

NBER WORKING PAPER SERIES

TIMING AND QUANTITY OF CONSUMER PURCHASES AND THE CONSUMER  
PRICE INDEX

Rachel Griffith  
Ephraim Leibtag  
Andrew Leicester  
Aviv Nevo

Working Paper 14433  
<http://www.nber.org/papers/w14433>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
October 2008

Financial support for Griffith and Leicester from the ESRC through the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS (CPP) is gratefully acknowledged. Griffith and Leicester, University College London and Institute for Fiscal Studies, Leibtag, US Department of Agriculture Economic Research Service, Nevo, Northwestern University and National Bureau of Economic Research. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research or the US Department of Agriculture.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by Rachel Griffith, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Timing and Quantity of Consumer Purchases and the Consumer Price Index  
Rachel Griffith, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo  
NBER Working Paper No. 14433  
October 2008  
JEL No. D12,E31,L11,L16

### **ABSTRACT**

A common approach to measuring price changes is to look at the change of the expenditure needed to purchase a fixed basket of goods. It is well-known that this approach suffers from problems and creates several biases in the measurement of price changes faced by consumers. Substitution and outlet bias, two commonly studied concerns, are both driven by consumer choices of what and where to buy. However, consumers also make other choices, including how much and when to buy. We discuss the implications of consumers' timing and quantity decisions have on standard practices of computing of computing a price index.

We use household-level data on quantities purchased and prices paid to construct a measure of the savings made by consumers' optimizing behaviour in the purchase of food. In particular, we compare the prices actually paid by the consumers to various alternatives that do not allow for substitution. Our analysis suggests that the average consumer makes significant, and comparable in magnitude, savings from the four dimensions of choice that we study. Furthermore, our data suggests significant heterogeneity in consumer behavior, and that this behavior is correlated with demographics. Our findings suggest that ignoring timing and quantity decisions, when computing a price index, can generate biases on the order of magnitude of substitution and outlet biases.

Rachel Griffith  
IFS  
7 Ridgmont Street  
London WC1E 7AE  
UK  
rgriffith@ifs.org.uk

Ephraim Leibtag  
U.S. Department of Agriculture  
FRED-ERS-USDA  
1800 M Street, NW Rm. N2137  
Washington, DC 20036  
eleibtag@ers.usda.gov

Andrew Leicester  
Institute for Fiscal Studies  
7 Ridgmount Steet  
London WC1E 7AE, UK  
andrew\_1@ifs.org.uk

Aviv Nevo  
Department of Economics  
Northwestern University  
2001 Sheridan Road  
Evanston, IL 60208-2600  
and NBER  
nevo@northwestern.edu

# 1 Introduction

One of the major challenges faced by applied economists is the measurement of price changes. The middle part of 2008 saw significant increases in the prices of food, vehicle fuel and household energy, for example. How do we measure how much prices have increased?

In practice, the most common approach to measuring price changes is to look at the changes of the expenditure needed to purchase a fixed basket of goods. This approach dates back at least to the early nineteenth century (Diewert, 1993), and has a large number of variations. It is well-known that this approach suffers from problems and creates several biases in the measurement of price changes faced by consumers.<sup>2</sup> Common concerns include substitution bias, outlet bias, quality change, treatment of new products, and heterogeneity across consumers.

Substitution and outlet bias, two commonly studied concerns, are both driven by consumer choices of what and where to buy. However, consumers also make other choices, including how much and when to buy. Many products are sold at non-linear prices: the price per unit, say an ounce, is typically lower for larger pack sizes. Thus, a consumer has a choice of not just what product to buy and where to buy it, but also what size to buy (and what price to pay). Similarly, many products have temporary price reductions – sales – that allow the consumer to purchase more today and potentially stockpile for future consumption. Both these dimensions of choice have been mentioned as potential sources of bias in price indices (e.g., Feenstra and Shapiro, 2003; Triplett, 2003; and references therein), but are discussed less often than the first two biases.

---

<sup>2</sup> See for example, Boskin et al (1998), Deaton (1998), Diewert (1998).

The standard approach to price measurement does not properly account for the fact that consumers do not face a single price, but rather face a distribution of prices (Baye, 1985; Reinsdorf, 1994). The variation in prices can be across goods, stores, size and over time. Different consumers will react to this variation differently, depending on their preferences and costs. For example, some consumers might have lower travel costs and therefore be more likely to shop at a range of different stores. Other consumers might have lower storage and transport costs and will therefore take greater advantage of quantity discounts and temporary price reductions. Accurate description of price changes has to move beyond the “average” consumer, and explore heterogeneity across consumers (Pollak, 1998).

If the relative price of different goods remains fairly constant, there is probably little variation in inflation across households and so distributional issues are not so important. In periods of high inflation, and in particular when inflation is driven by only certain commodities, heterogeneity across households is likely to be more important. For example, Deaton and Muellbauer (1980, Table 7.1) report that during 1975-76, when inflation in the UK was 15 percent, the rate for the poor was two points higher than for the rich.<sup>3</sup>

In this paper we contribute to the literature on the empirical relevance of these issues. We use household-level data on quantities purchased and prices paid to construct a measure of the savings made by consumers’ optimizing behaviour in the purchase of food. We examine the savings made by consumers timing their purchases to buy on sale, buying larger package sizes, buying generic brands or switching between alternative outlets. Furthermore, we document how

---

<sup>3</sup> Crawford (1996) and Crawford and Smith (2002) report similar more recent results.

this behaviour, and the implied savings, varies by demographics including income, family size, and employment status.

We observe data, collected by a marketing firm, on all food purchases brought into the home for a large, nationally representative, sample of UK households in 2006. Compared to previous studies, our data is more comprehensive – not limited to a subsample of goods – and more detailed regarding the brand, package size, location, whether on sale and time of purchase. For each purchase we know exactly what was bought (as measured by the bar code), the price paid and quantity purchased, the purchase date and the store it was purchased in. We also observe household demographics.

Using these data, we are able to compare the prices actually paid by the consumers in the sample to various alternatives. For example, to measure the savings from the timing of purchases we can compare the actual expenditure to the expenditure of a consumer who purchases the same bundle (in terms of products and outlets), but did not buy when the item was on sale. We show how the savings vary with income, age and employment status, as well as other demographics.

Our analysis suggests that the average consumer makes significant, and comparable in magnitude, savings from the four dimensions of choice that we study. This demonstrates, on a larger scale, the savings associated with consumer behaviour, found by previous work (e.g., Hendel and Nevo, 2006a, for saving from timing purchases, and Hausman and Leibtag, 2007, for saving from the availability of Wal-Mart). Furthermore, our data suggests significant heterogeneity in consumer behaviour. For example, some households buy a significant fraction of their food on sale, others rarely do. This behaviour is correlated with demographics.

In the next section we describe the data that we use and present some preliminary descriptive statistics. In section 3 we consider the potential and actual savings that households make by

buying on sale, larger pack sizes, generic brands and in specific stores. In section 4 we compare the levels of savings and a final section makes some concluding remarks.

## **2 Data**

The data come from the TNS Homescan panel ([www.tnsfres.com](http://www.tnsfres.com)), a representative consumer panel of around 25,000 households resident in Great Britain. Households are recruited from those who had previously responded to a large paper-based consumer survey. Respondents that match the demographic characteristics required for a new household are sent a postal invitation to participate. Participants are rewarded with points redeemable for a range of products and services (though limited to items that should not directly affect grocery consumption patterns). Participating households are issued an electronic hand held scanner in their homes and asked to scan the barcodes of all grocery purchases (foodstuffs, alcohol, bathroom products, medicines, pet food etc.) that come into the house.

Information on purchases is downloaded once a week by TNS. In addition, households mail till receipts to TNS. These are used to collect prices and verify the information entered by the households. Information on loose weight items such as vegetables and fruit is collected by households scanning barcodes in a book and keying in the weight data. Purchases from all store types (supermarkets, corner stores, online, local speciality shops etc.) are covered by the survey. For larger stores, the exact store of purchase is recorded; for smaller stores only the store type is known. The data includes information on the characteristics of the product including price, brand, pack size, whether the item was bought on promotion and a number of characteristics of the product. Demographic information about the household is collected by an annually-updated telephone survey.

Our focus in this paper is on four particular types of behaviour - sales, size, generic and outlet. How do we measure each of these?

TNS gathers information on price reductions and promotions from a variety of sources, including the receipts sent in by households, fieldwork, and directly from the stores. Promotions typically take two forms: price promotions (50% off, £1 off) and quantity promotions (buy one get one free, 50% extra volume).

Package size is reported directly in the data on product characteristics. In order to compare across a wide range of food types we look at how price varies across the quintiles of the package size distribution within each food category.

Generic (store) brands are also recorded directly as part of the product characteristic information. The detail allows us to distinguish “economy” versions of generic items from “regular” generics and “luxury” generics. We discuss this more in the next section.

Information on stores is collected via the households, who scan a barcode representing each store fascia prior to entering the details of each shopping trip. Households report to TNS the specific store when they sign up (and for most large shopping trips) and this is matched against the fascia barcode scanned for each trip. For corner and local stores, the specific shop location is typically not recorded.

For the analysis below, we use data for the calendar year 2006. We observe expenditure for 23,877 households on purchases in 189 categories, effectively covering all food and beverage purchases. These households make a total of 5.6 million separate shopping trips. On average a single shopping trip involves the purchase of 4.2 items and £6.08 in expenditure. The average

duration between shopping trips (excluding multiple trips within the same day) is 4 days (with a median of 3 days).

The demographic characteristics that we focus on include family composition, how often the household shops by car, what the most common mode of transport is for shopping and household income. The distribution of these is shown in the following tables.

**Table 1: Distribution of household income**

Household income	Observations	Share of sample	Share of non-missing sample
£0 - £9,999	2,052	8.6%	13.4%
£10,000 - £19,999	4,344	18.2%	28.5%
£20,000 - £29,999	3,545	14.9%	23.2%
£30,000 - £39,999	2,309	9.7%	15.1%
£40,000 - £49,999	1,434	6.0%	9.4%
£50,000 - £59,999	787	3.3%	5.2%
£60,000 - £69,999	340	1.4%	2.2%
£70,000 +	448	1.9%	2.9%
Missing / unknown	8,618	36.1%	–

Source: Authors' calculations from 2006 TNS sample

Income information began to be collected by the market research firm in 2006 and in our data around one-third of households have not had their incomes recorded. For other households, incomes are recorded gross at a household level in one of eight bands. The distribution of income in these data are similar to those found in other UK surveys, for example, UK Expenditure and Food Survey (EFS), although overall average income is slightly lower.<sup>4</sup>

Household type is defined using information on each of the household members, and here we also see similar to patterns to EFS data for the percentage of households with at least one

---

<sup>4</sup> The TNS data includes demographic weights which correct for potential biases in recruiting and retaining some household types. We do not use these weights in this analysis, but control for observed demographic characteristics when looking at the savings from different channels in the next sections.



child, although the market research data does seem to contain somewhat fewer households headed by a pensioner and fewer single adult households.

