixed purchase cost is relatively more important in shaping demand compared to the estimated costs of holding inventory at home.

Third, I use the estimated model to evaluate counterfactual scenarios designed to approximate a situation where a small store is constrained to carry only one size of detergent. I simulate consumers' optimal dynamic choices under each of the six package sizes observed in the data, and calculate the expected discounted present value of consumer utility in each case.

In a scenario where a store is restricted to offering a single package size, offering the largest (5 kg) package would provide the highest consumer utility. The 2 kg package yields the second highest consumer utility, and at the same time keeps average inventory (and consumption) closer to what is observed in the data. By contrast, offering the smallest (250 g) package size would deliver one of the lowest values of consumer utility. This is in line with the estimation results which indicate that, in the data, many households only purchased small packages as part of a bundle when buying a larger total quantity. This result may also have implications regarding the most desirable package size offered in grocery chains that are currently expanding in low-income neighborhoods throughout Africa.

This paper continues a line of research on dynamic demand estimation that includes Lal and Rao (1997), Pesendorfer (2002), Erdem, Imai and Keane (2003), Arcidiacono and Miller (2011), Gowrisankaran and Rysman (2012), and Hendel and Nevo (2006, 2013). Unlike earlier studies, I consider a developing country application and explicitly account for the bundling opportunities present in some markets.

This paper also contributes to the development economics literature, where there is long-standing interest in interventions designed to motivate households to save. Even small amounts of savings can contribute to many desirable outcomes, such as enabling households to send their children to school, or to invest in cleaner and healthier cooking equipment. For example, in a series of experiments Karlan et al. (2016) design several interventions to motivate households to make a deposit into their savings accounts each month. In the Philippines (which has the closest mean household income to the South African setting), a door-to-door campaign resulted in 26 USD saved on average. In my data, over a year, a household that consumes the average amount of detergent and always buys the largest package will spend 11.08 USD less than if it always bought the smallest package. These are remarkably high savings on a single product category that typical households regularly purchase in a grocery store. Understanding the determinants of these decisions may thus have important welfare implications.

2 The role of storable goods in developing countries

Commonly studied storable goods in industrial organization, such as laundry detergents, play a special role in developing countries. Most of the time, these are the most expensive items a household buys in a typical grocery store, and in many cases households spend a high fraction of their income on these products. Additionally, these items are difficult to substitute or home-produce. Consequently, almost all households use and purchase these products continuously.

These products are typically sold in many different package sizes and, just as is documented in many studies using US data, they exhibit non-linear pricing. In most cases, the larger package size has a lower unit price, thus purchasing in bulk is a form of saving for households. Because these products are relatively expensive, the potential saving for an average household over a year is significant.

In my data, the average household consumes 1156 grams of laundry detergent per month. Over a year, if a household always buys the smallest package available, it will purchase 55.49 units of 250 gram packages. At the average price for this package size, the household will spend a total of 450.58 Rand. By contrast, if a household purchases only the largest (5kg) package over the year, it will need to purchase 2.77 units, at a total cost of 251.52 Rand using the average price of this package size. This corresponds to a 199.06 Rand or 11.08 USD saving for the household. Using the lowest price of the largest package and the highest price of the smallest package observed in my data, this expenditure gain increases to 23.97 USD per year. This corresponds to 3.1% of the mean or 12.2% of the median monthly household income in South Africa in the same time period. These are remarkably high savings on a single product category purchased by a typical household in a grocery store.

In the development literature, there is long-standing interest in interventions designed to motivate households to save. Even small amounts of savings can contribute to many desirable outcomes, such as enabling households to send their children to school, or to invest in cleaner and healthier cooking equipment. For example, in a series of experiments Karlan et al. (2017) design several interventions to motivate households to make a deposit into their savings accounts each month. Their experiments span 3 countries, and some of them last for 24 months. They find that in the Philippines, which has the closest mean household income to the South African setting studied here, a door-to-door marketing campaign resulted in 26 USD saved on average. In this sense, the potential savings that can be achieved by changing what size products households buy are larger than those produced by a typical information campaign. Consequently, understanding the determinants of how households choose among different package sizes can have widespread implications.

3 Data and background

3.1 Background

In 2011-12, Unilever's share of the laundry detergent market was stable at approximately 80-90 percent (Figure A.1 in the Appendix).[1] Among laundry care products, standard powdered detergent has been the leading category, reaching a value of R8 billion in 2011. The market share of bar soap and other low-cost cleaning agents has been small and declining.

The laundry detergent market has two distinct types of product: hand-wash and automatic detergents. These products do not substitute each-other. Hand-wash detergents are used for washing by hand as well as in low-efficiency washing machines typical in South Africa.[2] Automatic detergents are used in high-efficiency machines and are considered a niche product. Hand-wash detergents account for about 90 percent of all sales during my period of study.

Among hand-wash detergents, Unilever has a 80-95 percent market share (Figure A.2 in the Appendix). The remaining products on this market are either small brands with mostly

[1] The second largest producer, Procter & Gamble, entered the market in 2013, after the end of my data, with up to 10 percent share of the market.

[2] Around 20% of South African households owned washing machines in 2011, most of which are low-efficiency machines. These drain into a bathtub or are used outside the house.

Table 1: Market shares and prices

	Sunlight regular		Sunlight tropical		OMO	
	Market share	Price	Market share	Price	Market share	Price
250g	0.01	7.92	0.01	7.94	0.01	8.46
500g	0.04	15.30	0.01	15.35	0.02	16.66
1kg	0.12	24.66	0.05	24.82	0.06	26.25
2kg	0.24	39.41	0.10	39.50	0.19	40.69
3kg	0.08	55.00				
5kg	0.06	90.66				

Notes: Market share is share of total sales in Rand across the 5,255 markets in the data. Price is average price in Rand across these markets.

specialty detergents (such as detergents sold for baby clothes or dark clothes), or generic store brands. Small brands have at most a 2 percent share of the market, while generic brands have a combined share of at most 4 percent.

3.2 Scanner data

I obtained a unique dataset of supermarket scanner data from all Shoprite stores in South Africa. Shoprite is a South-African supermarket chain that has grown considerably in the past decades and is known for expanding in all areas of the county, including low income neighborhoods. Shoprite's presence in low-income neighborhoods sets it aside from other formal grocery store chains in the country.

The data contains monthly information at the store level based on scanner data. This includes quantity sold and price per package separately for all types and brands of Unilever detergents.

The data includes information on 330 stores over 16 month period, from July 2011 to October 2012. From now on, I refer to a store-month as a "market," and there are a total of 5,255 markets in my data (a few stores were not open during the entire sample period).

Additional information about the laundry detergent market comes from a weekly scanner dataset that contains aggregate information from all Shoprite stores. This data includes price per item and quantity sold for every product, including both Unilever and its competitors. It also includes information on the promotional activities of each of these products, including sales weighted regular price and promotional price, as well as the percentage of volume sold during the promotion.

3.2.1 Products and package sizes in the sample

The sample used in this paper includes 3 powdered hand-wash detergent brands sold by Unilever: Omo, Sunlight, and Sunlight Tropical Scent. This group of products comprises 70-85 percent of all detergent sales across Shoprite stores, and a similar market share nationally.[3] Within this group, Sunlight has 56 percent of the total market share. Omo, which is considered to be a higher-performance detergent with better cleaning ability, has 27 percent. This brand is more popular in urban areas. Sunlight Tropical, which has a market share of 17 percent, is a variation of the Sunlight brand - it is perceived to have a stronger, longer lasting scent. Perhaps because of this feature this is more popular in low-income and rural areas.

These three products are sold in various package sizes, ranging from 250g to 5kg, for a total of 14 brand-size combinations over the period of study. Table 1 shows the summary statistics of market shares for all 14 products in my main sample. The most popular size (both in sales value and volume) is the 2 kg package, which has a 53 percent market share. In contrast to popular belief, packages smaller than 1 kg are not very popular. None of the products below 1kg has a market share larger than 4 percent, and the six products in this group have a combined market share of only 9 percent.[4] The two largest sizes have 14 percent combined market share, despite the fact that these can only be purchased from one brand, Sunlight.

My period of study did not see the introduction of a major new product, and there is no clear trend in the market share of any of the sample products over time (Figure 1). There are very few store-months when a particular product had zero sales (Table A.3 in the Appendix). In these cases I cannot distinguish whether the product was physically in the store, but had no sales for an entire month, or was not available.

3.2.2 Prices, quantity discount and promotional activity

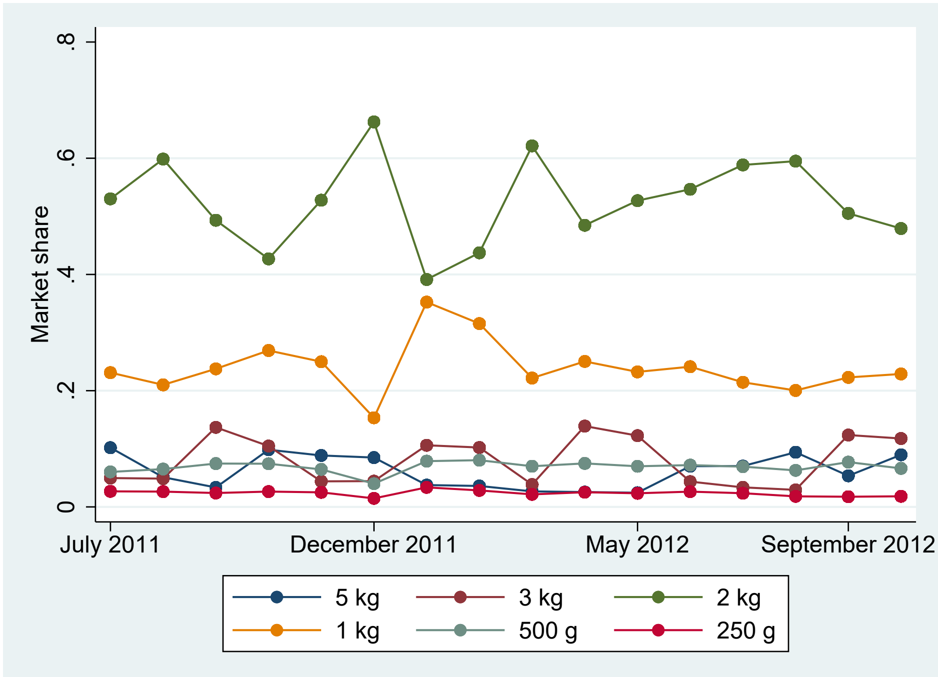
There are three features worth mentioning: (1) prices are not uniform across stores even though all stores belong to the same chain, (2) there is a substantial quantity discount embedded in the prices, and (3) promotions affect different package sizes at different times. I describe each of these features below.

Non-uniform prices across stores from the same chain. The average price of a given package size is similar across brands (Table 1). However, these similar average prices mask significant price variation both across stores and across time. This is in contrast to patterns

[3] The highest market share of other Unilever hand-wash detergents is 3 percent. For non-Unilever hand-wash detergents, it is 1 percent.

[4] In fact of all the *packages* sold, only 22 percent are less than 1kg.

Figure 1: Market share of Unilver products by package size



Notes: Market share is sales value of the given package size across all stores in a given month divided by the sales value of all 14 Unilever products across all stores in the given month.

documented in US data, for example in DellaVigna and Gentzkow (2019). On the top panel of Figure 2, I display the distribution of prices for selected products for a specific month to illustrate price dispersion of the same product across the 330 stores. The lower panel displays the same two product prices across all markets, which shows even more dispersion. Generally, smaller sizes have less price dispersion than larger sizes. Figure A.4 shows the three specific products with the highest market shares in selected stores over the study period. These graphs illustrate the differences between the price per package of the same product among selected stores.

These price differences across stores and over time provides the variation necessary to identify price elasticities.

