Journal of Economic Perspectives—Volume 23, Number 2—Spring 2009—Pages 99–120

Consumer Shopping Behavior: How Much Do Consumers Save?

Rachel Griffith, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo

There is substantial variation in prices across brands, stores, sizes, and over time, even for a narrowly defined product. For example, many products are sold at nonlinear prices: the price per unit, say an ounce, is typically lower for larger pack sizes. Similarly, many products have temporary price reductions—sales—that potentially allow the consumer to purchase more today at a low price and to stockpile for future consumption. This variation implies that, when deciding on purchases and consumption, a consumer has several choices to make. What to buy and where to buy it are two commonly studied choices; additional choices include how much and when to buy. The choices depend on preferences and costs. Some consumers have low travel costs and therefore will be more likely to take advantage of spatial price differences, while other consumers have lower storage and transport costs and will therefore take greater advantage of quantity discounts and temporary price reductions. Our main goal in this paper is to document the potential and actual savings that consumers realize from various dimensions of choice and how these vary with consumer demographics. Our focus is on four particular types of purchasing behavior: purchasing on sale; buying in

■ *Rachel Griffith is Professor of Economics, University College London, and Deputy Research Director, Institute for Fiscal Studies, both in London, United Kingdom. Ephraim Leibtag is an Economist with the Economic Research Service, U.S. Department of Agriculture, Washington, D.C. Andrew Leicester is a Research Student, University College London, and Senior Research Economist, Institute for Fiscal Studies, both in London, United Kingdom. Aviv Nevo is Professor of Economics, Northwestern University, Evanston, Illinois, and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. Their e-mail addresses are ⟨rgriffith@ifs.org.uk⟩, ⟨eleibtag@ers.usda.gov⟩, ⟨andrew_l@ifs.org.uk⟩, and ⟨nevo@northwestern.edu⟩, respectively.*

bulk (at a lower per unit price); buying generic brands; and choosing outlets. How much can and do households save through each of these behaviors?

Some of these dimensions of choice have been considered in earlier work—for example, Hendel and Nevo (2006a) on savings from the timing of purchases and Hausman and Leibtag (2007) on savings from the availability of Walmart stores—but the relative importance of size and brand has not been compared. Our analysis suggests that the average consumer realizes significant savings from the four dimensions of choice that we study, and that the savings are comparable in magnitude.

We use data collected by a marketing firm on all food purchases brought into the home for a large, nationally representative sample of U.K. 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 the date of purchase. For each purchase we know exactly what was bought (as measured by the barcode), the price paid, quantity purchased, the date, and the store of purchase. We also observe household demographics. Using these data, we are able to measure how much less consumers pay when they purchase on sale, buy larger packs, go to another retailer, or buy a generic brand. Combined with the purchasing pattern, we can compute a household-level savings measure from each choice dimension. We then show how the savings vary with income, household composition, size, age, and employment status.

Documenting these patterns has implications for many areas in economics. Our primary interest here is in the effect consumer choice has on the measurement of price changes. In practice, the most common approach to measuring price changes is to look at the change in expenditure needed to purchase a fixed basket of goods. This approach, with variations, dates back at least to the early nineteenth century (Diewert, 1993) and generates a price index with well-known biases that overestimate the rise in the true cost of living faced by consumers. For example, a fixed basket of goods does not take into account possibilities for substitution from more expensive to less expensive goods within the same product category, improvements in product quality, the introduction of new goods, or a move to purchasing at lower-priced retail outlets. These biases and their implications for national statistical agencies have been well-documented, including in this journal by Boskin, Dulberger, Gordon, Griliches, and Jorgenson (1998), Deaton (1998), Hausman (2003), and Schultze (2003).

We study two additional choices—to buy in bulk and to buy on sale—made by consumers and ask how they compare to other forms of substitution that have been emphasized in the literature. These dimensions of choice have been mentioned as potential sources of bias in price indices (for example, Feenstra and Shapiro, 2003; Triplett, 2003, and references therein), but are discussed less often than the other biases. While we do not offer an estimate of the bias generated by this form of substitution, our results do suggest that the savings from these dimensions of choice are comparable to those obtained from brand and outlet choice, and therefore the bias might be comparable.

Households differ in their abilities to select where in the distribution of prices they purchase; therefore, a focus on the "average" consumer in describing price changes will ignore interesting heterogeneity across consumers (Pollak, 1998). This may have important consequences, for instance, in the measurement of relative living standards (poverty and inequality) in comparing real purchasing power across time and consumers; and in deciding whether mandatory annual increases in state benefits that are pegged to national inflation rates are adequate for their recipients. These distributional issues are probably less important if the relative price of different goods remains fairly constant. In periods of high inflation, and in particular when inflation is driven by a subset of commodities, heterogeneity across households is likely to be more significant. For example, Deaton and Muellbauer (1980) report that during 1975–76, when inflation in the United Kingdom was 15 percent, the inflation rate for the poor was two percentage points higher than for the rich. More recently, Crawford (1996), Crawford and Smith (2002), and Leicester, O'Dea, and Oldfield (2008) report similar results.