**Table 2: Distribution of household type**

Household type	Observations	Share of sample
Single pensioner	1,940	8.1%
Pensioner couple	2,246	9.4%
Single adult	2,209	9.3%
Couple without children	2,835	11.9%
Other childless household	5,778	24.2%
Lone parent	1,008	4.2%
Couple with children	4,516	18.9%
Other household with children	3,345	14.0%

Source: Authors' calculations from 2006 TNS sample

**Table 3: Frequency of shopping by car**

Frequency	Observations	Share of sample
5 or more times / week	495	2.1%
3 – 5 times / week	3,309	13.9%
1 – 2 times / week	14,495	60.7%
At least once a month	1,897	7.9%
Less than once a month	1,541	6.5%
Never	2,140	9.0%

Source: Authors' calculations from 2006 TNS sample

These figures match very closely to other survey data on how people travel for shopping. Data from the 2002/3 National Travel Survey (Department for Transport, 2005) showed that around 79% of households usually did their main food shopping by car or taxi, 12% on foot and 10% by other means, remarkably similar to the figures from the 2006 TNS sample shown above.

**Table 4: Most commonly used method of transport for shopping**

Transport	Observations	Share of sample
Car or taxi	19,056	79.8%
Public transport	1,307	5.5%
Foot	2,964	12.4%
Other	550	2.3%

Source: Authors' calculations from 2006 TNS sample

Unsurprisingly, the car is overwhelmingly the most popular form of shopping transport with eight in ten households using it as their main mode of transport for shopping. This is roughly the same proportion of households that record using their car to go shopping at least once a week.

### **3 Household savings**

In this section we quantify the savings that households can and do make by purchasing on sale, buying in bulk (at a lower per unit price), buying generic brands and choosing outlets. We document the relative importance of these various dimensions, which may suggest where future work on price indices should focus efforts. One measure of the importance of the different dimensions is the amount saved by households from each dimension, ignoring the costs they may incur. For example, how much more would consumers have paid had they *not* purchased on sale, ignoring the costs of carrying inventory from one period to the next? Another important dimension is the degree of heterogeneity across households in any particular behaviour. We discuss the potential biases that may arise in a price index due to consumer behaviour.

#### **3.1 Sales and Stockpiling**

Many grocery items exhibit price variation over time. If we focus on a narrowly defined product (a particular brand and size sold at a particular store), much of the price variation over time is due to temporary price reductions. It has been documented that for many products consumers respond to this price pattern by stockpiling for future consumption (see Hendel and Nevo 2006a, Boizot et al (2001) and references therein).

When buying on sale a consumer faces a tradeoff between paying a lower price today for a product that will be consumed in the future, and incurring a storage cost until the product is consumed. The benefits from buying on sale depend on future consumption needs and on future

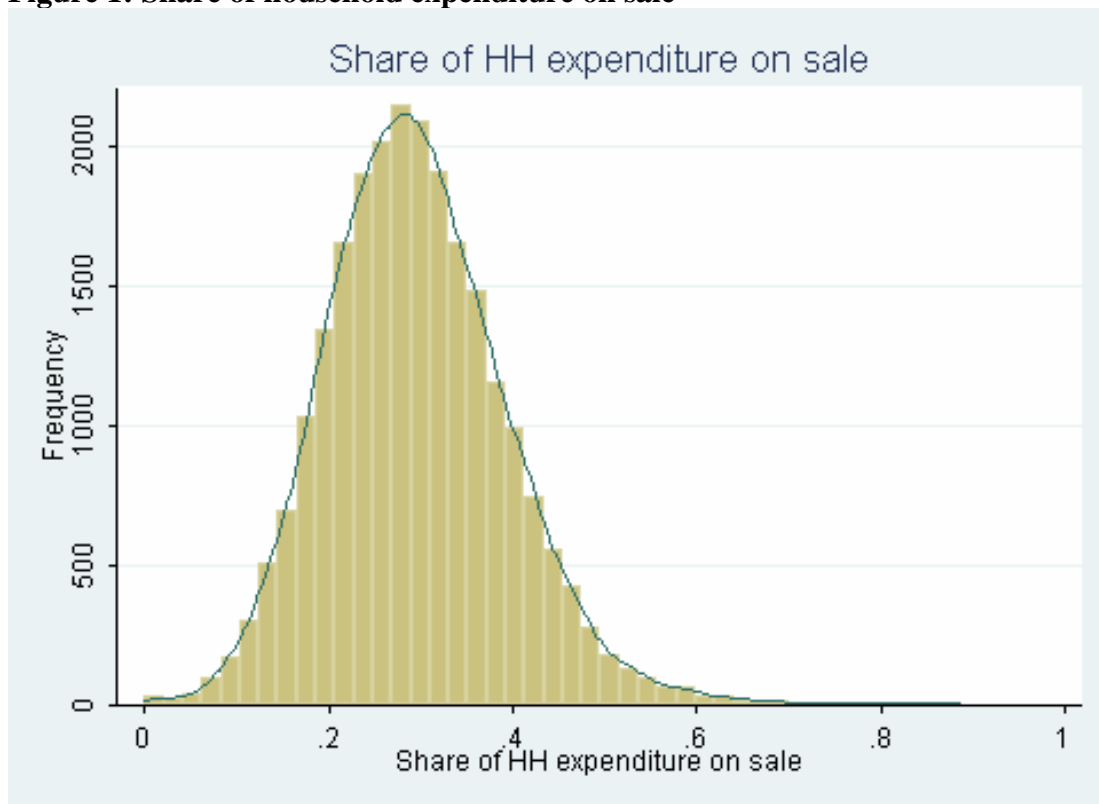
prices. Different consumers will make different choices, and a given consumer will make different choices for different goods.

In the following subsections we present evidence on heterogeneity across consumers in the propensity to purchase on sale. We then provide measures of the savings consumers obtain and use these to motivate the importance of controlling for this dimension of consumer choice. Finally, we discuss the implications for measurement of price changes.

### 3.1.1 *How much do households buy on sale?*

The average UK household spends around 29.5% of their total annual expenditure on items that are “on sale.” As expected, there is considerable variation across households (see Figure 1), with the household at the 10th percentile purchasing 17.7% and the 90th percentile purchasing 42.0%.

**Figure 1: Share of household expenditure on sale**



Note: Histogram of the share of total observed household expenditure in 2006 that was purchased on sale; sample includes 23,877 households.

Some of the variation across households is explained by observed demographics. In particular, when we regress the share of expenditure on sales on family type we see that retired households tend to buy less on sale (single pensioners buy 3.2% less on sale than single young households, while pensioner couples buy 2.9% less on sale than young childless couples, and over 5% less than couples with children). Families with children tend to buy more on sale than childless families and households with fewer adults.

Households that shop by car buy approximately 2% more of their food on sale than households that shop by public transport or on foot, or that shop less frequently by car. This is consistent with pensioners buying less on sale since they are less likely to shop by car.

The fraction of buying on sale is related to income in a non-monotonic way: the lower income households buy the least on sale and the middle income households buy the most. This is perhaps expected – low income households do not have the flexibility, in terms of storage, transport or liquidity, to take advantage of sales. On the other hand, the highest income households have a lower marginal utility of income and a higher value of time, and so do not find it worthwhile to take advantage of sales.

Overall, however, observed demographics explain less than 10% of the variation in the propensity to purchase on sale.

### ***3.1.2 How much do consumers save by buying on sale?***

We now turn to the question of how much consumers save by buying on sale. We examine savings made on all food products. As described in Section 2 there are 189 different food categories (e.g. bacon, eggs etc). Within each group there are many separate products (bar codes).

We want to compute a single saving figure for each household, which we do in two steps. First, we estimate the saving made by purchasing a specific product on sale using the following regression model:

$$\ln p_{iht} = \beta_j s_{it} + \eta_i + t_{jt} + r_{jr} + e_{iht} \quad (1)$$

where  $i$  indexes bar codes,  $j$  food categories,  $h$  households,  $t$  time and  $r$  region;  $p$  is unit price,  $s$  is a dummy variable indicating that the product was purchased on sale,  $\eta_i$  captures barcode specific characteristics (allowing us to control for differences in product characteristics) and  $e$  is an idiosyncratic error. As the subscripts make clear, we estimate the regression separately for each food category. The propensity to have sales and the discount level differ significantly across categories: some product groups have deep discounts while others barely any discounts. Allowing the coefficients to vary by category captures this heterogeneity. This procedure yields 189 beta coefficients, all of which are negative (sale prices are lower) and all but three are statistically significant at the 5% level (and all but five at the 1% level). Using the estimated coefficients we can compute the discount when buying in sale. This discount varies across food categories from 14% for the 10<sup>th</sup> percentile to 29% for the 90<sup>th</sup> percentile, with an average and median of 22%. The categories with low discounts are fruit fillings (2%), lard (4%) and sugar (10%), while examples of the categories with high discounts are savoury snacks (32%), breakfast cereal (31%) and baked beans (31%).