Non-linear pricing. Table 2 shows the unit price of each package size relative to the 1 kg package of the same brand. The values shown are average across markets, computed using regular (non-promotional) prices. For example, based on the first column, purchasing 1 kg of Sunlight detergent as four 250 g packages is 25% more expensive than purchasing a single 1 kg package. Table 2 indicates that quantity discounts across brands are very similar. In addition, although there are up to 5 different sizes from the same brand, there seem to be only three groups of unit prices. Smaller products are on average 25-30 percent more

Figure 2: Distribution of prices of selected products

Notes: Store level monthly prices in Rand. The upper panels show the price distribution across stores for two products in April 2012 (the middle of the sample period). The lower panels show the corresponding distributions over the entire sample (16 months, 330 stores).

expensive, and larger products are 20-25 percent cheaper compared to a 1 kg package.[5]

There is variation in quantity discounts over time and across stores. To illustrate this, Figure 3 plots the quantity discount of a 2 kg package versus buying two 1 kg packages across the 330 stores for April 2012 (the middle of the period). The x-axis shows the relative prices (price of a 2 kg package divided by the price of two 1 kg packages). There are substantial (up to 35 percent) differences in the amount of quantity discount for this month. In most cases, these measures are below 1, as one would expect. There are some values above 1, where purchasing two 1 kg packages is less expensive than one 2 kg package. This feature arises because of temporary discounts (see below) and is extensively discussed in Section 4.

Temporary promotions. On Figure 4, I plot the percentage of volume on promotion over time for the most popular brand. This shows that in almost all cases, either zero or 100 percent of the volume was sold at promotional prices, which strongly suggests that promotions were conducted nationally by the chain, rather than individually at the store level. This also suggests that the price variation I observe across stores cannot be explained

[5]This pricing behavior by the firm could be caused by price discrimination or cost differences across sizes - see, e.g., Cohen (2008). I also computed the same figures separately for different areas based on living standard measures, and I do not find any noticeable differences compared to Table 2.

Table 2: Quantity discounts

	Sunlight regular	Sunlight tropical	OMO
250g	1.30	1.29	1.30
500g	1.25	1.24	1.27
1kg	1	1	1
2kg	0.80	0.80	0.78
3kg	0.74		
5kg	0.74		

Notes: Unit prices (per 1 kg) relative to a 1 kg package of the same brand. Based on regular (non-promotional) prices.

by promotions taking place in some stores but not others. This point is further supported by Figure A.5 in the Appendix.

Note that promotional periods for the 2kg/3kg and the 5kg package typically do not coincide. For example, on the week indicated with the vertical line, the 2kg package was on promotion and the 5kg package was not. Because the regular unit price of these packages tends to be similar (Table 2), a promotion on the 2 or 3kg package creates periods where it is cheaper to buy a 2+3kg than 5kg. This feature of the data will be discussed extensively in Section 4 below.

3.2.3 Store characteristics

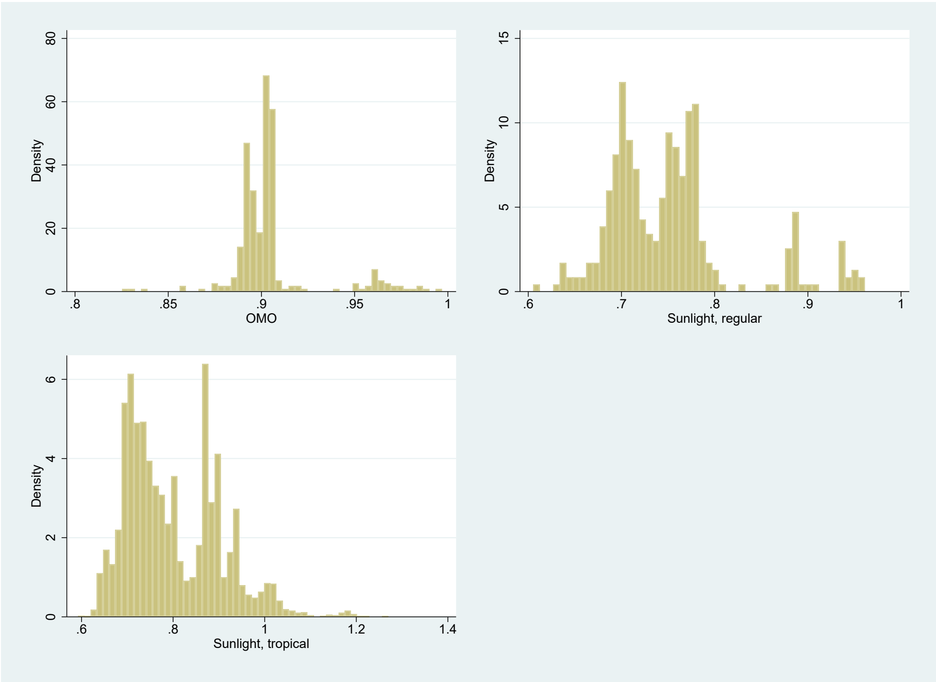
The data allowed me to collect the GPS coordinates of each store using Google Maps. Knowing the exact location of the stores is an advantage of my dataset relative to, e.g., the popular US scanner database from Nielsen.

There are Shoprite stores in all nine South African provinces (Figure A.6), and there are stores are located in urban, suburban and rural areas as well (Figure A.7).

Since I know the identity of each store, I was able to collect store characteristics from individual stores' websites. I collected the following information: store is located in a shopping mall, store is located in a city center, and Sunday opening hours (Table A.8 in the Appendix). Approximately 9 percent of the stores are in a shopping mall. There is large variation in Sunday opening hours across stores, ranging from closed all day to open until 9 pm. I use these variables to proxy for unobservable market characteristics, such as popularity or accessibility of the store.

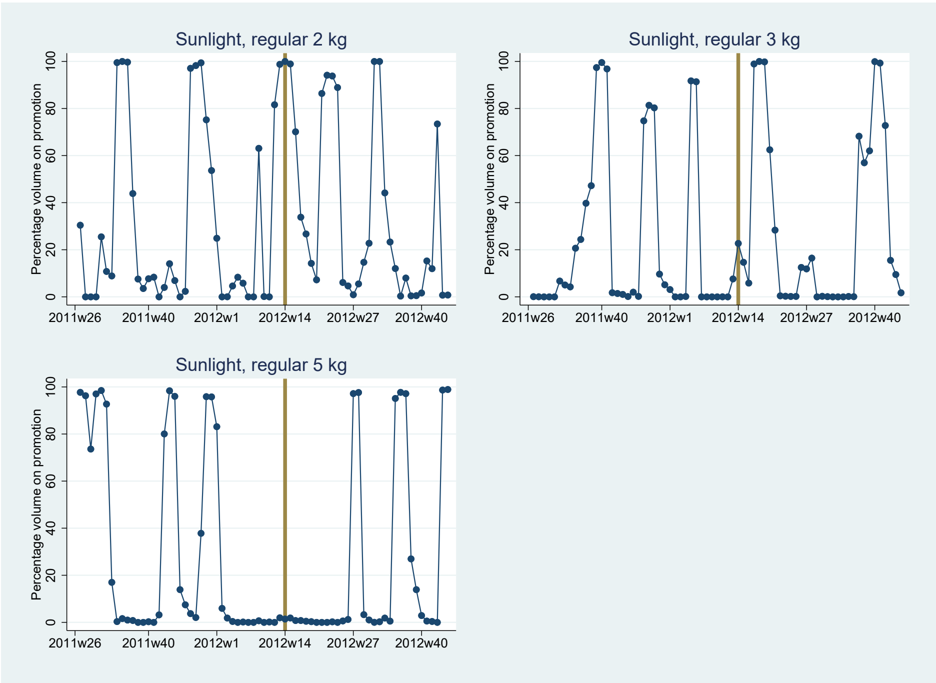
Unilever uses a Living Standard Measure (LSM) to categorize stores based on their location for internal marketing purposes. This measure is designed to approximate consumers'

Figure 3: Store level variation in quantity discounts



Notes: Prices across stores of one 2 kg vs two 1 kg packages of the same brand for April 2012 (middle of sample period).

Figure 4: Percentage of volume on promotion for selected products



Notes: Total across all stores. The vertical line indicates the same week on each panel.

purchasing power, and places each store into one of three main categories.[6] For simplicity, in the rest of the paper I refer to low-, middle- and high-income areas based on this LSM definition. In my data, 17 percent of the stores fall into the low-income category, 44 percent into the middle-income category, and 39 in the high-income category (Table A.8).[7] Average household income in the lowest category is 2285 Rand (310 USD) per month, which makes this group similar to low-income households elsewhere in Africa.

3.3 Demographic data

I have two additional sources of data which I combine with the store level scanner data. First, I use the full version of the latest South African Census Community Profile dataset. The Census was undertaken exactly in the time period covered by the scanner data, in 2011, and it allows matching households to very precise geographical areas. Specifically, the data contains an identifier called small area code which covers covers only 100-500 households (typically 2-3 street blocks). This allows me to work with a very precise market definition: since I know the GPS coordinates of each store, I can define markets based on distance to the store. I describe describe this market definition in Section 6.1. Since I observe individual households, I can directly sample individual households for estimation.

The Census contains household level information on annual income, type of main dwelling, population group (ethnicity) and gender of the household head, and ownership of various appliances. In the demand estimation I found it useful to include ownership of a washing machine and ownership of a car. Ownership of a car can be important in explaining whether households choose larger package sizes that are more difficult to transport.

Table A.10 provides summary statistics and compares households on the market of Shoprite stores to the entire South African population. Although Shoprite stores are somewhat over overrepresented in and around larger urban areas, 15% of the households these stores' markets live in rural areas and 14% of them live in informal dwellings.

3.4 Survey data

The second source of information on households is from two rounds of a survey I conducted specifically for this paper to gather information on (i) how households purchase laundry

[6]The Living Standard Measure is created by the South African Audience Research Foundation and is used widely by companies for marketing purposes. Based on multiple measures, it divides the population into 10 LSM groups, from 10 (highest) to 1 (lowest). In the data, stores are categorized into low LSM (1-4), medium LSM (5-6) and high LSM (7-10) areas. More details about LSM are available at http://www.saarf.co.za/lsm/lsms.asp

[7]The share of the adult population in South Africa in each of these categories is 32, 18 and 50 percent, respectively.

detergents, (ii) the amount of laundry detergent they keep at home (i.e., their inventory), and (iii) their storage costs. The survey asked the households where and how frequently they usually buy laundry detergent, and what kind of transportation they use. The surveyors also asked households to show them the detergents they had at home, and recorded the package size and how much was left in the package (e.g., half empty, almost gone, etc). This which allows me to determine the household's exact inventory.

The literature using US data typically relies on house size or urban/suburban location of the household to proxy the (physical) cost of storage. In a developing country setting there might be other considerations. For example, a larger package may get ruined or stolen, or invite neighbors asking to use it. These are potentially important determinants of whether a household decides to keep a large package at home, and I therefore collect information on all these factors as well.

For the survey, I randomly selected 3 stores, one in each LSM income area, and randomly sampled 100 households around each of these stores. In total, through the two rounds of the survey, I have 600 inventory observations for 300 households. The first round of the survey took place in December 2020 and the second in March 2022, about 16 months apart, which also corresponds to the length of my data, although in a different time period.

4 Savings due to bundling

As documented in the previous sections, the product features non-linear pricing, with similar unit prices within each of three broad size categories, small, medium and large. At the same time, temporary promotions, which usually target only some sizes, can substantially decrease prices.[8] In many markets, these two facts combined result in situations where it is cheaper to buy two smaller packages instead of a larger one to obtain the same total quantity. That is, the relatively flat quantity discount within size groups combined with temporary promotions that affect only one or two sizes per brand induce non-monotonicity in the nonlinear pricing schedule.

In the data 14 products are being sold, where a "product" is a particular package size of a particular brand. By bundling smaller packages, the quantities corresponding to these package sizes can be purchased in a total of 30 different combinations. For example, 1 kg of Sunlight Tropical could be purchased as four 250 g packages or two 500 g packages (as well as a non-bundled 1 kg package). Table 3 shows all these possible product bundles.[9]

[8] From the firm's point of view, the use of promotions on some package sizes but not others of a given product is consistent with Aguirregabiria (1999).

[9] Since substituting a larger package with a mixture of *different* small packages can never be optimal (e.g., buying two 250g and one 500g package instead of one 1kg package), such combinations are not listed.