Food Purchases and Household Characteristics

The data we use come from the TNS Worldpanel (described at ⟨http://www.tnsglobal.com/market-research/fmcg-research/consumer-panel⟩), a representative consumer panel of around 25,000 households in Great Britain. Participating households are issued an electronic handheld scanner in their homes and asked to scan the barcodes of all grocery purchases—food, alcohol, bathroom products, medicines, pet food, and so on—that come into the house. Ongoing participation is rewarded with points redeemable for a range of products and services (though limited to items that should not directly affect grocery consumption patterns).

Information on purchases is downloaded once a week by TNS. In addition, households mail cash register receipts to TNS, which matches the exact price paid to each purchase and acts as a check on the data as entered by the households. Information on loose nonbarcoded items such as vegetables and fruit is collected by households scanning barcodes in a book and keying in the weight data.[1] Purchases from all store types—supermarkets, corner stores, online, local speciality shops and so on—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 physical

[1] Between 2005 and 2006, the TNS sample size was increased from roughly 15,000 to 25,000 households; most of the newly recruited households were issued a new type of scanning technology that TNS believes makes the recording process easier and means that households are more likely to record all of their purchases. However, these households are not required to scan nonbarcoded items.

product characteristics such as flavor. Demographic information about the household is collected in an annually updated telephone survey.

Our analysis uses data for calendar year 2006. We observe expenditure for 23,877 households on purchases in 189 categories, effectively covering all food and beverage purchases. These households made 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 four days (with a median of three days). Table 1 provides demographic characteristics on household income, household composition, how often the household shops by car, and what the most common mode of transport is for shopping. This data is similar to that found in other national surveys. For example, the distribution of income and the household types in these data are similar to those found in other U.K. surveys, like the UK Expenditure and Food Survey (EFS), although overall average income is slightly lower and the market research data seem to contain somewhat fewer households headed by a pensioner and fewer single adult households.[2] The figures on shopping transport mode match very closely to Department for Transport (2005) figures.

Sales and Stockpiling

If we focus on a narrowly defined product—such as a particular brand and size sold at a particular store—much of the price variation over time is due to temporary price reductions. In many cases, consumers respond to this price pattern by stockpiling for future consumption (Hendel and Nevo, 2006a; Boizot, Robin, and Visser, 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 expected future prices. Different consumers will make different choices, and a given consumer will make different choices for different goods.

How Much Do Households Buy on Sale?

The TNS data record detailed information on sales obtained 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 (half price off or £1 off, say) or quantity promotions (buy one, get one free; or double volume). On average, around 29.5 percent of total annual average food expenditures are on sale items. There is considerable variation across households, with the household at the

[2] The TNS data includes demographic weights that correct for potential biases in recruiting and retaining some household types and some deliberate oversampling of others such as multiple adult households with many shoppers. 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.

Table 1
Household Income, Type, and Transportation Choices

Household income			*Household type*		
	Observations	*Share*		*Observations*	*Share*
£0–£9,999	2,052	8.6%	Single pensioner	1,940	8.1%
£10,000–£19,999	4,344	18.2%	Pensioner couple	2,246	9.4%
£20,000–£29,999	3,545	14.9%	Single adult	2,209	9.3%
£30,000–£39,999	2,309	9.7%	Couple, no children	2,835	11.9%
£40,000–£49,999	1,434	6.0%	Other, no children	5,778	24.2%
£50,000–£59,999	787	3.3%	Lone parent	1,008	4.2%
£60,000–£69,999	340	1.4%	Couple with children	4,516	18.9%
£70,000 +	448	1.9%	Other, with children	3,345	14.0%
Missing/unknown	8,618	36.1%			
Frequency of shopping by car			*Method of transport for shopping*		
	Observations	*Share*		*Observations*	*Share*
5+ times/week	495	2.1%	Car or taxi	19,056	79.8%
3–5 times/week	3,309	13.9%	Public transport	1,307	5.5%
1–2 times/week	14,495	60.7%	Foot	2,964	12.4%
> Once/month	1,897	7.9%	Other	550	2.3%
< Once/month	1,541	6.5%			
Never	2,140	9.0%			

Source: Authors' calculations from 2006 sample of the TNS Worldpanel. The TNS Worldpanel (described at ⟨http://www.tnsglobal.com/market-research/fmcg-research/consumer-panel⟩) is a representative consumer panel of around 25,000 households resident in Great Britain.

10th percentile purchasing 17.7 percent of its food on sale and at the 90th percentile, 42.0 percent.