To compute a household-level savings measure we weight the category-level savings using household specific expenditure weights reflecting the share of purchases of each product category that are purchased on sale and the share of that product category in each households' budget. That is we calculate a household's total savings as:

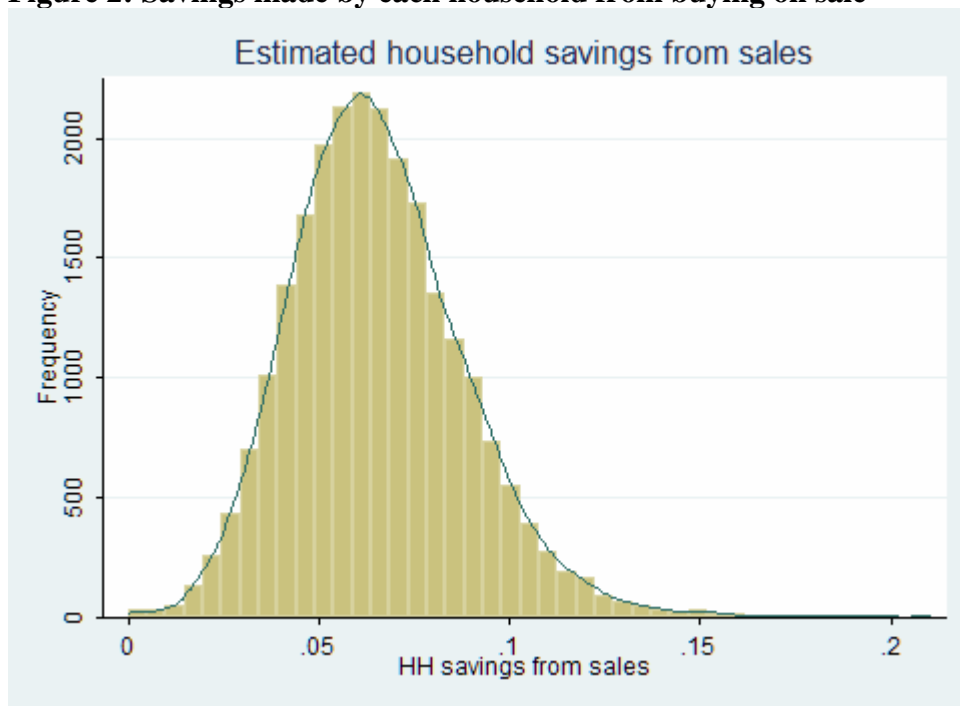
$$SaveSale_h = \sum_{jt} w_{hj} (1 - \exp(\beta_j)) \quad (2)$$

where  $w_{hj}$  is the expenditure of household  $h$  on items in product category  $j$  that were purchased on sale as a share of total expenditure across all categories:

$$w_{hj} = \frac{\sum_{it, i \in j} (\text{sale}_{iht} = 1) \text{expenditure}_{iht}}{\sum_{it} \text{expenditure}_{iht}} \quad (3)$$

This measure varies between 0 and 1 as long as all the betas are negative. It captures the fraction of expenditure the household is saving by purchasing on sale. Figure 2 shows the distribution of savings made by each household by purchasing on sales,  $SaveSale_h$ . Note that the variation across households in the savings measure stems from two sources. As we saw in Figure 1, households vary in the fraction they purchase on sale. Furthermore, households vary in their expenditures on different categories, and the discounts offered by sales differ by categories. Our measure captures both these sources of heterogeneity.

**Figure 2: Savings made by each household from buying on sale**



Note: Histogram of the savings each household made by purchasing on sale in 2006 as defined by equation (2); sample includes 23,877 households.

Households save between 0 and 21% of their annual expenditures, with a mean of 6.5%. This translates into a savings of up to £794 a year, with an average saving of £96 per year.

The amount saved through buying on sale varies by observed demographics in a very similar way to the variation in the proportion of expenditure bought on sale. Most notably, households where the head of household is retired and poorer households save substantially less by purchasing on sale than other households. One notable difference between the share purchased on sale and the extent of savings is that wealthier households appear to purchase more on sale (than poorer households), but do not seem to save much by doing so. However, as was the case with the fraction of purchases on sale, the observed demographics explained relatively little of the variation in the savings measure.

### 3.1.3 *How well does a standard price index capture sales?*

The sales behaviour documented in the previous subsections has several implications for the standard measurement of price indices. First, given the way statistical agencies collect data there is an issue about whether they correctly sample sales. Statistical agencies may over- or under-sample items on sale given the products included in the basket of goods relative to the sample of all products purchased by consumers. We provide evidence that this might be the case. Second, even if the correct products are chosen, and hence the true prices observed, there is an issue regarding what is the correct price index formula. We demonstrate the issue in the context of a simple example. Finally, there is a question of whether the difference between the ideal measure and the index actually computed varies across time and households.

As the previous figures show, sales are a significant factor in the expenditure of most households. In order to examine how the data collected by statistical agencies captures sales we looked at prices of the most commonly purchased (“most popular”) items within each food category, which we took to be an approximation of the set of prices most likely to be sampled by price collectors<sup>5</sup>. We find that the most popular products are *more* often on sale than the “average” product (see Table 5). This may be because the more popular products are typically branded items which go on sale more often, or perhaps because promoted items tend to be purchased more often. This suggests that the ONS data could misrepresent the distribution of prices paid by consumers.

---

<sup>5</sup> We aim to mimic the guidelines given by the UK Office of National Statistics to price enumerators, defining the most popular item as that most frequently purchased within a product group-store combination. The ONS samples food prices from the 6 largest national chain stores: Tesco, Sainsbury, Asda, Somerfield, Morrisons and Marks and Spencer. We consider the population of all other chain and local shops as a seventh “store”.



**Table 5: Frequency of promotion in 2006 TNS data, all and most popular items**

	<i>Share of expenditure (%)</i>		<i>Share of purchases (%)</i>	
	<b>All items</b>	<b>Most popular</b>	<b>All items</b>	<b>Most popular</b>
No promotion	70.35	64.14	71.94	67.61
Price promotion	12.70	14.18	13.37	14.78
Quantity promotion	16.96	21.70	14.69	17.62

Notes: Most popular items are those most frequently purchased in each of seven stores within each of the 189 food categories in the TNS data.

Source: Authors' calculations from 2006 TNS sample.

Putting aside the measurement issue – even if prices are collected in such a way as to accurately capture sale prices – there is an issue of how to compute a correct index in the presence of stockpiling. Consumer stockpiling implies a separation between a purchase-based price index and a consumption-based price index. Standard price indices are purchase-based, some explicitly so, but a utility based cost-of-living index should account for the ability to store the product<sup>6</sup>. To illustrate the difference consider the following example. Suppose a consumer consumes two products, A and B, at equal quantities. Product A always costs \$2, while product B is normally priced at \$2, but goes on sale for one period and is sold at a price of \$1. Normally, the consumer purchases one unit of each product each period and consumes both products in that period.

Suppose the consumer has a storage cost of \$0.25 per unit per period. During the sale of product B the consumer purchases 4 units and consumes one unit each week over the next few weeks. We assume there are no consumption effects: the consumer does not increase consumption in response to the sale price. For the calculation that follows we ignore the discount factor. The consumer saves on each of the units he purchases: for the last unit the consumer pays \$1.00 and stores it for 3 periods at a cost of \$0.75, for a total saving of \$0.25 relative to buying

---

<sup>6</sup> In what follows we assume that the goal is to compute a utility based cost-of-living index. Many statistical agencies explicitly state that their price index is not a cost of living index. However, we follow the standard view that sees a cost-of living index as a way to reason about issues involving the consumer price index.

the product at the regular \$2 price. The consumer, however, will not save from buying additional units because of the storage costs. If the consumer bought a fifth unit on sale the storage costs for 4 periods will exactly equal the savings (and we assume that in this case the consumer will not store the product).

Suppose we want to compute a cost-of-living index for this consumer. We set the base as the prices during non sale periods, so when consuming 1 unit of each product the base is \$4. The true consumer's price index for the period of the sale and the following weeks is  $(2.00+1.00)/4.00=0.75$ ,  $(2.00+1.25)/4.00 = 0.8125$ ,  $(2.00+1.50)/4.00 = 0.875$ ,  $(2.00+1.75)/4.00 = 0.9375$ , and 1.00 for every following week. The observed prices of product A are 2.00 at each period and observed purchases are 1 unit each period. The prices of product B are 2.00, 1.00, and 2.00 for every following period, while the purchasing patterns are 1 unit, 4 units, 0, 0, 0 and 1 unit for every following week. Consumption is constant every week. A standard price index will capture some price reduction in the week of the sale. The exact reduction in the price index depends on the quantity weight used to compute the index. For example, a fixed weight price index, with equal weights, will yield 0.75, for the week of the sale, and 1 for every following week. However, a standard index will not capture the effective drop in the price index in the weeks following the sale, and the problem cannot be "fixed" by adjusting the weights. The problem is that the price the consumer faces is a shadow price. Note, that aggregation across weeks, to construct a monthly price, will also not solve the problem, even if the timing is captured exactly right. In this case, the aggregation will *overestimate* the benefits from purchasing on sale since it will ignore the storage cost.

This simple example illustrates the issues with a standard price index. A more realistic model would allow for consumption effects and uncertainty regarding future prices (see Hendel

and Nevo, 2006b, for such a model). Such a model allows us to compute the unobserved shadow prices faced by the consumer in each period, but is very computationally intense and cannot be estimated on a large scale. Developing more tractable alternatives is a current area of research.

### **3.2 Bulk Discounting and Choice of Package Size**

Many grocery items are sold at non-linear prices. Larger package sizes are sold at higher prices but at lower per unit price. For example, Hendel and Nevo (2006a) report that the regular, non-sale, price of a 24-pack of soft drinks cans cost 2.7 times more than a 6-pack. This implies a discount of over 30 percent in the per period unit price.

A typical treatment of non-linear prices in the price index is to focus on unit values. This corrects for the difference in the actual price between sizes, as opposed to the difference in the size per unit. However, it does not account for the tradeoffs consumers are making (Triplett, 2003).