Table 3: Possible bundles

	Stores	Months	Markets	Price saving	Or. price
Sunlignt, tropical					
Buy 2x250g instead of 500g	265	16	988	0.20	16.32
Buy 2x500g instead of 1kg	54	1	54	0.96	23.70
Buy 4x250g instead of 1kg	0	0	0		
Buy 2x1kg instead of 2kg	137	7	214	2.10	45.34
Buy 4x500g instead of 2kg	2	1	2	1.62	45.18
Buy 8x250g instead of 2kg	0	0	0		
All Sunlight, tropical	287	16	1207		
Sunlight, regular					
Buy 2x250g instead of 500g	290	16	974	0.22	16.00
Buy 2x500g instead of 1kg	13	1	13	0.88	26.55
Buy 4x250g instead of 1kg	0	0	0		
Buy 2x1kg instead of 2kg	140	6	220	1.46	45.30
Buy 4x500g instead of 2kg	0	0	0		
Buy 8x250g instead of 2kg	0	0	0		
Buy 1x2kg and 1x1kg instead of 3kg	308	7	967	3.46	66.37
Buy 3x1kg instead of 3kg	4	3	5	1.39	59.97
Buy 6x500g instead of 3kg	0	0	0		
Buy 12x250g instead of 3kg	0	0	0		
Buy 1x2kg and 2x500g instead of 3kg	11	1	11	1.48	66.08
Buy 1x2kg and 1x3kg instead of 5kg	329	16	1764	11.53	100.00
Buy 2x1kg and 1x3kg instead of 5kg	10	3	10	3.62	106.50
Buy 2x2kg and 1x1kg instead of 5kg	206	8	263	7.73	101.52
Buy 5x1kg instead of 5kg	2	1	2	0.33	95.82
Buy 10x500g instead of 5kg	0	0	0		
Buy 20x250g instead of 5kg	0	0	0		
Buy 2x2kg and 2x500g instead of 5kg	0	0	0		
All Sunlight, regular	330	16	3355		
OMO					
Buy 2x250g instead of 500g	283	16	1849	0.42	17.30
Buy 2x500g instead of 1kg	0	0	0		
Buy 4x250g instead of 1kg	0	0	0		
Buy 2x1kg instead of 2kg	3	1	3	0.40	43.84
Buy 4x500g instead of 2kg	0	0	0		
Buy 8x250g instead of 2kg	0	0	0		
All OMO	283	16	1852		

Notes: Entries indicate the number of times a specific bundle is the cost-minimizing option in the data. Total number of stores is 330, total number of months is 16. Total number of markets is 5255.

Out of the 30 possible bundles, 16 are cost-minimizing at least once in the data, i.e., it is cheaper to purchase a bundle instead of a single package. The largest average price saving comes from purchasing a 2 kg and a 3 kg package instead of a 5 kg package. This is worth it in 1764 out of the 5255 markets in the data. There are markets where bundling can be cost-minimizing for multiple sizes and/or brands at the same time. For example, I see 363 markets where buying two 250 g packages of OMO instead of one 500 g package *and* buying a 3 kg and a 2 kg package of Sunlight instead of one 5 kg are both cost-minimizing bundles. In total, there are 57 possible combinations of these bundling opportunities in the data. The online Appendix lists all these cases.

As can be seen in Table 3, the number of unique months as well as the number of unique stores is almost always the same, reflecting the fact that bundling opportunities do not arise only is specific months or in specific stores. I investigate any potential correlation between bundling opportunities and market characteristics, including whether the store is in a city center or shopping mall, area income and other average household characteristics. The Online Appendix shows the adjusted R^2 from regressions when store and market characteristics are added separately or jointly to explain bundling opportunities for each package size. These 14 characteristics explain no more than 1 percent of the variation on top of the basic controls such as state and month fixed effects. Variables such as area income or store location appear to have little to no correlation with bundling opportunities.

The presence of bundling opportunities has important implications in the context of African countries, where the common belief is that due to severe financial constraints, households often purchase products with the lowest package price rather than the lowest unit price. This appears to be supported by the large market shares of small packages. However, looking purely at market shares can be misleading, as the case studied here illustrates. A consumer might optimally purchase multiple small packages because this bundle is less expensive than a larger package.

In the next section, I first address this issue in the context of demand estimation. I present a method to estimate the share of small packages that may have been purchased as part of a bundle. The estimation relies on the fact that I observe a large number of markets. For each size, I have over 1000 markets where there are no bundling opportunities (Table 4, so that one can proceed with a conventional discrete choice demand model. I use these estimates to infer counterfactual consumer demand on markets with bundling opportunities if such opportunities were absent.

Table 4: Markets with no bundling opportunities

	Unique stores	Unique month	Markets (store × month)
250g	326	16	1064
500g	326	16	1088
1kg	326	16	1088
2kg	326	16	1088
3kg	324	16	1049
5kg	325	16	1029
All	330	16	5255

Notes: Total number of stores is 330, total number of months is 16. Total number of markets is 5255.

5 Model and estimation

Since laundry detergent is a storable product and prices change over time because of temporary discounts, the consumer may purchase the product in a dynamic manner. Besides these obvious features, the demand model should also incorporate the fact that laundry detergent is a differentiated product. The estimation needs to take into account of the fact that purchase is observed at the market level rather than the individual level, and that there are bundling opportunities present on some, but not all markets. The model and estimation below incorporate all these features.

5.1 Model setup

Consider the following model based on Hendel and Nevo (2006). On a given market, consumer h makes monthly choices between brands $j = 1, ..., J$ sold in different package sizes x. The good is storable, and any quantity not consumed is stored as inventory.

The per period utility is

$$\begin{aligned} U(c_{ht}, x_{hjt}, i_{h,t+1}) &= u(c_{ht}, v_{ht}, x_{hjt}, \gamma) - C(i_{h,t+1}, \boldsymbol{\theta}^C) - F(x_{hjt}) \\ &\quad + \boldsymbol{\beta}_x \mathbf{a}_{jxt} + \xi_{jxt} + \mu(\mathbf{a}_{jxt}, \mathbf{D}_h, \boldsymbol{\theta}_x^{\mu}) + \varepsilon_{hjxt} \end{aligned}$$

where c_{ht} is the quantity consumed (from all brands) by the consumer at time t, v_t is randomness in the consumer's needs, $C(.)$ is the storage cost of inventory i, $F(.)$ is size-specific cost of purchase (such as a transportation cost), $\mathbf{a}_{jxt}$ are observed product characteristics (price and brand), ξ_{jxt} is the valuation of unobserved characteristics, $\mu(.)$ is the valuation of product characteristics as a function of consumer demographics $\mathbf{D}_h$, and ε_{hjxt} is a Logit error. The parameters to be estimated are the marginal utility of consumption γ, storage cost parameters $\boldsymbol{\theta}^C$, and the valuation $(\boldsymbol{\beta}_x, \boldsymbol{\theta}_x^\mu)$ of product characteristics.

The inventory transition is given by

$$i_{h,t+1} = i_{ht} - c_{ht} + x_{hjt} \tag{1}$$

The consumer's problem is:

$$V(\sigma_1) = \max_{\{j, x_{hjt} | \sigma_t\}} \sum_{t=1}^{\infty} \delta^{t-1} E\left[U(c_{ht}, x_{hjt}, i_{h,t+1}) | \sigma_1\right] \tag{2}$$

s.t. $0 \leq i_{ht}$, $0 \leq c_{ht}$, $0 \leq x_{hjt}$ and (1), where the state variables σ_t are (i_{ht}, v_{ht}, c_{ht}), the vectors of product characteristics $\mathbf{a}_{jxt}$ and ξ_{jxt}, and the vector of choice-specific shocks.

I make similar assumptions regarding the distribution and evolution of consumption shocks, choice-specific shocks and prices as Hendel and Nevo (2006). Namely, the consumption shock v_{ht} is independently distributed over time and across consumers, prices follow a first-order Markov process, and the choice-specific shocks are i.i.d. Type-I extreme value. See that paper for discussion and justification of these assumptions.

5.2 Towards estimation

The estimation builds on Hendel and Nevo (2006), which in turn builds on the nested algorithm of Rust (1987). The objective of the dynamic model is to describe the choice among different package sizes. Briefly, to reduce the dimensionality problem, the dynamic decision is based on a summary measure ("inclusive value") of the factors that enter the static demand estimation. In other words, given package size, the choice of which brand of detergent to buy is not part of the consumer's dynamic problem.

For each size sold, define

$$\widetilde{\omega}_{hxt} = \log\left[\sum_j \exp(\boldsymbol{\beta}_x \mathbf{a}_{jxt} + \xi_{jxt} + \boldsymbol{\mu}(\mathbf{a}_{jxt}, \mathbf{D}_h, \boldsymbol{\theta}^\mu))\right], \tag{3}$$

as the inclusive value (quality adjusted price index) for all brands of size x. The sum is over brands which come in size x. To use these inclusive values in the dynamic problem of the consumer an adjustment must be made to account for a different interpretation of the outside options. In the dynamic problem, the interpretation is standard, the value of buying a given package size compared to not buying anything. Since the static demand estimation is a series of BLP exercises, one for each package size, the inclusive value in (3) is interpreted as a the value of buying a specific size compared to anything else, which can be either a different size or not buying any detergent. This is discussed in Section 6.5. Using these adjusted inclusive values (ω_{hxt}), I can rewrite the dynamic programming problem as

$$\begin{aligned} V(i_{ht}, v_{ht}, c_{ht}, \boldsymbol{\omega}_{ht}, \boldsymbol{\varepsilon}_{ht}) \quad = \quad & \max_{x_{ht}|\sigma_t} u(c_{ht}, v_{ht}, \gamma) - C(i_{h,t+1}, \boldsymbol{\theta}^C) - F(x_{ht}) + \omega_{hxt} + \varepsilon_{hxt} \\ & + \delta E \left[V(i_{h,t+1}, v_{h,t+1}, \boldsymbol{\omega}_{h,t+1}, \boldsymbol{\varepsilon}_{h,t+1}) | i_{ht}, v_{ht}, \boldsymbol{\omega}_{ht}, \boldsymbol{\varepsilon}_{ht}, x_{ht} \right], \end{aligned} \quad (4)$$

with the additional assumption about the transition process

$$\Pr(\omega_{hx,t+1} | \boldsymbol{a}_t, \boldsymbol{\xi}_t, \mu_{ht}) = \Pr(\omega_{hx,t+1} | \boldsymbol{\omega}_{ht})$$

For estimation, I need to specify functional forms for the objects defined above. Following Hendel and Nevo (2006), I specify

$$u(c_{ht}, v_{ht}, \gamma) = \gamma \log(c_{ht} + v_{ht})$$

$$C(i_{ht}, \boldsymbol{\theta}^C) = \theta_1^C i_{ht} + \theta_2^C i_{ht}^2 + \theta_3^C i_{ht}^3$$

The distribution of ω_{hxt} is assumed to be Normal, with standard deviation σ_x and mean

$$\iota_0 + \sum_{x'} \iota_{x'} \omega_{hx',t-1}.$$

To proceed with the dynamic estimation, one needs to compute inclusive values for all markets. The inclusive values (3) contain marginal utilities of the prices and product characteristics from the static demand. Since individual purchase data is not available, a BLP (1995) type discrete choice estimation is the natural approach to recover these marginal utilities. However, the discrete choice assumption is violated in markets where there is a potential bundling opportunity. I now describe a possible solution to this issue.

5.3 Estimating static parameters and correcting the market shares

First, consider markets with no bundling opportunities (i.e., where based on the observed prices it is not optimal to purchase multiple smaller packages instead of a bigger package). For these markets, I estimate β_x,ξ_{jxt} and $\mu_h(\mathbf{a}_{jxt})$ (that is, all the ingredients for the inclusive values in (3)) using BLP. Since the consumer's static brand choice is conditional on size, a separate BLP estimation is run for each set of products of a given size x. In each case, the interpretation of the outside option is "a choice other than size x."

Next, I consider markets with bundling opportunities. On these markets, I first need to estimate the share of packages that were purchased as part of a bundle. Next, I use these estimates to "correct" the market shares, and compute all necessary ingredients to compute inclusive values and proceed with the estimation of a dynamic demand of the consumer.

I begin by computing the total price of buying each profitable bundle (for example, the price of buying a 2 kg and a 3 kg package instead of a 5 kg package of the same brand when the latter is more expensive). Call this the "effective price" of the given total quantity (in this example, the effective price of 5 kg of the product).