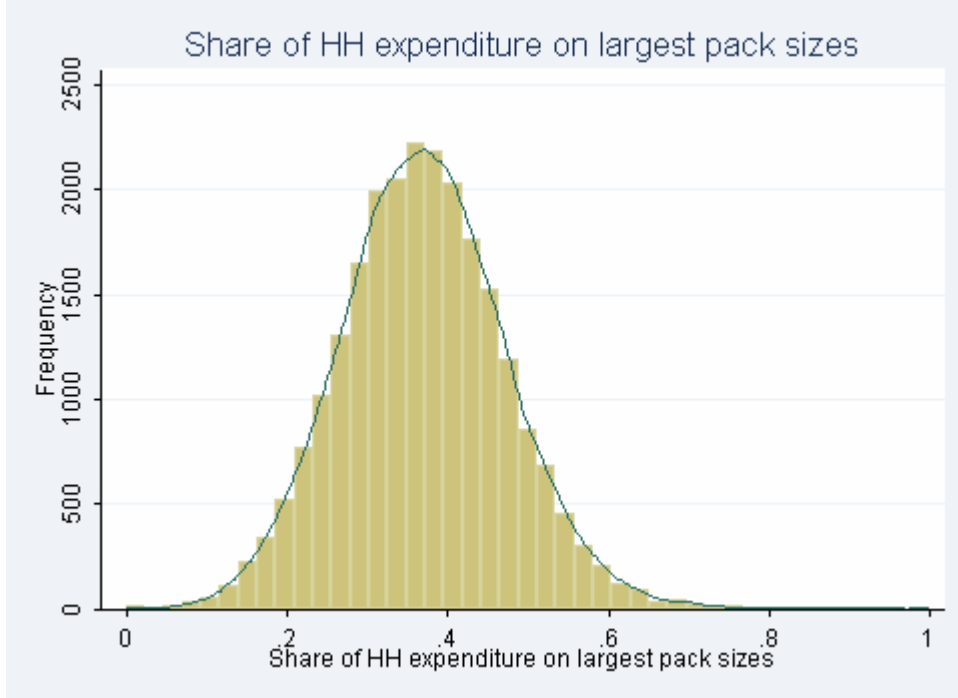
The tradeoffs facing the consumers are similar to those we discussed in the case of sales. Consider a consumer deciding between purchasing a smaller unit and a larger one at a lower per unit price. Unless she can consume the additional quantity before her next store visit the consumer has to weigh the benefits of the lower price with the costs of storing the product longer and the depreciation in the quality of the product. Different consumers will make different choices depending on their marginal utility from income, the cost of storage and transport of the product, and their future consumption needs. A given consumer will make different choices for different products, depending on storage costs, durability, expected consumption and the price schedule.

### ***3.2.1 How much do households buy in bulk?***

As described in section 2 we use quintiles in package size within each food category to measure the extent of bulk discounting. The average household spends 15.8% of their total annual expenditure on the largest package sizes, and 21.2%, 21.3%, 26.8% and 14.9%, on the other sizes from largest to smallest, respectively. There is considerable variation across households in these fractions. For ease of exposition we focus on the two largest quintiles as “bulk” sizes and compare the savings made from purchasing in those quintiles to purchases made in the second largest size group. Figure 3 shows the distribution of the proportion of total household expenditure that was spent of the largest two size quintiles.

Households purchase along the entire range of between 0 to nearly 100% of their groceries in large package sizes, spending on average 37% of their budget in these size groups. Unsurprisingly single person households purchase less in bulk than multi person households. Single pensioners make even less use of bulk discounts, spending on average 2.2% less on large pack sizes than single younger households. Households that shop by car buy in bulk more often and the middle income categories make greater use of bulk discounts. Overall, purchasing behaviour of larger package sizes is similar to purchasing on sale, and in fact the two shares are positively correlated at the household level, with a correlation coefficient of 0.23.

**Figure 3: Share of household expenditure on items in largest size quintiles**



Note: Histogram of the household expenditure shares within the largest 2 size quintiles in 2006. Sample includes 23,877 households.

### 3.2.2 How much do consumers save by buying in bulk?

In order to compute how much consumers save by buying in bulk we first estimate a series of regressions:

$$\ln p_{iht} = \beta_j^s \text{sale}_{it} + \sum_{p=2}^5 \beta_j^p q_i^p + \eta_k + t_{jt} + r_{jr} + e_{iht} \quad (4)$$

where price is unit price,  $q_i^p = 1$  if the pack size of product  $i$  is in the  $p^{\text{th}}$  quartile of all products in product group  $j$ , and zero otherwise, and where  $k$  indexes brand<sup>7</sup>. As in the case of sales, we estimate this regression separately for each product category in order to account for the difference

---

<sup>7</sup> Unlike in the sales case, we cannot use product fixed effects as clearly an individual barcode will always be within a given size quintile. For some product groups, “brand” does not exist as a product characteristic – in these cases we use store fixed effects. There are a few product groups for which there is insufficient size variation to generate quintiles in the size distribution; we ignore these groups in the analysis of this section.

in the degree of non-linear pricing across categories. The average savings are substantial for the larger package sizes. Across all product groups, the average saving on the unit price from buying in the largest size quintile relative to the second is almost 37%, and from the second largest size quintile is 28%.

To compute a household-level savings measure we take the same approach as for sales, but now considering the two largest size quintiles. In particular, we weight the category-level savings from each size using household-size specific expenditure weights. We calculate the household's total savings as

$$SaveSize_h = \sum_{p=4}^5 \sum_{jt} w^p_{hj} (1 - \exp(\beta_j^p - \beta_j^2)) \quad (5)$$

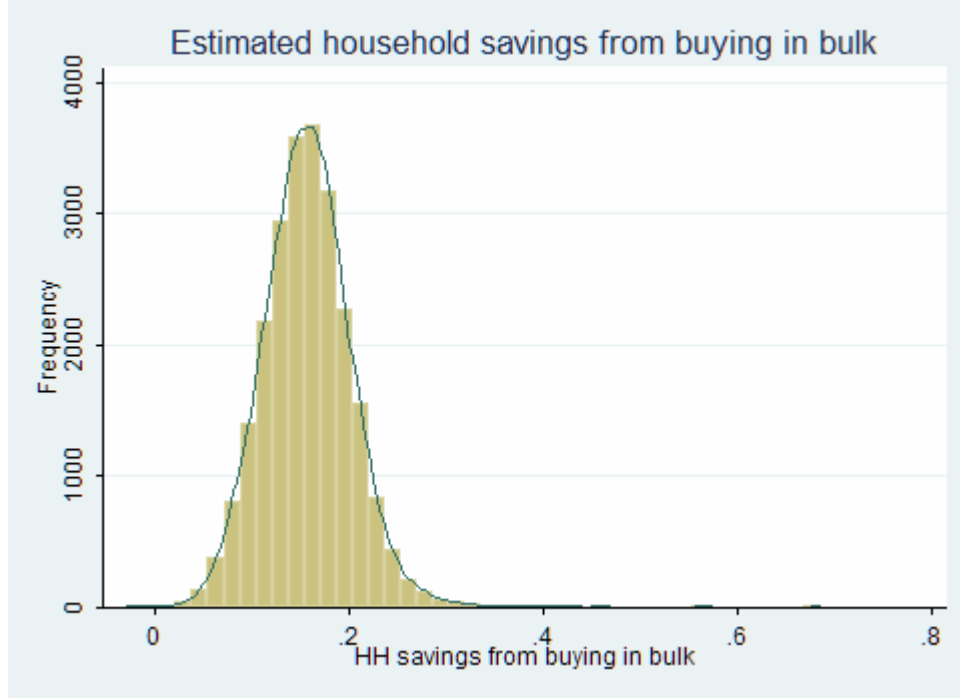
where  $w^p_{hj}$  is the expenditure of household  $h$  on items of size  $p$  in product category  $j$

$$w^p_{hj} = \frac{\sum_{it, i \in j} \text{expenditure}_{iht}^p}{\sum_{it} \text{expenditure}_{iht}} \quad (6)$$

This measure computes the savings from purchasing the largest sizes relative to the second smallest size. If instead we computed the saving relative to the smallest size, the savings would be even larger. We focus on the second largest size because it is the most popular size. On the other hand, the smallest size is not very popular and in many cases only some, not very common, brands are available in this size.

The average household saves 16% of their annual expenditure from buying the largest package sizes, which translates into savings of £224 per year. Figure 4 shows the distribution of savings by households, which varies from just less than 0 to almost 70%<sup>8</sup>.

**Figure 4: Savings made by each household from buying larger package sizes**



Note: Histogram of the savings each household made by purchasing larger pack sizes in 2006 as defined by equation (5); sample includes 23,877 households.

The amount saved varies with household size, with larger households saving more. Households that shop by car, but only infrequently (once a month) save the most in bulk purchases, and households in lower income categories also save more. The demographic variables, however, are able to explain very little of the variation in savings from bulk purchasing, less than 1%.

---

<sup>8</sup> For some product groups the ‘saving’ from buying in the larger pack sizes is negative which in a few cases translates into a negative household saving. This is probably because in some groups larger pack sizes are for branded or higher quality goods and we are failing to adequately capture quality via store or brand effects in equation (4).

### ***3.2.3 How well does a standard price index capture bulk pricing?***

The issues in measurement of a price index are very similar to those we saw in the case of sales. First, there is an issue of whether the statistical agencies are correctly sampling all the relevant prices. In the UK, the specification of food items to be collected as part of the basket of goods and services used to calculate inflation rates typically contains an exact size that must be priced such that it is unusual for different sizes of the same product to be sampled. In this case there is no way to account for a change in the relative price of different sizes. As prices change, the tradeoffs between different sizes, and therefore consumer choices, will change. For example, as prices increase consumers might substitute towards larger sizes. Without sampling different sizes statistical agencies will miss this effect and compute an index that over estimates price increases.

Occasionally, statistical agencies will sample different package size, for example, because firms change the sizes they offer. A common practice for statistical agencies is to chain the price from different sizes using unit values, i.e., price per ounce. This practice is justified if prices are linear in size, which is rarely the case. As pointed out by Triplett (2004), this will generate an under estimate in the price index.

Finally, there is considerable variation across households in the propensity to buy in bulk. This suggests that the biases will vary significantly across households.

### ***3.3 Generic Brands and Product Choice***

Product choice creates a well known bias in standard price indices. Consider, for example, an increase in price. A standard index measures the difference in prices times the initial period quantities. Such an index overstates the change in utility since it ignores the changes in quantities in response to the price increase. For example, Boskin et al (1998) suggest “that this source of



substitution bias ... leads to a combined substitution bias of 0.4 percentage points per year in using the CPI to measure changes in the cost of living.” In principle, the problem can be handled by using a weighted average of the two period quantities (Diewert, 1976). Indeed, this, among other reasons, has led some researchers to claim that this is a second order problem (Hausman, 2003).

To provide a point of comparison to this literature we look in our data at this form of substitution. In order to keep the empirical exercise tractable below we will focus on substitution towards store (‘generic’) brands.<sup>9</sup> As we will show below, the fraction of store brands varies across households and over products. Heterogeneity across income levels presented suggests that there is substitution towards store brands as economic conditions worsen. Gicheva, Hastings and Villas-Boas (2007) provide evidence that consumers buy more store brands when gas prices go up. Caronia (2008) finds that the income shock caused by the Argentinean 2002 Peso devaluation caused a flight from branded products towards store brands.

### ***3.3.1 How much do households buy store brands?***

In the UK store brands are more popular and are often of higher quality than US store brands. The data allow us to distinguish between ‘economy’ and ‘standard’ store brands. Economy brands are most similar to generic brands in the US. Figure 5a shows the share of household expenditure on economy items, ranging from 0 to 100% with an average of 3.8%. Figure 5b shows how ‘standard’ store brands are much more popular, also ranging from 0 to 100% but averaging 41%.

---

<sup>9</sup> Given the scale of the analysis, it is not feasible to study substitution between branded products.

Families with children spend more on economy brands, as do households on lower incomes. However, families with children are *less* likely to buy standard brands suggesting some substitutability between them.

### 3.3.2 How much do households save by buying store brands?

In order to measure the saving we run the following regression:

$$\ln p_{iht} = \beta_j^e \text{econ}_i + \beta_j^s \text{stan}_i + \eta_{jf} + t_{jt} + r_{jr} + e_{iht} \quad (7)$$

where  $f$  indexes fascia and  $\text{econ}_i = 1$  if the product is a economy own brand and  $\text{stan}_i = 1$  if the product is a standard own brand. Letting the coefficients vary by category allows for different quality of the store brands across categories. On average, the economy store brand is almost 39% cheaper and the standard store brand is 25% cheaper.

To compute a household-level savings measure we weight the category-level savings using household specific expenditure weights as:

$$\text{SaveGeneric}_h = \sum_{jt} w_{hj} (1 - \exp(\beta_j)) \quad (8)$$

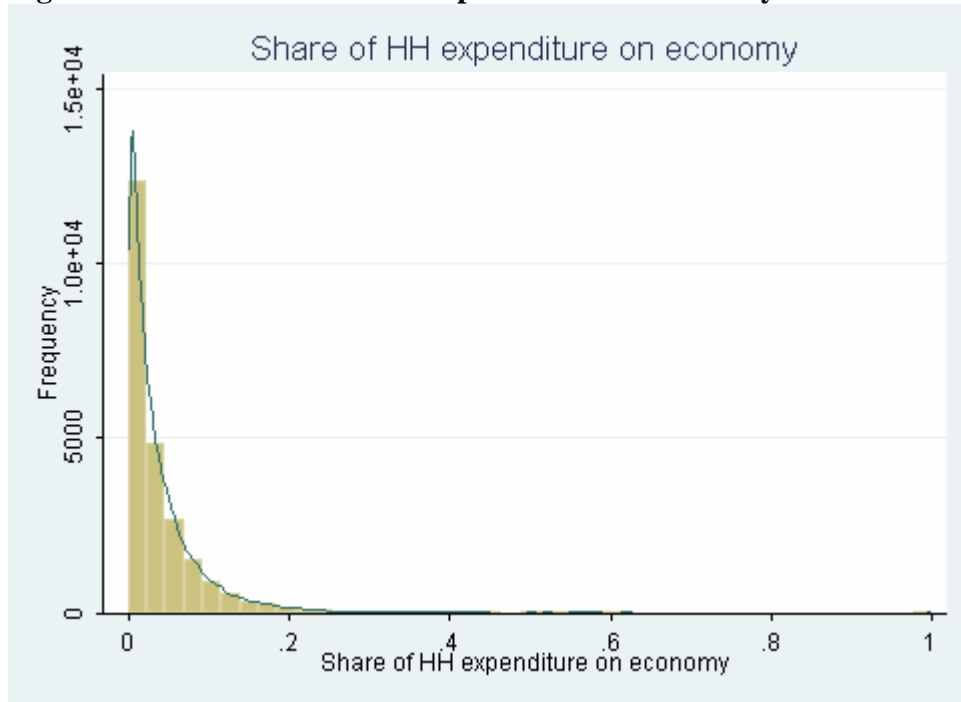
where:

$$w_{hj} = \frac{\sum_{it, i \in j} (\text{generic}_{iht} = 1) \text{expenditure}_{iht}}{\sum_{it} \text{expenditure}_{iht}} \quad (9)$$

Households save on average 2% of their annual expenditure by buying store economy brands, with consumers who buy standard store brands saving on average 3.7%. This translates into an average saving of £25 for economy own brands and £50 for standard own brands on average. The savings from standard are larger, despite the lower discounts, because its share of expenditure is

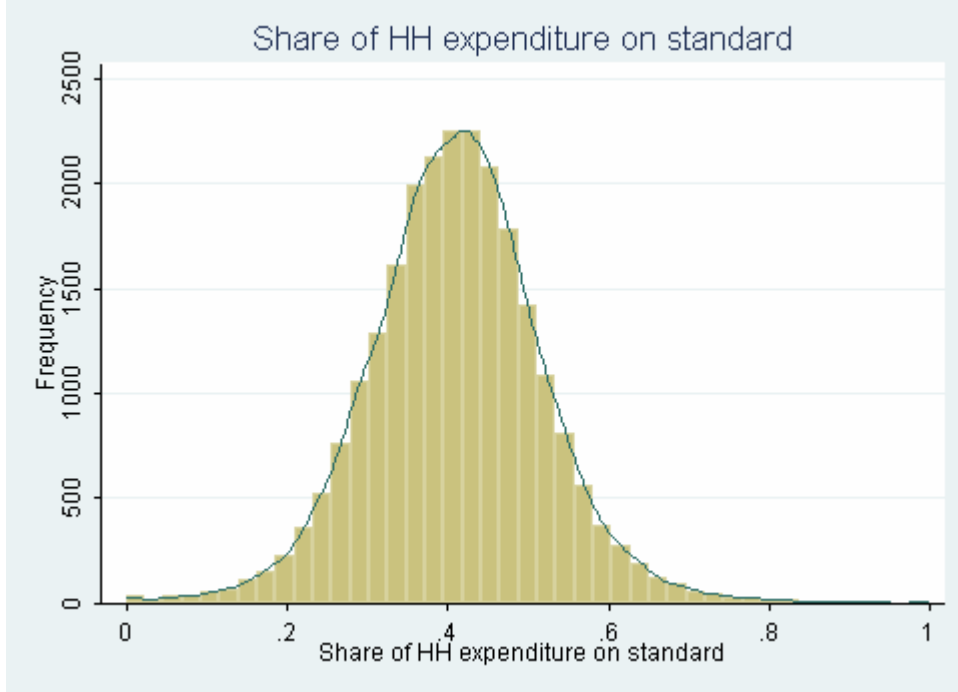
much higher. Again, some households make negative ‘savings’ from buying standard generic brands – this illustrates the point that in the UK store own brands can often be of a comparable quality to branded goods. Savings from economy store brands are always positive, suggesting a more obvious quality differential.

**Figure 5a: Share of household expenditure on “economy” items**



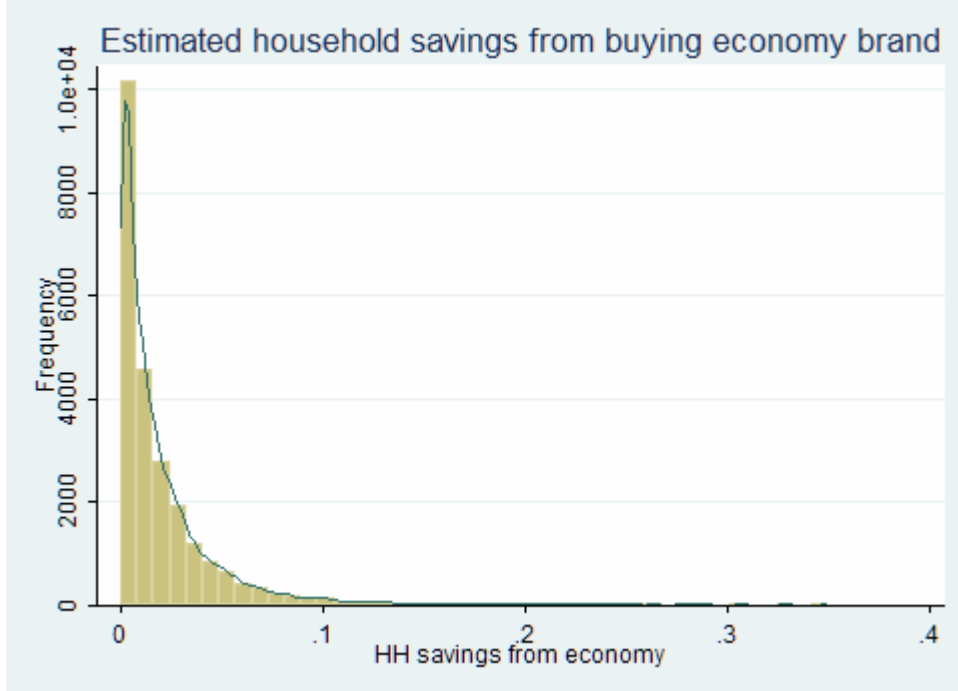
Note: Histogram of the household expenditure shares on economy own brands in 2006. Sample includes 23,877 households.

**Figure 5b: Share of household expenditure on “standard” own label items**



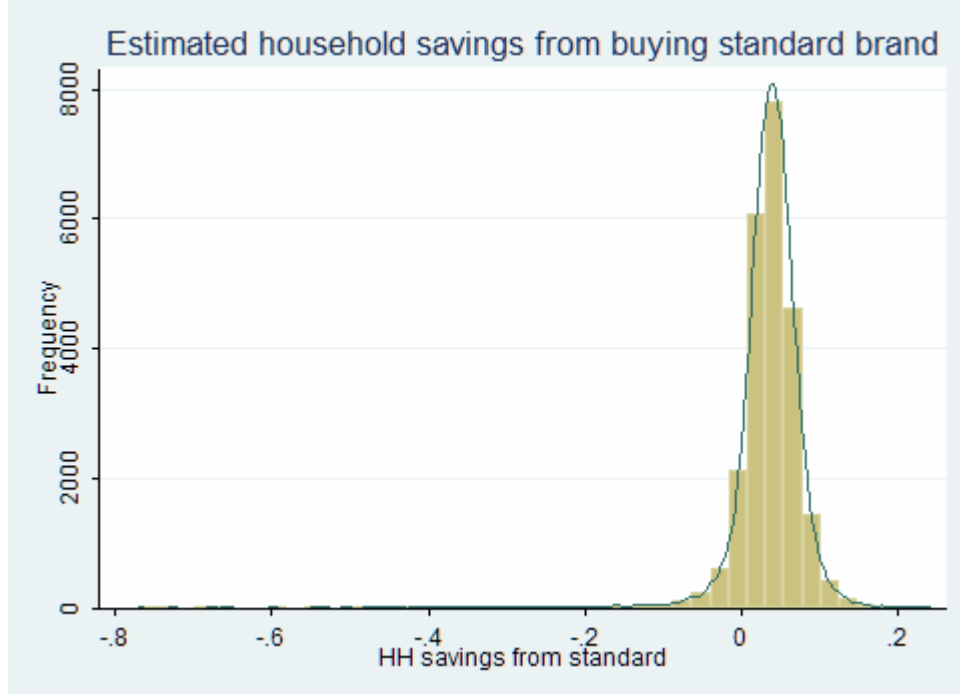
Note: Histogram of the household expenditure shares on standard own brands in 2006. Sample includes 23,877 households.