Using these effective prices together with the BLP parameter estimates, it is possible to compute the "corrected" market share of, say a 2 kg package - the share of 2 kg packages that were not purchased as part of a bundle. To do this, I use the parameter estimates to compute the demand increase for a 5 kg package when its price drops to the effective price. This is the extra demand for 5 kg of the product that will be fulfilled through the purchase of bundles. In this example, if the observed market share of the 5 kg packages is $S_{jt}^{(5)}$ and their predicted market share under the effective price is $\hat{s}_{j5t}$, then the corrected market share of the 2 kg and 3 kg packages will be $\hat{s}_{j2t} = S_{jt}^{(2)} - (\hat{s}_{j5t} - S_{jt}^{(5)})/2$ and $\hat{s}_{j3t} = S_{jt}^{(3)} - (\hat{s}_{j5t} - S_{jt}^{(5)})/2$, respectively.

To compute $\hat{s}_{j5t}$, I simulate 1000 values of the ξ_{j5t} estimated for the markets with no bundling, compute the market shares $\hat{s}_{j5t}$ for each simulation (using the actual observed prices and the effective prices), and take the average across simulations. Note that this procedure retains that attractive feature of BLP that no distributional assumptions are made on ξ.[10]

What remains is the computation of the ξ's. To obtain these, I use the BLP parameter estimates β_x and $\mu_h(\mathbf{a}_{jxt})$ and solve systems of equation of the form

$$s_{jxt}(\mathbf{a}_{jxt}, \beta_x, \xi_{xt}, \mu_h(\mathbf{a}_{jxt})) = S_{jt}^{(x)},$$

i.e., equating the model-predicted market shares to those observed in the data. For brands/-sizes involved in bundling (for example, the 2, 3, and 5 kg packages in the example above), I use the corrected market shares $\hat{s}_{jxt}$ obtained above.

5.4 Estimation of the dynamic parameters

As described above, the inclusive values in (3) allow simplifying the dynamic problem, which now requires keeping track only of the expected utility conditional on size. Since the (simulated) households have specific characteristics, the inclusive values will be household-specific

[10]In robustness checks, I extend the support of ξ_{j5t} used in these simulations in multiple ways, such as adding +/- 2 standard deviations. See the detailed results in the Appendix.

as well. Below, I start by estimating a single process for all households; I then relax this assumption and allow separate processes for household groups by income area.

The estimation proceeds by solving the consumer's dynamic programming problem for each trial of the dynamic parameters. Consumers choose package size as a function of their state variables. These state variables include inventory, consumption shocks, random shocks to the consumer's size choice, and inclusive values. Consumption is assumed to be exogenously given. The consumer trades off the cost of holding an inventory with its benefits. The benefits arise on the one hand from the quantity discount embedded in purchasing a larger size, and on the other from the ability to exploit temporary price discounts. The consumer's inventory is an endogenous state variable which is determined by the previous level of inventory, purchased size and current consumption.

The Bellman equation used in the estimation is given by equation (4). The value function is approximated by a polynomial function of the state variables (inventory, consumption shocks, random shocks to the consumer's size choice, and inclusive values). The estimated value function is an input to compute the dynamic choice probabilities of each simulated consumer.

The model implies a set of choice probabilities of each package size for each (simulated) consumer and market:

$$\Pr(x_{ht} = x) = \frac{\exp(\omega_{hxt} + M(c_{ht}, i_{h,t+1}, v_{ht}, \boldsymbol{\omega}_{ht}, x))}{\sum_{x'} \exp(\omega_{hx't} + M(c_{ht}, i_{h,t+1}, v_{ht}, \boldsymbol{\omega}_{ht}, x'))}$$

where

$$\begin{aligned} M(c_{ht}, i_{h,t+1}, v_{ht}, \boldsymbol{\omega}_{ht}, x) &= u(c_{ht}, v_{ht}, \gamma) - C(i_{h,t+1}, \theta^C) - F(x) \\ &\quad + \delta E\left[V(i_{h,t+1}, v_{h,t+1}, \boldsymbol{\omega}_{h,t+1}, \boldsymbol{\varepsilon}_{h,t+1}) | i_{ht}, v_{ht}, \boldsymbol{\omega}_{ht}, x\right] \end{aligned}$$

Since in the current application data is at the store level, simulated consumers are aggregated across markets. Parameters are computed using a simple simulated minimum distance estimator, minimizing the squared distance of model-predicted and observed market shares. Let N_H denote the number of simulated consumers on a market, and define

$$Q(\boldsymbol{\theta}) = \sum_x \left(S^x - \frac{1}{N} \sum_{h=1}^{N} \Pr(x_{ht} = x) \right)^2$$

The estimator minimizes $Q(\boldsymbol{\theta})$ summed across markets.

6 Estimation details

6.1 Market definition

There are two important practical implications of how one defines the market (i.e., of choosing the market size). First, the market size plays a role in determining the market share of the outside good. If the market size chosen is too small (smaller than the quantity of laundry detergents actually sold), estimation becomes impossible. If the chosen market size is too large, that also has undesirable consequences. In a typical static demand estimation, the outside good is a normalization. The case considered here is different: the outside market shares need to accurately reflect the probability that the consumer does not buy any laundry detergent in a given month to be consistent with the dynamic part of the demand estimation. I will discuss this issue in more details in Section 6.3.

Second, the market definition affects which demographic characteristics should be included in the estimation. In South Africa, many areas are segregated, but substantially different populations live in relatively close proximity to each other. Thus, a more precise market definition helps match the actual households that use particular stores. Similarly, having the right market definition can be important to know which characteristics may affect consumers' product size choices. For example, it is common for the households considered here to walk to the store, which makes the cost of carrying a large package an important consideration in the dynamic model.

I define markets based on both time period and geography. The scanner data is at the store-month level, which is the right frequency to model laundry detergent purchase. Based on my survey, 93.7 percent of households said that they purchased detergent once a month, and only 2 percent said that they made more frequent purchases. Since I know the exact location of the stores, I can define the market based on distance. Stores are in areas with different population density. Even though I do not know the location of all possible supermarket chains in South Africa, I know the distance to the nearest store of the same chain. This distance varies substantially in my data, which suggests that using a simple definition of the market is not useful. For example, it is not useful to define the entire metropolitan area of Tshwane (the area around the capital city of Pretoria) as a market, since there are are 26 stores in this area and almost all stores have a neighbor within 2 kilometers.

I define the market size for each store as follows. Based on the Census, for each store, I compute the number of households N_r who live within r km, where $r = 1, 2, ..., 50$. I also compute the maximum number of laundry products the store sells in a given month over the time period observed in the data, L. I set the market size to be $\min_r \{N^r | L \cdot 1.1 \leq N^r\}$, i.e., the radius where the number of households exceeds the total number laundry products sold, plus 10 percent. Using this definition, the average market size is 11,432 households. The market is defined as a radius of 5 km or less from the store for 94.8 percent of the stores. The radius is 1 km for 43.9% of the stores, and it is 2 km for 32.7%. The distribution of the market radius is shown on Figure A.9. I also use this market definition to create control variables in the estimation, specifically, the number of stores from the same chain in a given market, and the distance from the store to its nearest neighboring store.[11]

6.2 Details of the static demand estimation

Estimation of the static demand parameters follows the Generalized Method of Moments (GMM) algorithm proposed by Berry, Levinsohn and Pakes (1995). Detailed treatments of the procedure can be found in Berry, Levinsohn and Pakes (1995) and Nevo (2001). Briefly, consider a dataset with information on product and market characteristics $\mathbf{a}_{jxt}$ and actual product shares $S_{jt}^{(x)}$. Berry et al. (1995) show that, for given $\mu_h(.)$, it is possible to numerically solve for $\psi_{jxt} \equiv \beta_x \mathbf{a}_{jxt} + \xi_{jxt}$ from the equations $s_{jxt}(\mathbf{a}_{jxt}, \beta_x, \xi_{xt}, \mu_h(\mathbf{a}_{jxt})) = S_{jt}^{(x)}$, i.e., equating the model-predicted market shares to those observed in the data. Using the resulting values of ψ_{jxt}, one can express the unobserved product characteristics as $\xi_{jxt} = \psi_{jxt} - \beta_x \mathbf{a}_{jxt}$, a nonlinear function of the model parameters. Identification relies on moment conditions $E[\xi_{jxt}|\mathbf{Z}_{jxt}] = 0$ where the $\mathbf{Z}_{jxt}$ are suitable instruments, and estimation is via GMM.

Linear variables. The linear part of the utility, $\beta_x \mathbf{a}_{jxt}$, includes price, brand dummies, store characteristics and demographics. Store characteristics are included as they are likely to be correlated with consumer choices as well as prices and/or the violability of a given package size. For example, if stores located in city centers are more likely to offer smaller package sizes, then we need to control for the location of the stores. Similarly, I control for whether the store is in a larger shopping mall, opening hours on Sunday, and the distance to the nearest store to control for popularity and accessibility. For example, stores closed on Sunday and located in the city center might be used by people returning home from work and looking for smaller package sizes.

The linear part of the utility also includes a set of market-level average household demographics that exhibit little to no variation within markets. This lack of variation in some demographic variables is due to features of the South African setting combined with the relatively small markets I consider. In this application, perhaps uniquely among similar discrete choice demand applications, I see very segregated markets. This arises from the segregation

[11] The median distance to the nearest store is 4.7 km, the mean is 17.8 km.

of South African neighborhoods, and it implies that there is little to no variation in, e.g., population groups across individuals within markets. In addition, as described above, markets are defined as relatively small areas around each store, and consequently characteristics such as "urban" or "rural" do not change across individuals within markets. The full set of variables included are shown in Table A.16 under "Linear parameters."

Nonlinear variables. To model individual level heterogeneity among simulated consumers, I use demographic characteristics that vary within market and are relevant in the current context. I use household income, a binary variable indicating whether the household head is male, and four categories based on whether the household owns a washing machine and/or a car. Table A.10 illustrates the variation in these variables. For example, there is a meaningful portion of the population that owns a car but does not own a washing machine. Since I am able to draw individuals from the Census, I do not need to estimate the covariance of the different demographics and can instead use their empirical distribution.

The nonlinear part of the utility, $\mu_h(\mathbf{a}_{jxt})$, interacts these household demographics with price and with a constant. Interactions with the constant capture heterogeneity in the valuation of the outside good among individuals. The interpretation of the outside good in each BLP estimation is that the consumer decides to purchase no product (or a product not modelled here), or a different size. For example, the fact that a household has a car is probably an important determinant of which size they purchase.

I do not include interactions between brand dummies and household demographics. Including these would only be useful if there was variation in either the availability of brands across markets, or market shares, correlated with the demographics included in the model. Here, there is little variation in the availability of brands across markets. In addition, all the products I consider are powdered hand-washing detergents, and the biggest differences between them (other than the brand) are likely to be factors such as scent. Preference for these factors is unlikely to be explained by the demographic variables that are typically available in Census data. For the full set of nonlinear variables included, see Table A.16 under "Non-linear parameters."

Instruments. To identify the model, instruments are needed for two reasons: to control for price endogeneity, and to identify the nonlinear parameters. Some of the previous attempts to instrument for price are not feasible in the current context. There are no product characteristics in $\mathbf{a}_{jxt}$, only brand dummies. None of the product characteristics change across markets (either across months or across stores) and there is little variation in the availability of different brands conditional on size across stores. This prevents the use of instruments based on exogenous product characteristics. Instead, similar to Nevo (2000), I construct instruments using the prices of the same product in different markets. Specifically, I identify the 15 closest neighbors of each store.[12] I regress stores' price on these 15 prices and rank neighbors based on the size of these correlations. To determine the number of neighboring prices used as instrument, I compute F statistics and include neighbors until there is no further increase in the F statistic.

To identify the nonlinear parameters, I use the market-level average of the demographic characteristics that enter the nonlinear part of the utility. This includes household income, male household head, and the fraction of households owning cars/washing machines. In some specifications, I found it useful to add the interaction of some of these instruments and the most relevant price instrument.

Other considerations. Demand specifications are the same across all sizes. For each market, I simulate 400 individuals, and the same individuals are used in estimating the demand for each size. Standard errors reported below are clustered at the store-month level, to allow for both heteroskedasticity and correlation of the shocks ξ_{jxt} across products within a market. For each specification, I report the J overidentification statistic, and the Newey-West D-test for the null hypothesis that the nonlinear parameters are jointly 0. The numerical stability of the estimates is ensured through the choice of optimizer, convergence criterion, and by eliminating a source of instability in the typical codes used to implement the procedure following Ujhelyi, Chatterjee and Szabó (2021), Appendix D.5.

6.3 Details of the dynamic model

Inventory and consumption. The solution of the dynamic programming problem requires identifying the dependence between purchased package size, inventory, and consumption. Typical applications must address the issue that only households' purchases are observed, while consumption and inventory are unobserved. For example, Hendel and Nevo (2006) generate an initial distribution of inventories and compute optimal consumption based on the model (see also Erdem, Imai and Keane (2003)). In my survey, I directly ask households about each of these elements, and use this information combined with the scanner data to identify the dynamic parameters. I use three variables from the survey: consumption, quantity of detergent in inventory, and the package size(s) in inventory.

For consumption, I ask households how often they typically do laundry (daily/weekly/-monthly), and I separately ask how many loads of laundry they typically do each time. Based on these, I compute how many loads of laundry a household typically does in a month, and the associated consumption of detergent (one load of laundry requires about 100 gram of

[12] If that particular brand is not sold in that particular store in a month, than the next store prices are used. Thus I select the 15 closest store where the particular product was actually sold.

powdered detergent, based on the directions on the package). Table 5 shows the summary statistics.

Table 5: Detergent consumption by income area in the survey

	Mean	Median	St.dev.	Min	Max	N
Low-income area	1229	1200	589.01	200	3200	100
Middle-income area	1106.06	1200	535.82	400	3200	99
High-income area	1133	1200	459.7	400	2400	100

Notes: Powdered detergent consumption in the survey by income area. Values are in gram.

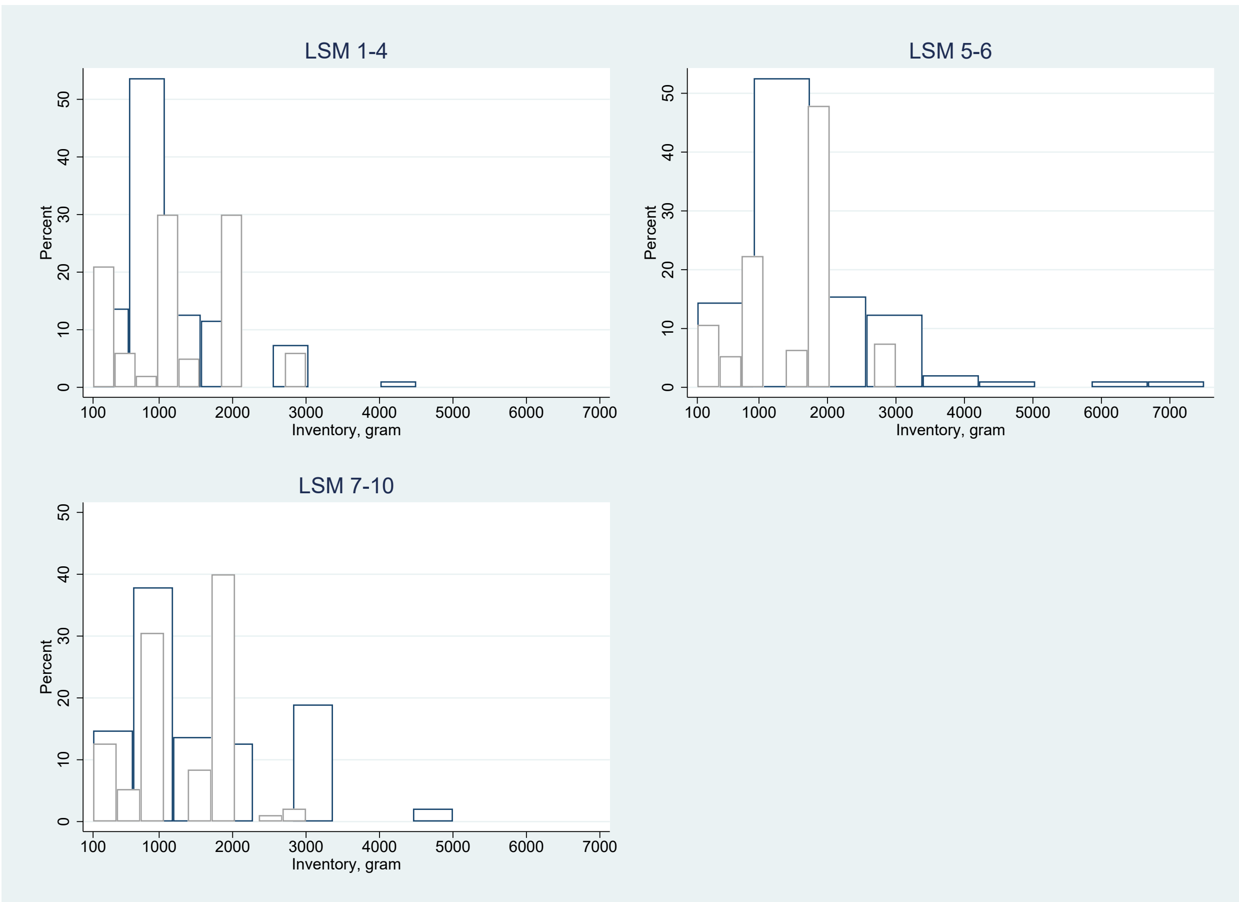
For inventory, the surveyors asked households to show them what kinds of laundry detergent they had at home. The surveyors recorded a description of the item(s) shown to them, including the package size and the amount left in each package - for example, "Sunlight 500 grams in a yellow plastic, dry, half empty" "OMO 2KG in plastic package, dry, ordinary washing powder quarter full" and "OMO washing powder regular 2kg just opened and dry." (No respondents refused to answer this question.) Based on this information, I compute the current inventory of detergents for each household. In some cases, the surveyor recorded that the package was "almost full" or "almost empty." In the first case, I subtract 100 g from the package size, and in the second case, I use 100 g. These numbers correspond to the suggested amount of detergent used for one load. There are 289 recorded inventory levels from the first round of the survey, and 287 from the second round. Respondents without an inventory level showed either liquid detergents or laundry bars (soaps); all households had at least some detergent at home. Figure 5 below shows the distribution of inventories by LSM area. The data reveals some differences in inventory based on income level.

Since I have two visits per household, I see that households do not always buy the same package size, i.e., there is variation between the package sizes recorded during the two visits. Interestingly, the aggregate inventory of the 300 households is very similar at the two time periods.

Further details on how the survey data was combined with the scanner data, and additional details on solving the dynamic programming problem can be found in the Appendix.

Inclusive value process. Once the static demand parameters are estimated, I compute inclusive values using equation 3. I assume a first-order Markov process for the inclusive values of each package sizes. Since the demand parameters contain individual level heterogeneity, the inclusive values vary at the individual level as well. To reduce the number of processes to be estimated, I group individuals. In the simplest specification, I group all individuals, and estimate a single process for each size. In the main specification, I estimate separate processes for each of the three LSM areas. I also show results where the grouping is

Figure 5: Household inventory of detergent in the survey



Notes: Powdered detergent inventory in the survey by income area.

done based on the ownership of washing machines and cars. These are the variables which capture most of the individual level heterogeneity in the demand estimation, therefore I use the same variables to capture the heterogeneity in inclusive values. To relax the first-order Markov process assumption, I also include the second lag of the inclusive value for each size. Alternatively, I take the sum of five lags for each size, as Hendel and Nevo (2006).

Other considerations. The per-period utility also contains a set of dummy variables for each package size. This has two objectives. First, it is possible that there is a one-time utility gain or loss from purchasing a specific size. For example, the largest sizes might be difficult to carry, especially when households walk to the store. This one time transportation cost (which is different from the inventory cost) would be accounted for by the size dummies. Second, the parameters used to compute the inclusive values which enter the dynamic model are obtained from 6 different static BLP estimations. Although the specifications are identical across these estimations, the (average) level of utility may differ across specification. The inclusion of size dummies is useful to account for any possible level effects.

The estimation of the dynamic model uses only markets where all choices are available. Consumption shocks are assumed to follow a log-Normal distribution.

Because the parameters of the inclusive value process enter the dynamic estimation in a complex way, and because the estimation involves multiple simulated objects, it is practical to use a bootstrap to estimate standard errors. I draw datasets of the same size with replacement, and repeat the estimation 30 times. This includes resampling the dataset on which the inclusive process is estimated.

6.4 Identification

The identification of the static demand parameters follows the identification of the standard BLP procedure, with instruments discussed in Section 6.2. This section provides a description of the identification of the dynamic problem.

Estimating the dynamic parameters, i.e. the parameters of the utility from consumption and the parameters of the storage cost functions requires solving the dynamic programming problem for each parameter trial. This results in two sets of parameters. The first are the parameters of the polynomials used to approximate the value function. In the main estimation, I use 26 terms for approximating the value function, which includes the state variables of the consumer. Besides the constant, I include a cubic function of log inventory, the inclusive values, and quadratic functions of both choice specific extreme value shocks and consumption shocks. The second set of parameters which can be identified from the dynamic programming problem is a set of size fixed effects. These fixed effects can be interpreted as a one time cost of purchasing a specific size compared to the outside option. More formally, approximate the value function in (4) as

$$V(i_t, v_t, \boldsymbol{\omega}_t, \boldsymbol{\varepsilon}_t) = \boldsymbol{\varphi}\mathbf{r}_t,$$

where $\mathbf{r}_t$ is a vector of polynomial terms, and $\boldsymbol{\varphi}$ are the parameters to be estimated. Specifically,

$$\boldsymbol{\varphi}\mathbf{r}_t = m_t + \eta F(x_t) + \delta\boldsymbol{\varphi}\mathbf{r}_{t+1}, \tag{5}$$

where

$$m_t = \gamma \ln(c_t + v_t) - (\theta_1^C(\log i_t) + \theta_2^C(\log i_t)^2 + \theta_3^C(\log i_t)^3) + \omega_{xt} + \varepsilon_{xt}.$$

Identification relies on data from the household survey, which provides data on current inventory, consumption, purchase, and household characteristics. In typical scanner data studies, only purchase is observed, and there is no information on either inventory at home or consumption. This makes it difficult to identify the dependence between purchase size, inventory, and consumption. The information contained in my survey alleviates this difficulty. In this case, it is not necessary to simulate initial household inventory, or compute the optimal consumption for each potential size purchased. Similarly, I can directly compute future inventory as current inventory plus size purchased, minus consumption.

In addition, the survey has the same demographic information about the households (for example, area income, ownership of washing machine and cars) as the elements which enter in the static demand estimation. This makes it possible to match the inclusive values obtained from the static demand estimation to surveyed households, and thus to their inventory, consumption and purchase choices.

Using these survey data and inclusive values, for a given realization of shocks and a given draw of dynamic parameters, (5) can be written as

$$m_t = \boldsymbol{\varphi}(\mathbf{r}_t - \delta\mathbf{r}_{t+1}) - \eta F(x_t)$$

where m_t, $\mathbf{r}_t$, $\mathbf{r}_{t+1}$ and $F(x_t)$ are all known. Thus, the parameters $(\eta, \boldsymbol{\varphi})$ can be estimated using OLS.[13]

With both inventory and consumption observed, identification of the remaining dynamic parameters γ and $(\theta_1^C, \theta_2^C, \theta_3^C)$ follows the standard arguments in Rust (1996), Magnac and Thesmar (2002), and Aguirregabiria (2005) (see Hendel and Nevo (2006, p.1653)). In prac-

[13] Hendel and Nevo (2006) also include size specific dummies $F(x_t)$ in the per period utility, but they cannot identify these through the dynamic programming problem. The difference is that they do not know optimal consumption of the households. Because of this, they estimate these fixed effects as part of their full optimization procedure.

tice, the identification of the consumption parameter γ comes from two sources of variation in the data, both of which produce variation between planned and realized consumption levels. First, there is a consumption shock to planned consumption. Second, some purchases do not allow for the full planned consumption, because consumption cannot be higher than the current inventory of the households.