**Figure 6a: Savings made by each household from buying economy brand**



Note: Histogram of the savings each household made by purchasing economy own brands in 2006 as defined by equation (10); sample includes 23,877 households.

**Figure 6b: Savings made by each household from buying standard brand**



Note: Histogram of the savings each household made by purchasing standard own brands in 2006 as defined by equation (10); sample includes 23,877 households.

### ***3.3.3 What are the implications of substitution towards store brand for price indices?***

If prices of the store brand are collected regularly, as part of the basket then the usual substitution bias applies – the key issue is what quantity weight is used to weight the store brand prices. However, if the store brand prices are not collected regularly, as is the case for many food items, then the shift towards the cheaper products will be mostly missed by standard price indices.

In the UK, for most food items, there is no particular instruction to collect branded or own brand products when sampling prices (one exception is sliced bread where only branded items can be chosen). Product choice is at the discretion of the price collector (the Consumer Price Index Technical Manual (ONS, 2007) suggests that the product selected should be “...representative of what people buy in your area from all products matching the specification of each item to be priced in that outlet.” (page 21)). Without access to the raw price data collected

by ONS it is not clear whether branded products are over- or under-sampled in the UK food price index. It is likely that ‘representative’ items may well be branded more often than not: the guidance given on the sorts of products to collect sometimes specifies particular branded products that match the item description (in breakfast cereals, butter and soft drinks, for example) which might to those products being sampled. Confectionery is the only example where specific *branded* products are explicitly priced (KitKat, Polo Mints, Smarties, Mars Bar, Fruit Pastilles, Crunchie, Dairy Milk) with no opportunity to price own brand alternatives. In some cases, ‘economy’ own brands are explicitly ruled out (for example for canned goods, condiments and ready meals) but these can sometimes make up a substantial share of spending and may be obvious substitutes when economic conditions worsen or relative prices shift.

### **3.4 Outlet Choice**

Probably the largest change over the last decade in food retailing in the US and the UK is the increase of the market share of a single firm: Walmart in the US and Tesco in the UK. Walmart is the largest food retailer in the US, with sales higher than Kroger, Supervalu-Albertsons and Safeway, which are the largest supermarket chains. In the UK Tesco has similarly gained market share rapidly. The Competition Commission (2008) reports the Tesco grocery sales share increasing from 20.2% in 2002 to 27.6% in 2007. Based on slightly different data, TNS (2008) reported a Tesco grocery market share for August 2008 of 31.6%.

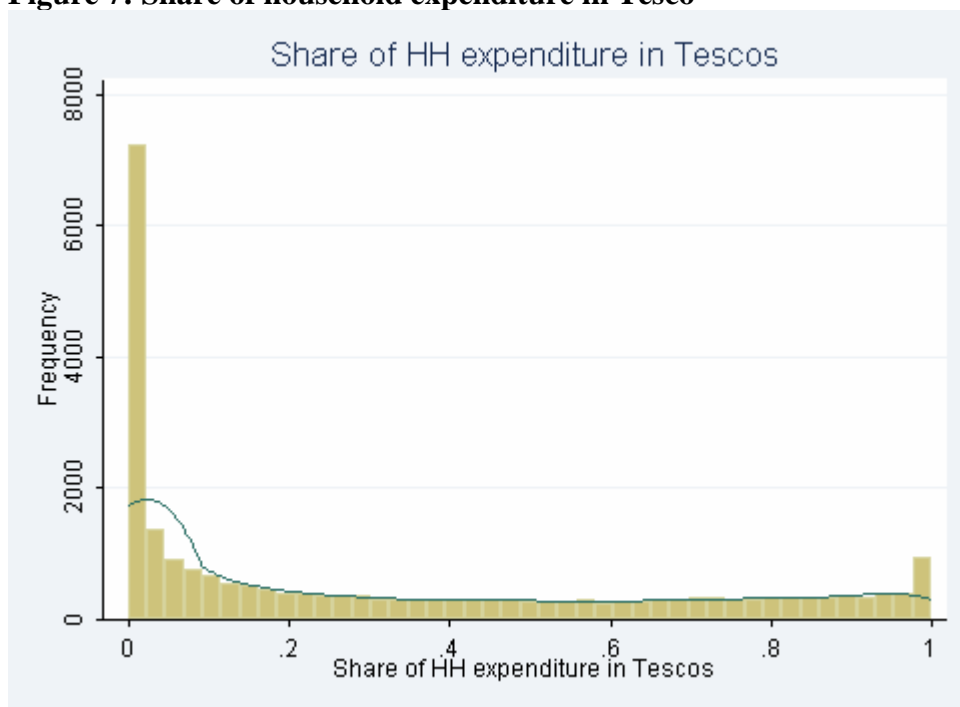
However, as we show below, the fraction of expenditure in Tesco varies considerably across households and over products. The implications of outlet substitution have been well documented (for example, see Boskin et al, 1998, Hausman, 1998, Hausman and Leibtag, 2004). In the US the way that prices are collected does not fully capture price changes in retailers like

Walmart, or they do not properly account for quality differences (treating the price difference as purely a quality difference). The situation is somewhat different in the UK where the ONS does attempt to reflect these changes in market share.

### 3.4.1 How much do households buy in Tesco?

Households vary substantially in the share of their purchases that are made at Tesco. The average household spends 32% of total annual expenditure at Tesco. However, nearly 20% of households spend nothing at Tesco, and 1.7% of households spend all of their budget there.<sup>10</sup> Figure 7 shows the distribution of the share of expenditure at Tesco.

**Figure 7: Share of household expenditure in Tesco**



Note: Histogram of the household expenditure shares at Tesco in 2006. Sample includes 23,877 households.

Couples and families with children buy a larger share of their groceries at Tesco, as do households that shop more often by car. Higher income households are more likely to shop at

---

<sup>10</sup> Just as a comparison, using similar consumer-level data from the US we find that the average household spends 16% of their expenditure at Walmart, while 23% of households do not purchase any food at Walmart.

Tesco, though lower income households are much less likely to. None of these demographics explains much of the household level variation in the share of groceries purchased at Tesco.

### 3.4.2 *How much do consumers save by shopping in Tesco?*

We compare prices paid in Tesco, the UK's largest supermarket chain, using a regression of the form:

$$\ln p_{iht} = \beta_j \text{Tesco}_{iht} + \eta_k + t_t + r_r + e_{iht} \quad (10)$$

where  $\text{Tesco}_{ihts} = 1$  if household  $h$  bought good  $i$  at time  $t$  in a Tesco store. Relative to the other dimensions, examined above, the potential savings are much more modest. The average discount in Tesco is 1.6% and the median is 1.0%. Furthermore, there seems to be much less heterogeneity in the savings across product categories, particularly compared to savings from generic brands and bulk discounting.

To compute a household-level savings measure we weight the category-level savings using household specific expenditure weights as:

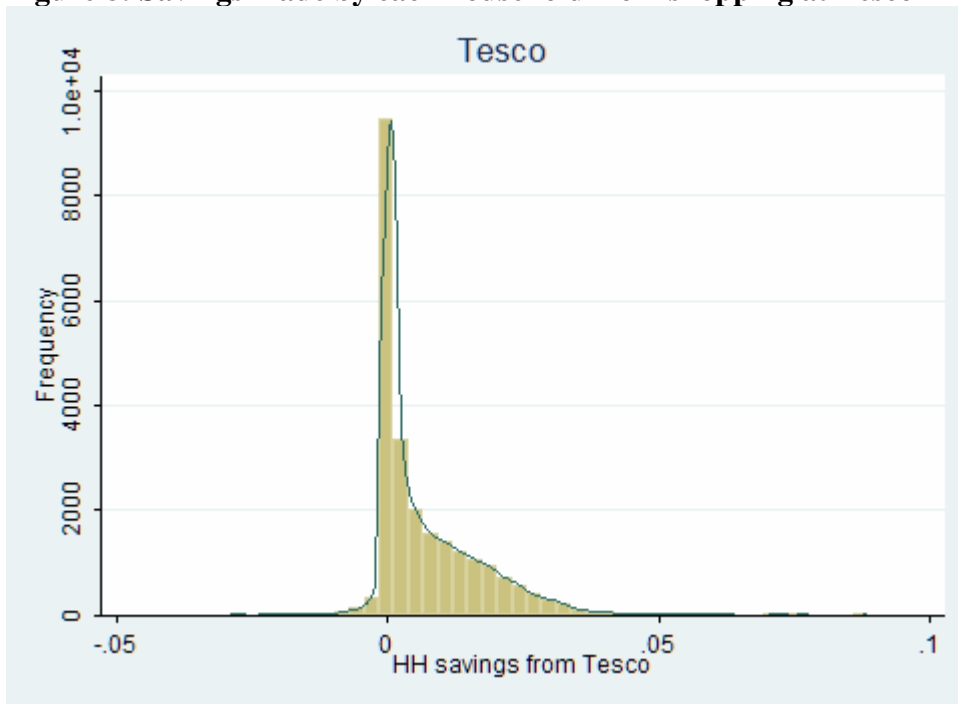
$$\text{SaveTesco}_h = \sum_{jt} w_{hj} (1 - \exp(\beta_j)) \quad (11)$$

where:

$$w_{hj} = \frac{\sum_{it, i \in j} (\text{Tesco}_{iht} = 1) \text{expenditure}_{iht}}{\sum_{it} \text{expenditure}_{iht}} \quad (12)$$



**Figure 8: Savings made by each household from shopping at Tesco**



Note: Histogram of the savings each household made by purchasing standard own brands in 2006 as defined by equation (13); sample includes 23,877 households.