6.5 Computing inclusive values for the dynamic problem

In the dynamic problem the interpretation of inclusive values ω_x is relative to the outside option of not buying any detergent, the value of which is normalized to zero. Since the static demand estimation is a series of BLP demand estimations, one for each package size, the interpretation of the inclusive values are different. For a given size, the computed inclusive value $\widetilde{\omega}_x$ is the value of buying the specific size compared to everything else (including the outside option), rather than the value of compared to only the outside option. To directly apply the inclusive values computed from the static demand estimation in the dynamic problem, these two different definitions need to be reconciled. More specifically, the relation between inclusive values across the two problems is given by:

$$\omega_x - \sum_{j \neq x} \omega_j = \widetilde{\omega}_x$$

for package sizes $x = 0.25, 0.5, 1, 2, 3, 5$. Solving yields

$$\omega_x = \left(3\omega_x - \sum_{j \neq x} \widetilde{\omega}_j \right) /8.$$

7 Estimation results

7.1 Static demand estimates

I begin by documenting parameter estimates of the static choice between brands, conditional on size. In Table A.16, each column shows the results from a separate BLP estimation for the choice of brands given a specific package size. As described above, these estimates use data from markets where there are no bundling opportunities. Each of these specifications passes the J-test for the validity of the moment conditions, and the Newey and West (1987) D-test always rejects the null that the nonlinear parameters are jointly 0.

In general, the parameter estimates are sensible. The price coefficient has the expected negative sign and is statistically significant in each specification. Households headed by a male have less elastic demand. This could reflect the fact that these households are wealthier, or it could reflect differences in preferences (e.g., men could pay less attention to laundry detergents when shopping).[14]

Consumers have higher utility from buying larger packages in stores that are open on Sunday: the corresponding coefficient is negative for smaller packages and positive for larger packages.

Coefficients on the (constant $\times$ car only) interaction term are positive and significant for the larger packages (1kg and above) and switch signs for smaller packages. This indicates that large packages are valued particularly by consumers who own a car but do not own a washing machine.[15] This makes sense, as these are the consumers who use hand-wash detergents more and can more easily transport the larger packages. Coefficients on (constant $\times$ no car or washm) are positive and significant for the smaller packages and switch signs for larger packages (2kg and above). This too makes sense, as these are again the consumers who use hand-wash detergents, but face higher costs of transporting larger packages.

The significance of car ownership in explaining demand for laundry detergent suggests that transportation costs are an important factor in this setting.

7.2 The effect of bundling opportunities

Table 6 shows the effect of accounting for possible bundling. The upper panel describes markets with bundling opportunities. Each column corresponds to a total quantity purchased by a consumer at a given time. The "Quantity sold" row shows the average number of packages sold of a given size in the data. The "Corrected quantity sold" row shows the estimated number of consumers purchasing that quantity, either by buying one package of the corresponding size, or through bundling. For example, while on average 51.3 units are sold of the 5 kg packages, I estimate that on average 154.09 consumers buy a total of 5 kg - that is, I estimate that, on average, 102.79 ($= 154.09 - 51.3$) consumers find it profitable to buy 5 kg of detergent by creating a bundle of smaller packages. Similarly, although the average number of 250 g packages sold is 385.26, some of these are purchased as part of a bundle - accordingly, I estimate that the average number of consumers buying a total of 250 g is only 349.58.

I calculate the ratio of Corrected sold and Sold separately for each market; the table shows the average and the median of these measures. On the median market, there are over 3 times as many households who purchase 5 kg of detergent as households who purchase a

[14]More generally, the lower price elasticity of male shoppers is consistent with the findings of Fitzpatrick (2017) in Uganda.

[15]The excluded category is composed of consumers who own both a car and a washing machine.

5 kg package. This suggests that looking purely at the sales data can be very misleading regarding the share of households who purchase larger quantities.

For comparison, the lower panel of Table 6 shows the average number of units sold on markets with no bundling opportunities in the data. For some package sizes, average quantity sold differs substantially from the top panel - which is likely explained at least in part by the presence vs. absence of bundling opportunities. For example, in the data there are many more 5 kg packages sold in markets with no bundling opportunities than on markets with bundling opportunities (119.73 vs 51.30). However, the Corrected sold values indicate that this difference could be due to the fact that, on the latter, it is profitable to buy 5 kg detergent in other ways than as a 5 kg package.

Note that in different markets the corrected market share of a particular package size could be larger or smaller than the market share observed in the data. In some cases, more consumers could be buying 2 kg of detergent than 2 kg packages, and in other cases consumers could be using 2 kg packages to create bundles, resulting in more 2 kg packages sold. As shown in Figure A.15, there is indeed substantial variation in the ratio of corrected/original quantity sold for each size.

Because the sizes most affected by bundling are also the ones with the smallest market shares, the impact of the bundling correction on market shares is small, as shown by the comparison of the Market share and Corrected market share rows of Table 6.

Table 6: Effect of buying in bundles

	5 kg	3 kg	2 kg	1 kg	500 g	250 g
Markets with bundling opportunities						
N of markets	2040	2816	3057	1352	2752	2709
Quantities sold	51.30	290.87	2319.79	1476.44	613.18	385.26
Corrected quantities sold	154.09	241.44	2241.21	1408.01	634.93	349.58
Average ratio	5.46	1.28	0.96	0.94	1.06	0.85
Median ratio	3.02	0.88	0.97	0.97	1.02	0.93
Market share	0.010	0.056	0.401	0.277	0.123	0.084
Corrected market share	0.037	0.045	0.391	0.265	0.130	0.077
Markets without bundling opportunities						
N of markets	1029	1049	1088	1088	1088	1064
Quantities sold	119.73	164.28	1592.165	1508.52	750.03	514.30
Market share	0.025	0.033	0.305	0.336	0.176	0.130

Notes: Average number of units sold per market.

7.3 Inclusive values

I compute the inclusive values based on the parameter estimates in Table A.16. Table A.17 shows the estimated processes when a single process is estimated for all simulated individuals.

Table A.18 shows the adjusted R^2 from other specifications. First, I estimate separate processes by income area or car/washing machine ownership. Dividing households into these groups does not improve the fit of the regressions. In each case, for every size x, the largest coefficient estimate is on ω_{t-1}^{x}, i.e., the own lagged inclusive value, similarly to specification shown in Table A.17. To relax the first-order Markov process assumption, I also estimate the inclusive value process in two alternative ways: adding the second lag of each size, and adding the total of five lags of each size, as in Hendel and Nevo (2006). As shown in the bottom section of Table A.18, neither of these improves the fit meaningfully compared to the baseline specification.

7.4 Results of the dynamic model

This section summarizes the results from the estimation of the above dynamic model. I begin by discussing the estimated inventory cost and fixed cost of purchase and then turn to evaluating model performance.

Table 7 reports the dynamic parameters from various specifications. Columns (1)-(4) assume a flexible three-degree polynomial for storage cost as a function of inventory. These specifications also contain six fixed costs parameters, one for each size. Column (1) is estimated using all stores and assumes a single inclusive value process for all (simulated) individuals. This specification uses survey data from all income areas. In columns (2)-(4), I present results separately for stores in different income areas, where column (2) results are for low income, (3) for middle and (4) for high income areas. In these cases, I estimate separate inclusive value processes for each income area and use only survey data from households who live in the given income area.

To interpret the inventory cost parameters, suppose that inventory at the beginning of the period is 1 kg (tne median in the data). Then according to column (1), buying a 500 g package increases the storage cost from 10.14 to 12.86, compared to 14.23 if buying a 1 kg package. The average savings due to non-linear pricing from buying the 1 kg package is 6.19.

There is variation in the one-time purchase cost across different sizes and across different income areas. These costs can be interpreted, for example, as transportation costs. Buying a larger package may require hiring a taxi rather than walking home from a store. The estimated fixed cost increases in package size: the cost associated with buying a 5 kg package is about 10 times higher than the cost associated with the smallest package. These costs are similar in low- and middle-income areas, but they are between 20-40% lower in high-income areas. In high-income areas more households own a car compared to in low- and middle-income areas, and it is reasonable to assume that they face lower transportation costs.

Continuing the example above, based on column (1), transportation costs are estimated to add an additional 5.41 Rand (=14.14 - 8.73) to the cost of purchasing a 1 kg package compared to buying a 500 g package. In this case, the sum of the fixed cost of purchase and storage costs (5.41 + 1.37) is similar in magnitude to the total quantity discount.

Overall, the estimates indicate that the fixed costs are larger in magnitude than the inventory costs (for the inventory levels observed in the data). This will have implications in my policy experiments: decreasing these fixed costs, such as through the introduction of mobile stores, can have important welfare effects (see Section 8).

To interpret the inventory cost function predicted by these specifications, note that in the model, inventory costs and the package size dummies explain why a household does not purchase the product that has the lowest unit price. Thus, the identification of the inventory cost parameters is connected to differences in unit prices across package sizes. In a typical application, the unit price is monotonically decreasing in size. In these typical cases a consumer would incur a loss if they decided not to purchase the less expensive, larger product. When, in the data, some consumers do not buy the larger size, the model would explain this by estimating a positive storage cost. With decreasing unit prices, we expect to find increasing storage costs as a function of inventory.[16]

In my case, there is no or very little price discount in the data for packages larger than 2 kg (i.e., 3 and 5 kg) - even though these larger sizes account for about 14 percent of the market. In some cases, the unit price of 5 kg is higher.[17] (See Figure A.19 in the Appendix for the distribution of unit prices across markets.) This feature of the data implies that the storage cost is potentially flat or even declining after the 2 kg package.

Implied inventory costs are plotted in Figure A.20 for the specifications in columns (1)-(4). I consistently find declining inventory costs for larger sizes. As expected, the turning point is around the 2 kg size. I also add a quadratic and a fourth-degree polynomial cost function specification (corresponding to column (2)). Estimation using the quadratic specification yields a minimized objective function value that is 3 times as high as the corresponding cubic specification, indicating an inferior fit for the data. Nevertheless, the quadratic specification

[16] This is indeed what Hendel and Nevo (2006) finds.

[17] Recall that I corrected for this feature of the data in situations where "bundling" is possible. However, in some cases the 5 kg package has a higher unit price than a 3 kg package but there is no bundling opportunity (because the smaller packages a consumer might use to create a 5 kg bundle have even higher unit prices).

also shows a declining part, showing that this feature of the cost function is not due to the functional form assumptions.

Columns (5) - (8) show alternative specifications of the dynamic model for the lowest income area to illustrate specific features of the model and to check robustness in various ways. In column (5) I double the potential consumption shock (setting the mean to 100 g), in column (6) I restrict the maximum inventory level to 6.2 kg, the maximum I observe in the survey data (the median is 1 kg). I do not find any noteworthy changes in implied costs in either specification.

In column (7), I do not include any fixed costs of purchase. This results in negative consumption utility, which, based on the model, is needed to explain why households are not opting for the lowest unit-priced options.

In column (8), I show results corresponding to the specification in column (1) when I ignore bundling in the static estimation and I do not correct the market shares which enter in the dynamic estimation. Failing to account for bundling changes the dynamic parameters only slightly - this is due to the fact that, although bundling affects many markets, in most cases the difference between the corrected and observed market shares is small.

To investigate the performance of the model, I use the estimated parameters to plot the model-implied purchase patterns over time. I simulate 100 consumers for each market over 16 months. Figure 6 shows the fit of the model separately for each of the package sizes for stores in low income areas. As expected, purchases of package sizes with larger market shares are explained better by the model compared to less popular sizes. Nevertheless, the model appears to capture well the over-time fluctuations in market shares due to the temporary price promotions. Similar graphs for the other income areas areas are presented in the Appendix.