Clearly the savings made from Tesco shopping are smaller than those from some of the other channels discussed so far. On average, the saving is just 0.7% (median 0.2%) and even those households that save the most through Tesco shopping save less than 10% by doing so. In cash terms, the mean saving is around £10 per year (median £). Around 7.5% of households make negative ‘savings’ through Tesco shopping – that is, the items they buy from Tesco are more expensive than buying from the same product group elsewhere.

These results show an important distinction between the UK and US experiences of a single retailer coming to have a substantial market share. In the US, Wal-Mart’s growth was largely attributable to having lower prices. In the UK, Tesco is not on average a low-price store: across the 189 product groups in our analysis, the average saving in Tesco 1.6%, but is positive (Tesco prices are higher) in 79 categories.

Pensioners and childless households save the least from Tesco shopping, whilst families with children save the most. Larger savings are also made by those who shop by car, in particular those who shop once or twice a week by car. Richer households also make larger savings. Once again, however, these demographics are unable to explain much of the variation in the savings made through Tesco purchases.

## 4 Comparing the savings

Up to this point we discussed each of the savings measures separately. We now compare the measures to each other in order to gain an understanding of the relative importance of the different dimensions.

### 4.1 Comparison of potential savings

The various regressions we described above tell us about the potential savings that are available to consumers from each of the forms of substitution considered. The coefficients measure the discount from sales, large sizes, generic brands and outlet choice, controlling for quality differences across products. Table 6 shows the distribution of these savings implied by the beta coefficients (an observation is one of the food categories, i.e. these are not weighted by quantity).

**Table 6: Potential savings through different channels**

Savings channel	Mean	Std. dev.	10 <sup>th</sup> percentile	Median	90 <sup>th</sup> percentile
Sale	21.7%	6.0%	14.1%	22.4%	38.8%
4 <sup>th</sup> size quintile	28.1%	27.7%	-2.2%	27.1%	65.2%
5 <sup>th</sup> size quintile	36.8%	33.5%	8.0%	35.5%	72.7%
Economy generic brand	38.8%	30.7%	0.0%	46.0%	75.1%
Standard generic brand	25.4%	22.6%	-2.0%	25.4%	54.3%
Tesco purchase	1.6%	9.9%	-8.5%	1.0%	12.7%

Notes: savings based on the coefficients obtained from equations (1), (4), (7), (10); for each channel there are 189 regressions, one for each food category. Savings in the size quintiles are relative to purchases in the second largest size quintile.

Source: Authors' calculations from 2006 TNS sample

The potential savings are highest from economy generic brands, followed by bulk purchases, standard generic brands, economy brand, sales and then Tesco. Except for the savings from Tesco, these savings seem to be of comparable magnitude.

These numbers show the potential savings from each dimension. For example, if a household bought all of their purchases on sale we would measure them as saving roughly 22%, assuming equal expenditure in each category. This is significantly higher than the actual savings made which depend on the expenditure shares devoted to each savings channel.

#### 4.2 Comparison of shares

Given these potential savings, we ask how much do households use these different ways to save. Table 7 summarizes the distribution of household expenditure on different savings channel. Note that the distribution is across households.

**Table 7: Household expenditure shares on various savings channels, 2006 TNS data**

Savings channel	Mean	Std. dev.	10 <sup>th</sup> percentile	Median	90 <sup>th</sup> percentile
Sale	29.5%	9.8%	17.7%	28.9%	42.0%
Largest two size quintiles	37.0%	10.3%	24.1%	36.8%	50.0%
Economy generic brand	3.8%	4.9%	0.1%	2.2%	9.4%
Standard generic brand	41.1%	10.7%	28.0%	41.1%	54.0%
Tesco purchase	31.6%	34.1%	0.0%	16.3%	88.4%

Source: Authors' calculations from 2006 TNS sample

The relative distribution of expenditure shares does not track the relative potential saving. For example, economy generic brands offered the greatest potential for savings yet expenditure shares on these brands are relatively small. This should not be surprising: economy brands are probably of lower quality. Similarly, Tesco offers small potential savings, yet is very popular, reflecting the fact that consumers shop at Tesco for reasons other than savings.

Table 8 displays the correlation of expenditure shares of the different dimensions. The correlations are intuitive. Consumers who buy more on sale also purchase in bulk, which is reasonable since both these choices are driven by storage costs. On the other hand, households that buy on sale tend to spend less on generic.<sup>11</sup> Similarly, households that buy economy generic tend to purchase in bulk.

**Table 8: Correlation between expenditure shares on various savings channels**

	Sale	Largest two quintiles	Economy generic	Standard generic	Tesco
Sale	1.000				
Largest two size quintiles	0.226	1.000			
Economy generic brand	0.033	0.255	1.000		
Standard generic brand	-0.119	-0.092	-0.051	1.000	
Tesco purchase	0.058	0.019	0.129	-0.084	1.000

Source: Authors' calculations from 2006 TNS sample

### 4.3 Comparison of actual savings

In order to measure the actual, as apposed to potential, savings we present our computed measures in Table 9. These measures combine the potential savings, presented in Section 4.1, with the choices consumers make, presented in the previous section.

**Table 9: Household savings from various channels**

Savings channel	Mean	Std. dev.	10 <sup>th</sup> percentile	Median	90 <sup>th</sup> percentile
Sale	6.5%	2.3%	3.8%	6.4%	9.5%
Largest two size quintiles	15.6%	4.5%	10.1%	15.6%	21.2%
Economy generic brand	2.0%	2.6%	0.1%	1.1%	4.9%
Standard generic brand	3.7%	4.3%	0.1%	3.9%	7.6%
Tesco purchase	0.7%	0.9%	0.0%	0.2%	2.0%

Source: Authors' calculations from 2006 TNS sample

<sup>11</sup> Leibtag and Kaufman (2003) report a similar finding for the US.

The results suggest the savings from bulk purchasing are the largest, followed by sales, purchases of generics and shopping at Tesco. These results follow directly from what we saw in the previous sections. Bulk purchases offered the largest potential savings and made up a significant share of total expenditures. It is not totally surprising that the actual savings are therefore largest from buying in bulk. The figures in the table also demonstrate the importance of computing the savings measures and not just settling for the numbers in the previous sections. The potential savings from standard generic brands was relatively high, as was the share of expenditure on these brands. However, the actual savings are relatively low. This is driven by a negative correlation between potential savings and expenditure within categories: the categories where the expenditure share is higher are those where the potential savings are relatively low.

A couple of caveats are in place in interpreting these numbers. First, these measures do not account for the costs of savings. In the case of sales and bulk purchases costs involve storage and transport cost. In the case of generic brands, the cost could include quality differences. And for outlet choice they include travel costs as well as potential quality differences between stores. These costs are likely different for the different dimensions and therefore in a way make the savings non-comparable.

An additional caveat is that the savings measures are not orthogonal. For example, if larger sizes tend to be more on sale, the savings from them are potentially counted twice, both in bulk and in the sales measures (although note that in equation (4) we have controlled for sale when measuring the savings from bulk purchasing). In order to explore this further we examine in Table 10 the correlation between the various savings measures.

**Table 10: Correlation between household savings from various channels**

	Sale	Largest two quintiles	Economy generic	Standard generic	Tesco
Sale	1.000				
Largest two size quintiles	0.275	1.000			
Economy generic brand	0.058	0.257	1.000		
Standard generic brand	0.064	0.230	0.070	1.000	
Tesco purchase	0.146	0.120	0.107	0.060	1.000

Source: Authors' calculations from 2006 TNS sample

As before, consumers who save by buying on sale also save by purchasing in bulk, with both decisions driven by storage costs. Households that save by purchasing on sale also save more when purchasing in Tesco. This is perhaps what we would expect: Tesco is not an 'every day low price' retailer but instead uses frequent in store price and quantity promotions more heavily. Consumers that save with the standard generic brand also save bulk discounts, but seem to save less by timing their purchases (i.e., buying on sale) and by switching stores. The correlation between savings from economy and standard generics is also quite low, again suggesting they are close substitutes for each other.

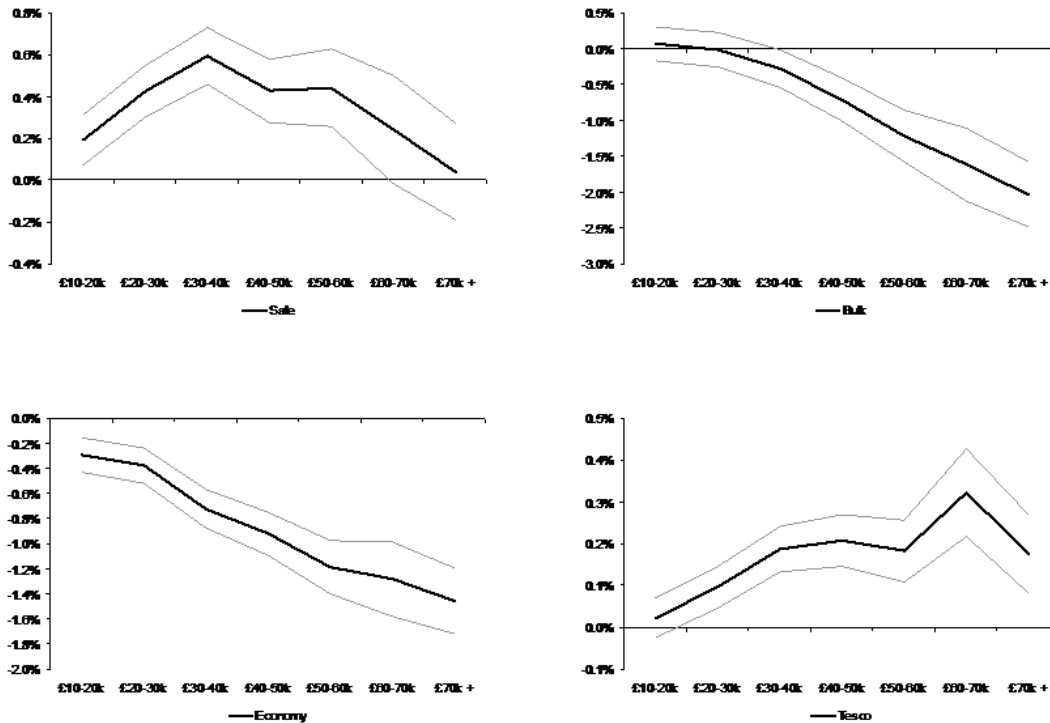
In order to further explore the correlation across savings measures we examine how savings vary with demographics, focusing on income and family composition. The figures below plot the coefficients and confidence intervals from a regression of the household saving from each channel on a categorical dummy for the income (figure 9) and family type (figure 10).