Table 7: Dynamic parameter estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cost of inventory								
Linear	136.202	131.801	165.725	124.566	128.147	94.233	-181.093	213.516
	(10.95)	(7.708)	(13.469)	(15.751)	(6.711)	(2.611)	(144.233)	(15.152)
Quadratic	-369.559	-356.981	-444.782	-347.419	-348.205	-237.270	495.237	-595.646
	(32.571)	(24.347)	(38.017)	(46.906)	(21.768)	(7.631)	(578.251)	(43.934)
Cubic	222.137	223.268	268.114	209.539	219.918	145.941	-237.900	358.399
	(20.68)	(17.238)	(23.54)	(29.14)	(15.61)	(6.005)	(-6277.375)	(27.442)
Utility from consumption								
	1213.676	1202.263	1477.798	965.362	1028.818	747.196	-6277.375	1865.538
	(100.257)	(69.178)	(132.969)	(125.933)	(55.17)	(30.196)	(2786.54)	(150.3)
Size dummies								
250 g	-5.113	-5.852	-5.534	-4.388	-5.581	-4.856		-8.055
	(0.272)	(0.181)	(0.349)	(0.327)	(0.156)	(0.098)		(0.392)
500 g	-8.733	-9.709	-9.755	-6.911	-9.163	-6.901		-11.525
	(0.497)	(0.418)	(0.639)	(0.647)	(0.345)	(0.177)		(0.738)
1 kg	-14.141	-15.097	-15.960	-11.402	-14.123	-9.849		-21.421
	(1)	(0.772)	(1.261)	(1.256)	(0.679)	(0.319)		(1.501)
2 kg	-25.876	-27.021	-29.772	-21.105	-25.029	-16.457		-40.189
	(2.028)	(1.577)	(2.516)	(2.63)	(1.387)	(0.617)		(3.065)
3 kg	-41.147	-39.912	-47.916	-35.508	-36.824	-22.617		-67.224
	(3.633)	(2.713)	(4.366)	(5.027)	(2.414)	(0.924)		(5.309)
5 kg	-61.180	-62.427	-71.057	-51.057	-57.415	-36.718		-96.780
	(5.081)	(3.938)	(6.23)	(6.777)	(3.465)	(1.476)		(7.613)

Notes: Column (1) is for all stores, columns (2)-(4) are for stores in low, middle and high income areas, respectively. Column (5) shows results with increased consumption shocks. Column (6) caps inventory at 6.2 kg. Column (7) is a variation of column (2) with no size dummies. Column (8) repeats column (1) without correcting for bundling. Standard errors are bootstrapped as described in the text with $N = 30$.

8 Simulating the effect of small-scale stores

In Africa as well as many other parts of the world, an emerging trend is for large grocery chains to open small-scale stores in low-income areas. Traditionally, these areas were served only by a collection of informal stores, and high travel costs meant that consumers had limited access to formal supermarkets - not unlike the situation of "food deserts" in the US (see Marshall and Pires (2018)).

From consumers' perspective, the presence of these stores can reduce transportation costs - but this benefit is only realized if the store carries the product the consumer wants to buy. Since these small-scale stores have limited space, they typically do not carry a full selection of products, and often restrict the package sizes offered. It is common for stores to carry the smallest package sizes based on the notion that consumers look at these stores to satisfy their temporary demand. However, this ignores the trade-offs highlighted by the model above, namely, the trade-off between quantity discounts, fixed costs of purchase, and inventory costs.

What is the impact of increased consumer access, and what would be the welfare-maximizing package sizes for these stores to carry? To study the impact of increased access and associated reduction in transportation costs, I use the estimated model to analyze a counterfactual scenario where consumers' fixed cost of purchase decreases. To simulate the impact of a grocery store that only has space to carry one package size, I study the problem of a social planner choosing the welfare maximizing package size.

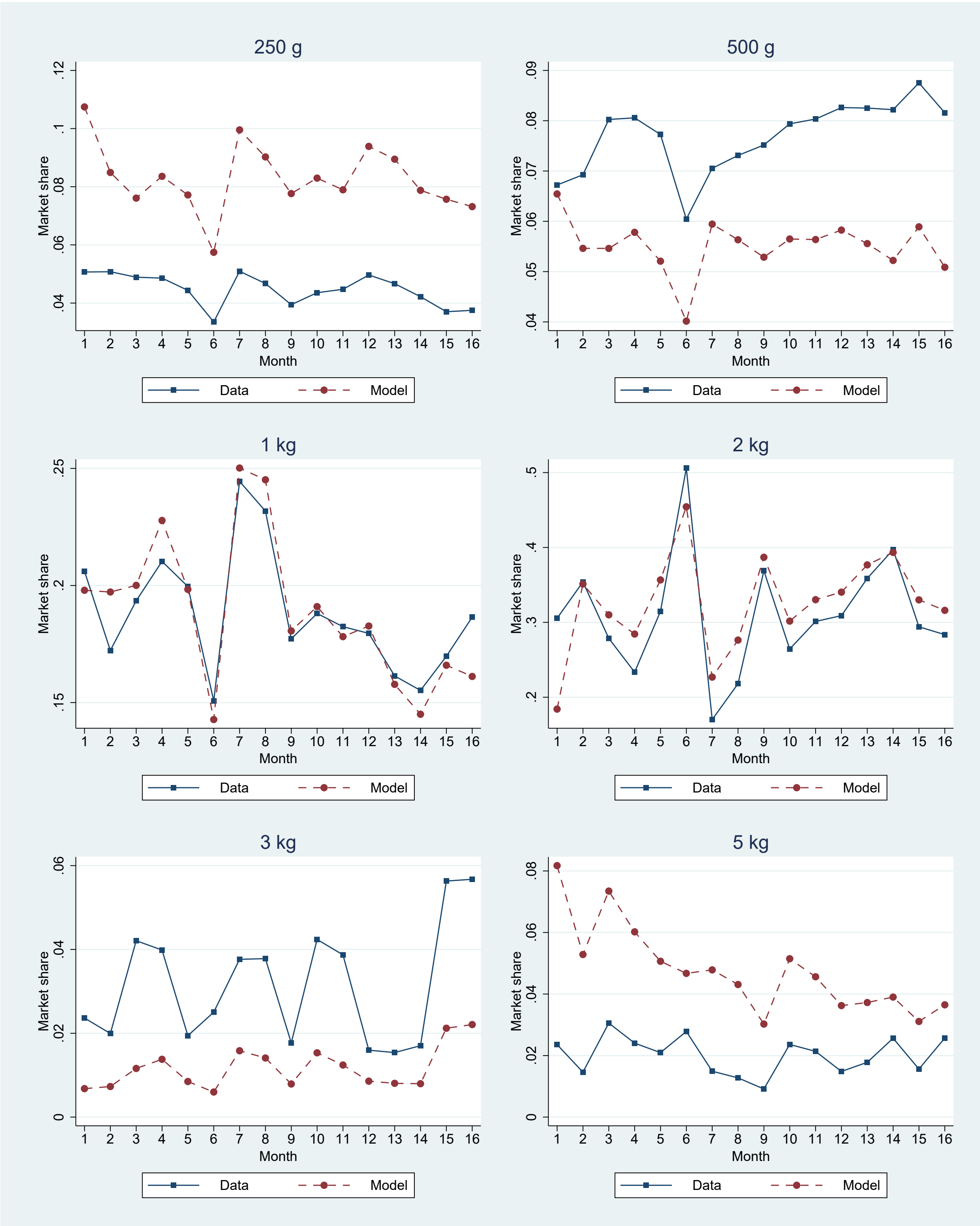
8.1 Counterfactual: reducing transportation costs

In this section, I ask how demand would change if transportation cost was reduced, which I model by reducing the estimated fixed cost of purchase. Specifically, I take the estimated model, and predict purchase probabilities for 16 months with fixed costs reduced by 25, 50, or 75 percent relative to the baseline estimate, or eliminated completely. I compare the results with the baseline predictions.

Table 8 displays the results of these counterfactual experiments. The findings indicate that both small and large package sizes can gain market share. Throughout, the market share of the 2 kg package decreases and that of the 1 kg package increases. The latter makes sense as mean consumption is also around 1 kg: once the fixed cost of purchase is reduced, households make more frequent purchases in order to avoid the storage cost.

In the extreme scenario where all fixed costs are eliminated, there is a large increase in the market share of the largest (5 kg) package size from 3-5 to 17 percent, with a corresponding large decline for the smallest sizes. This substitution reflects consumers taking advantage

Figure 6: Model fit, low-income areas



Notes: Market shares of different sizes observed in the data and predicted by the dynamic model. 100 simulated consumers for each store. Stores are in low-income areas.

Table 8: Market shares when reducing the fixed cost of purchase

	Baseline	Base*0.75	Base*0.50	Base*0.25	Zero
Low-income area					
No purchase	0.27	0.17	0.10	0.08	0.02
250 g	0.09	0.09	0.10	0.11	0.07
500 g	0.06	0.08	0.11	0.14	0.11
1 kg	0.23	0.28	0.31	0.32	0.22
2 kg	0.30	0.31	0.30	0.25	0.37
3 kg	0.01	0.00	0.01	0.02	0.04
5 kg	0.04	0.06	0.07	0.09	0.17
High-income area					
No purchase	0.19	0.13	0.09	0.08	0.02
250 g	0.11	0.11	0.11	0.13	0.09
500 g	0.16	0.17	0.18	0.20	0.14
1 kg	0.28	0.27	0.28	0.27	0.20
2 kg	0.24	0.27	0.26	0.21	0.31
3 kg	0.00	0.00	0.01	0.01	0.06
5 kg	0.01	0.06	0.07	0.09	0.17

Notes: Each column corresponds to a different scenario where the consumer's choice set is restricted to the given size (or the outside option). The simulations span a period of 16 months, with 50 individuals per store.

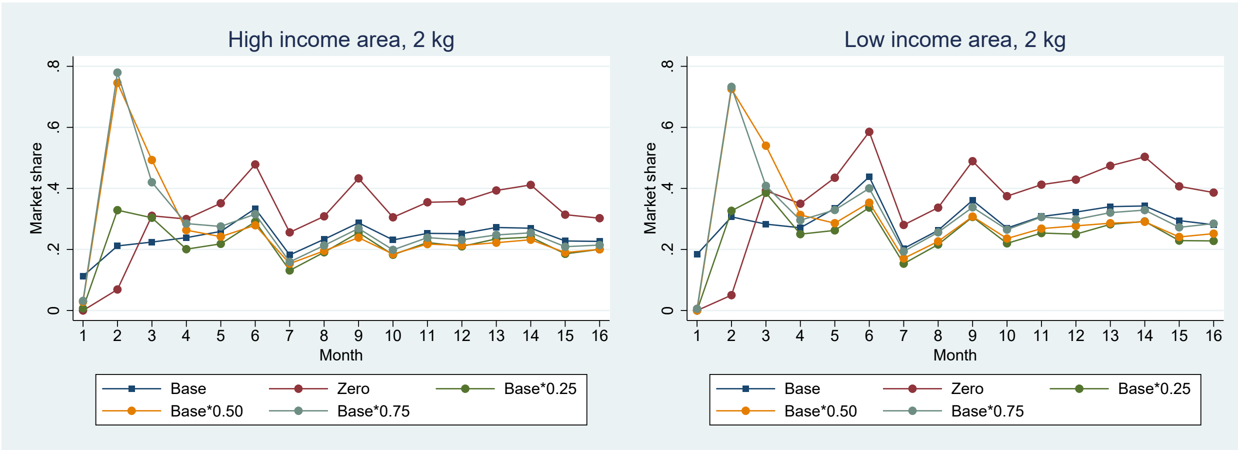
of the quantity discount offered by the larger packages when this does not require them to incur additional transportation costs.

The results also show important dynamics of the substitution between sizes when fixed costs are reduced. Figure 7 shows the simulated market shares over time for the 2 kg package size separately for low and high income areas, and Figures A.23 and A.24 in the Appendix show the patterns for all package sizes. These figures reveal a decline of the smallest sizes relative to the baseline in the first 3 months, with a large increase in the largest sizes during the same period. This raises the average inventory of the households, and since purchase depends on inventory, demand for the larger packages eventually returns to the baseline level. After 3 months, demand for the 2 kg and higher package sizes is similar to baseline levels. At the same time, there is a clear and sustained increase in demand for the 1 kg package, which on average indicates households purchasing about the quantity they intend to consume. These patterns seem to be more pronounced in low income areas, where the estimated fixed cost was higher.[18]

Table A.25 shows the associated consumption and inventory changes under these coun-

[18] Clearly, care should be taken in interpreting the purchase probabilities presented above. This exercise does not take into account potential longer term consumption effects, such as households using more detergent as a result of increased access.

Figure 7: The effect of reducing the fixed cost of purchase



Notes: 50 simulated consumers for each store.

terfactuals. In the no cost scenario, inventory held by the households increases by up to a factor of 5 compared to the baseline. Note however that I did not impose any constraint on inventory accumulation. As such, this overlooks factors like limited storage space, a preference for fresh products to avoid spoilage, etc. Inventory decisions by the households may involve additional considerations than those in the model - particularly for inventory levels not observed in the data.