The correlation for savings from bulk purchases and economy brands follow similar patterns: they decrease monotonically with income. Relatively, these dimensions seem to be used more by low income households. On the other hand, saving from sales, and to a lesser extent saving from Tesco, seem to be non-monotonic in income. Low income households do not seem

to save, maybe because they are unable, and neither do high income households, maybe because it is not worth their time.

**Figure 9: Savings by income group (relative to households with incomes under £10,000)**

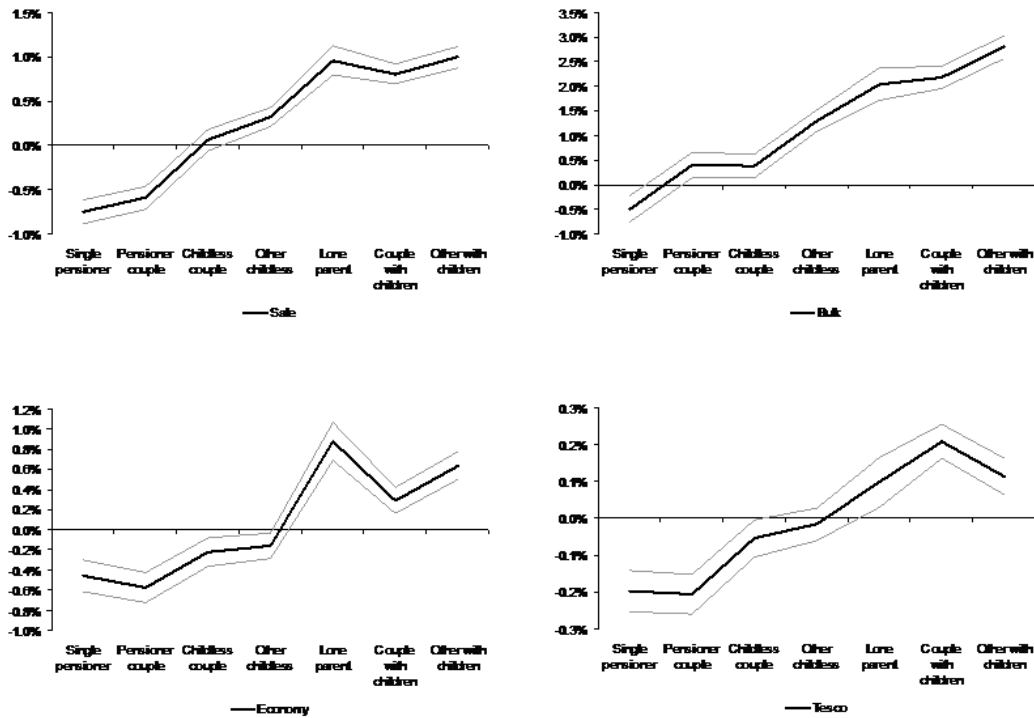
Savings by income (relative to £0-10k)



Source: Calculated from 2006 TNS sample

The results for family types suggest that families with children save more through each of the savings channels than those without. The relative gains are highest for bulk discounting and sales. This makes sense: larger families have larger consumption needs and so they are able to take advantage of sales and larger packets, for example, even if their storage costs are high. Retired households tend to be less likely to save in all dimensions, except bulk purchases, relative to other household groups.

**Figure 10: Savings by family type (relative to single childless adult)**  
 savings by family type (relative to single adult)



Source: Calculated from 2006 TNS sample

## 5 Concluding Comments

We use a rich household-level data set to document purchasing patterns. We focus on four dimensions of consumer choice behaviour: purchase on sales, purchase in bulk, purchase of generic products and store choice. For each dimension we compute an estimate of how much the consumers save. Our findings suggest that all dimensions of choice yield significant and comparable savings.



The data we used in this paper comes from the UK. An interesting question is whether the results we find are also present in other countries as well. We also have similar data from the US. Preliminary analysis suggests similar findings in the US, with some interesting differences. Sales and bulk purchases, and the savings they entail, seem to be more significant in the US. This should not be surprising. On average, US homes are larger and consumers tend to shop more with cars, so transport and storage costs are lower. As a result, in the US larger sizes are offered and purchased (for example, gallon size ice cream pack, which are quite popular in the US, are basically not available in the UK).

Store brands are also different in the US. They are a larger share of expenditure than the economy brands in the UK, but significantly less than the combined economy and standard brands in the UK. Overall, it seems to us store brands are less important in the US. As we discussed above, the savings from Walmart seem to be more significant than the savings from shopping at Tesco. In summary, it seems the conclusion – that savings from sales and bulk purchases are important – holds even stronger in the US.

The statistics we offer measure the monetary savings from behaviour along various dimensions. The measures ignore the cost of substitution. When buying on sale or in bulk households incur a transportation cost and a storage cost, which we ignore. We also ignore the cost of travelling to various stores and the potentially lower quality of store brands. To fully account for these costs one needs to estimate a demand system. Based on our results we hope future work will go in this direction. Indeed, our measures suggest that the biases due to sales and bulk purchasing might be of the same (or even a greater) order of magnitude as substitution and outlet bias. Yet they have received much less attention. We second Triplett (2003) in suggesting that this is a fruitful area for future research.

## References

- Abraham, K.G. (2003), "Toward a Cost-of-Living Index: Progress and Prospects", *The Journal of Economic Perspectives*, 17(1), 45-58.
- Baye, M. (1985), "Price Dispersion and Functional Price Indices", *Econometrica* 53(1), 213-23.
- Boizot, C., J-M Robin and M. Visser (2001) "The Demand for Food Products: An Analysis of Interpurchase Times and Purchased Quantities", *The Economic Journal*, 111, 391-419
- Boskin, M. et al (1998), "Consumer Prices, the Consumer Price Index, and the Cost of Living", *The Journal of Economic Perspectives*, 12(1), 3-26.
- Caronia, M. (2008), "The Microeconomics of Large Devaluations", *Northwestern University*, mimeo.
- Competition Commission (2008), "Market Investigation Into the Supply of Groceries in the UK", Appendix 3.1,  
[http://www.competition-commission.org.uk/rep\\_pub/reports/2008/fulltext/538\\_3\\_1.pdf](http://www.competition-commission.org.uk/rep_pub/reports/2008/fulltext/538_3_1.pdf)
- Crawford, I. (1996), "UK Household cost-of-living indices, 1979-92", in John Hills (ed.) *New Inequalities: The Changing Distribution of Income and Wealth in the United Kingdom*. Cambridge: Cambridge University Cambridge.
- Crawford, I. and Z. Smith (2002), "Distributional Aspects of Inflation", IFS Commentary 90, London: IFS
- Deaton, A. (1998), "Getting Price Right: What Should Be Done?" *The Journal of Economic Perspectives*, 12(1), 37-46.
- Deaton, A. and J. Muellbauer (1980), *Economics and Consumer Behavior*, Cambridge University, Cambridge.
- Department for Transport (2005), "Focus on Personal Travel 2005 Edition", London: The Stationery Office  
(<http://www.dft.gov.uk/pgr/statistics/datatablespublications/personal/focuspt/2005/focusonpersonaltravel2005edi5238>)
- Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4:2 , 115-145.
- Diewert, W.E. (1993), "The Early History of Price Index Research", in *Essays in Index Number Theory*, Volume I, Contributions to Economic Analysis 217, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North Holland, 33-65.
- Feenstra, R.C. and M.D. Shapiro (2003), "High-Frequency Substitution and the Measurement of Price Indexes", in *Scanner Data and Price Indexes*.
- Gicheva, D., J. Hastings and S. Villas-Boas (2007), "Revisiting the Income Effect: Gasoline Prices and Grocery Purchases", *NBER Working Paper* 13614.
- Hausman, J. (2003), "Sources of Bias and Solutions to Bias in the Consumer Price Index", *The Journal of Economic Perspectives*, 17(1), 23-44.

- Hausman, J. and E. Leibtag (2004), "CPI bias from Supercenters: Does the BLS Know that Wal-Mart Exists?", *NBER Working Paper* 10712.
- Hausman, J. and E. Leibtag (2007), "Consumer Benefits From Increases Competition in Shopping Outlets: Measuring the Effect of Wal-Mart", *Journal of Applied Econometrics*, 22, 1157-1177.
- Hendel, I. and A. Nevo (2006a) "Sales and Consumer Inventory", *The RAND Journal of Economics*, 37(3), 543-561.
- Hendel, I. and A. Nevo (2006b) "Measuring the Implications of Sales and Consumer Inventory Behavior", *Econometrica*, 74(6), 1637-1673.
- Leibtag, E. and P. Kaufman (2003), "Exploring Food Purchase Behavior of Low-Income Households: How Do they Economize?", *Agriculture Information Bulletin No 747-07*, <http://www.ers.usda.gov/Publications/AIB747/aib74707.pdf>
- Office for National Statistics (2007), "Consumer Prices Index Technical Manual", [http://www.statistics.gov.uk/downloads/theme\\_economy/CPI\\_Technical\\_Manual.pdf](http://www.statistics.gov.uk/downloads/theme_economy/CPI_Technical_Manual.pdf)
- Pollak, R.A. (1998), "The Consumer Price Index: A Research Agenda and Three Proposals", *The Journal of Economic Perspectives*, 12(1), 69-78.
- Reinsdorf, M. (1994), "The Effect of Price Dispersion on Cost of Living Indexes", *International Economic Review*, 35(1).
- TNS (2008), "Price Now Clearly Driving Market Shares", Press Release 19<sup>th</sup> August 2008, [http://www.tnsglobal.com/assets/files/TNS\\_Market\\_Research\\_Supermarket\\_Share\\_data\\_mmentary.pdf](http://www.tnsglobal.com/assets/files/TNS_Market_Research_Supermarket_Share_data_mmentary.pdf)
- Triplett, J. (2003), "Using Scanner Data in Consumer Price Indexes: Some Neglected Conceptual Considerations", in *Scanner Data and Price Indexes*.
- Triplett, J. (2004), "Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products", OECD Science, Technology and Industry Working Papers, 2004/9, OECD Publishing.