8.2 A planner's problem: Only one size in the stores

In this section, I use the above estimates to answer the following question: Assuming that a store can carry only one size of laundry detergent, what size would a social planner choose?[19]

Specifically, I consider a social planner problem where the planner chooses one package size to maximize the sum of households' utility based on Equation (2). Let $V_h(\sigma_1|x)$ denote the solution of problem (2) for consumer h given that only products of package size x are available (as well as the outside option of not buying the product). The planner solves

$$\max_{x \in \{0.25, 0.5, 1, 2, 3, 5\}} \sum_h V_h(\sigma_1|x)$$

where maximization is over one of the package sizes currently being sold. To compute $V_h(\sigma_1|x)$, I calculate consumers' utility streams over a period of 16 months.

I consider the planner's problem separately by income area in order to model the pos-

[19]The social planner perspective on the choice of package size could be different from the question of which package size maximizes profits. On the latter, see Björnerstedt and Verboven (2016) in the context of merger analysis.

sibility that a store could decide to tailor the package size to local demand. The potential value of such customization has recently been highlighted by Klopack (2022) in the context of US fast-food chains.

The results are in Table 9. Each column corresponds to a different scenario with the consumer's choice set restricted to the specific size (or the outside option). The results show the average monthly values of consumption, inventory, purchase probability and utility level. (The results for the middle-income area are very similar to those for the low-income area, and are displayed in the Appendix.)

In the data, the market share of the smallest, 250 g package is 4.22% in low-income areas. According to the corresponding simulation, this market share would increase to 48% if this size was the only one available to the consumer (i.e., on average 52% of consumers would choose the outside option in a given month). The counterfactual increase for the largest, 5 kg package is much smaller: from 2.10% observed in the data to 17%. The counterfactual market share of this package is smaller in high-income areas, 14%, reflecting lower observed consumption and preferences towards smaller sizes.

Counterfactual purchase probabilities are generally higher in the high-income area for 1 kg and smaller sizes. At the same time, the purchase probabilities are lower for 2 kg and larger sizes. The increase in utility from the smallest to the largest package size is more pronounced in low-income areas. If the only consideration is expected utility, the 5 kg package would be the welfare-maximizing option in both areas. This would be the optimal size for a planner limited to offering only one size to consumers.

According to this model, offering only the smallest package size is not the optimal choice in either income area (in fact, offering only the 250 g package delivers the lowest average utility in both income areas). This observation is in line with the results of Section 7.2 which indicated that many households only purchased small packages as part of a bundle that allowed them to save money when buying a larger total quantity.

It is worth pointing out that if only the 250 g or the 500 g product is offered, both median and mean consumption is lower than what is observed in the data. The reason is that these quantities are smaller than the mean monthly consumption. Since this exercise maintains the discrete choice assumption that consumers purchase only one product (once a month), consumption must necessarily decrease. Would it be reasonable to assume that, in this case, the consumer would prefer to purchase these smaller package sizes multiple times rather than opting for a larger size? The estimated fixed costs help answer this: based on Table 7, the fixed cost of purchasing N packages of size X is almost always higher than the fixed cost of purchasing a single package of size NX. This suggests that a consumer would not want to make multiple small purchases.

What would be the welfare maximizing package size if the fixed cost of purchase went down? This corresponds to a situation where opening small stores makes access to the product easier, but at the same time, they can carry only one package size. To answer this question, I repeat the above exercise, but decrease the fixed cost of purchase by 25 percent for each package size. Table 10 shows the results. (In Table A.27 in the Appendix, I show corresponding results when the fixed costs are set to zero.)

In this case, there are important changes. First, there are very little differences in expected utility between the 2 kg, 3 kg and 5 kg sizes for low-income areas. Second, there is also little difference in households' associated inventory. The main difference is how frequently households wish to purchase the product. The average household would purchase the 2 kg package 10 times over the simulated 16-month period, the 3 kg package 6 times, and the 5 kg package 4 times. With lower fixed costs, there is an 8 percentage point increase in purchase probability compared to Table 9 if the largest package is offered, and purchase probability doubles if the 3 kg package is offered. When access improves, consumers are even more likely to buy the largest packages.

For high-income areas, the 2 kg package size provides the highest expected utility, in contrast to the 5 kg package in Table 9.

In summary, the welfare maximizing package size offered in a small store may be different from the most popular package size purchased in a regular store. Considerations such as ease of access, frequency of purchase, or storage costs all enter households' decision about which package size to buy. In addition, these results appear to show that offering one of the larger package sizes, rather than one of the smaller sizes, would yield higher consumer utility.

Table 9: Counterfactual simulations

	Low-income area						High-income area					
	250 g	500 g	1 kg	2 kg	3 kg	5 kg	250 g	500 g	1 kg	2 kg	3 kg	5 kg
Consumption												
Average	34.30	54.63	82.71	90.37	90.09	93.35	37.37	55.50	75.03	79.87	78.22	80.79
Median	25	50	97.59	88.56	89.11	92.88	25	50	79.16	76.87	75.02	77.69
Inventory												
Average	10.08	13.12	45.35	167.15	119.40	250.91	19.30	27.55	79.44	217.82	116.53	235.18
Median	0	0	0	92.85	130.20	265.49	0	0	21.18	105.25	117.21	241.82
Purchase probability												
Average	0.48	0.46	0.45	0.35	0.22	0.17	0.49	0.47	0.44	0.35	0.19	0.14
Median	0.48	0.46	0.43	0.32	0.18	0.15	0.50	0.47	0.44	0.31	0.16	0.11
Utility level (expected)												
Average	156.53	161.66	165.76	176.30	169.71	183.80	123.35	126.34	126.57	134.11	123.63	133.15
Median	154.80	160.05	163.82	169.59	166.86	181.07	119.66	122.79	122.49	127.37	120.46	130.03

Notes: Each column corresponds to a different scenario where the consumer's choice set is restricted to the given size (or the outside option). The simulations span a period of 16 months, with 50 individuals per store. Consumption and inventory are measured in 10 g.

Table 10: Counterfactual simulations with reduced fixed cost of purchase

	Low-income area						High-income area					
	250 g	500 g	1 kg	2 kg	3 kg	5 kg	250 g	500 g	1 kg	2 kg	3 kg	5 kg
Consumption												
Average	34.98	55.92	87.12	98.53	98.47	98.57	38.18	57.22	80.00	86.96	82.58	87.33
Median	25	50	98.22	98.59	98.52	98.65	25	50	82.68	83.62	79.04	84.02
Inventory												
Average	11.78	24.35	191.99	636.25	627.29	648.66	21.67	41.89	219.67	643.58	355.51	685.59
Median	0	0	147.45	674.34	634.58	640.19	0	0	174.38	710.12	236.62	678.95
Purchase probability												
Average	0.53	0.53	0.73	0.65	0.42	0.25	0.56	0.56	0.75	0.66	0.30	0.26
Median	0.53	0.55	0.74	0.62	0.36	0.20	0.56	0.56	0.76	0.63	0.27	0.19
Utility level (expected)												
Average	157.43	163.75	182.14	227.75	225.68	227.84	123.94	127.52	136.80	169.18	144.30	171.21
Median	155.88	162.42	180.22	225.76	224.49	225.71	120.25	123.76	133.22	167.22	140.72	168.76

Notes: Simulations for the base*0.75 case. Each column corresponds to a different scenario where the consumer's choice set is restricted to the given size (or the outside option). The simulations span a period of 16 months, with 50 individuals per store. Consumption and inventory are measured in 10 g.

9 Conclusion

This paper analyzes consumer choices between different package sizes of a storable product in South Africa. It provides a comprehensive dynamic demand estimation under nonlinear prices and, in addition, temporary promotions that often target only specific sizes. These two facts combined can result in situations where consumers are better off buying two smaller packages instead of a larger one, which can lead to misleading conclusions regarding consumers' preferences for small package sizes in developing countries.

The results are based on high-frequency scanner data for an entire product category, and geo-coded store locations linked to consumer demographics. The dataset includes both rural and low-income populations, and it is complemented with a survey that directly collects information on consumption and inventory to improve the identification of several dynamic parameters of the model.

My first finding is that accounting for bundling opportunities has important effects for the interpretation of observed market shares. I estimate that, on the median market, there are over 3 times as many households who purchase 5kg of the product as households who purchase a single 5 kg package (the largest available package size). These findings run against the common belief that in similar environments households often purchase products with the lowest package price rather than the lowest unit price.

Currently grocery chains are expanding in low-income neighborhoods throughout Africa by opening small-scale stores with limited space that sell a limited number of items. To study which package size is efficient, I solve the planner's dynamic problem when consumers are faced with the choice of either buying a single available package size or not buying anything. I find that if a store is restricted to offering a single package size, offering the 3 kg package would provide the highest consumer utility. This is the second largest package size sold in the data – by contrast, offering the smallest (250 g) package size would deliver one of the lowest values of consumer utility.

This paper also contributes to the development economics literature, where there is long-standing interest in interventions designed to motivate households to save. In my data, over a year, a household that consumes the average amount of detergent and always buys the largest package will spend 11.08 USD less than if it always bought the smallest package. These are remarkably large savings on a single product category that typical households regularly purchase in a grocery store. From a policy perspective, this suggests that making these saving opportunities salient, as is often done in other contexts, can be an effective tool to increase welfare.

References

Aguirregabiria, Victor. 1999. "The dynamics of markups and inventories in retailing firms." *The Review of Economic Studies*, 66(2): 275–308.

Arcidiacono, Peter, and Robert A Miller. 2011. "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity." *Econometrica*, 79(6): 1823–1867.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Björnerstedt, Jonas, and Frank Verboven. 2016. "Does merger simulation work? Evidence from the Swedish analgesics market." *American Economic Journal: Applied Economics*, 8(3): 125–164.

Cohen, Andrew. 2008. "Package size and price discrimination in the paper towel market." *International Journal of Industrial Organization*, 26(2): 502–516.

DellaVigna, Stefano, and Matthew Gentzkow. 2019. "Uniform pricing in us retail chains." *The Quarterly Journal of Economics*, 134(4): 2011–2084.

Erdem, Tülin, Susumu Imai, and Michael P Keane. 2003. "Brand and quantity choice dynamics under price uncertainty." *Quantitative Marketing and economics*, 1: 5–64.

Fitzpatrick, Anne. 2017. "Shopping while female: Who pays higher prices and why?" *American Economic Review*, 107(5): 146–149.

Gowrisankaran, Gautam, and Marc Rysman. 2012. "Dynamics of consumer demand for new durable goods." *Journal of political Economy*, 120(6): 1173–1219.

Hendel, Igal, and Aviv Nevo. 2006. "Measuring the implications of sales and consumer inventory behavior." *Econometrica*, 74(6): 1637–1673.

Hendel, Igal, and Aviv Nevo. 2013. "Intertemporal price discrimination in storable goods markets." *American Economic Review*, 103(7): 2722–2751.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the top of mind: How reminders increase saving." *Management Science*, 62(12): 3393–3411.

Klopack, Ben. 2022. "One Size Fits All? The Value of Standardized Retail Chains." *RAND Journal of Economics*, forthcoming.

Lal, Rajiv, and Ram Rao. 1997. "Supermarket competition: The case of every day low pricing." *Marketing science*, 16(1): 60–80.

Marshall, Guillermo, and Tiago Pires. 2018. "Measuring the impact of travel costs on grocery shopping." *The Economic Journal*, 128(614): 2538–2557.

Nevo, Aviv. 2001. "Measuring market power in the ready-to-eat cereal industry." *Econometrica*, 69(2): 307–342.

Newey, Whitney K, and Kenneth D West. 1987. "Hypothesis testing with efficient method of moments estimation." *International Economic Review*, 777–787.

Pesendorfer, Martin. 2002. "Retail sales: A study of pricing behavior in supermarkets." *The Journal of Business*, 75(1): 33–66.

Rust, John. 1987. "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica: Journal of the Econometric Society*, 999–1033.

Ujhelyi, Gergely, Somdeep Chatterjee, and Andrea Szabó. 2021. "None of the Above: Protest Voting in the World's Largest Democracy." *Journal of the European Economic Association*, 19(3): 1936–1979.