Some of the variation across households is explained by observed demographics. Retired households tend to buy less on sale: single pensioners spend 3.2 percentage points less of their food expenditure on sale than single young households, while pensioner couples spend 2.9 percentage points less of total food spending on sale than young childless couples, and over 5 percent 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 percentage points more of their food on sale than households that shop by public transport or on foot, or that shop less frequently by car. Overall, lower-income households buy the least on sale and middle-income households buy the most. A plausible explanation of this pattern is that 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 percent of the variation in the propensity to purchase on sale.

How Much Do Households Save by Buying On Sale?

We want to compute a single saving figure for each household. Our data allows us to identify whether or not a purchase was on sale, but not the value of the saving, and so we begin by estimating this saving. For each of our 189 food product categories, we use a regression model whose key explanatory variable is a dummy variable for whether the item was on sale. The model is:

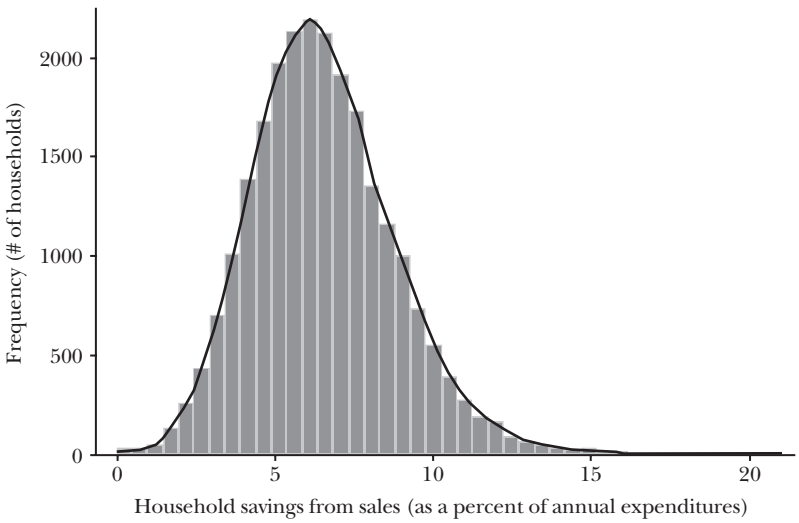
$$\ln p_{iht} = \beta_j s_{it} + \eta_{ij} + \tau_j + \rho_j + e_{iht}$$

where i indexes a detailed product (as defined by a unique barcode); h indexes households; t is time in weeks; j indexes food categories; p is unit price; s is a sales dummy variable equal to 1 if the product is purchased on sale; η_i captures barcode specific characteristics (allowing us to control for observed and unobserved differences in product characteristics); τ and ρ are dummies for time and region, reflecting that prices vary across time and space; and e is an idiosyncratic error. All the coefficients are allowed to vary by product category j.

This procedure yields 189 β coefficients which are estimates of the average percentage price discount obtained by purchasing items in that category on sale. Each of these coefficients is negative (sale prices are lower), all but three are statistically significant at the 5 percent level, and all but five at the 1 percent level. The discount when buying on sale varies substantially across food categories from 14 percent off at the 10th percentile to 29 percent at the 90th percentile, with a mean and median of 22 percent. Some categories with especially low discounts are fruit fillings (2 percent), lard (4 percent), and sugar (10 percent), while examples of the categories with high discounts are savoury snacks (32 percent), breakfast cereal (31 percent), and baked beans (31 percent).

To compute the savings for each household in each category, we multiply the potential savings in the category with the fraction of spending on sale by the household in the category. This calculation measures how much higher (in percentage terms) total expenditure would have been if purchases on sale had not been at their promoted prices but were instead at the average nonsale price. We then compute for every household a weighted average of the 189 category-level savings, where the weights vary by household and are equal to the share of each product category in each household's total budget. The result is a measure that captures the fraction of total expenditure the household has saved by purchasing on sale. Households that spend a large fraction of their budget on sale items will tend to have higher values, but even households that spend relatively little on sale items as a share of their budget can still make substantial savings if those purchases are concentrated in product categories where the value of sales is high. Figure 1 shows the distribution of savings made by each household by purchasing on sales. Households save between 0 and 21 percent of their annual expenditures, with a

Figure 1
Savings Made by Households from Buying on Sale



Source: Authors' calculations from 2006 TNS sample.
Notes: The histogram shows the estimated savings each household made by purchasing on sale in 2006. The sample includes 23,877 households.

mean of 6.5 percent. This translates into a saving of up to £794 a year, with an average 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, poorer households and households in which the head of household is retired save substantially less by purchasing on sale than other household types. A higher propensity to buy on sale does not necessarily mean higher savings; for example, higher-income households purchase more on sale than lower-income households but do not seem to realize higher savings as a result, suggesting they buy on sale in product categories in which the savings from sales are relatively low. However, as was the case with the fraction of purchases on sale, the observed demographics explain relatively little of the variation in the savings measure.

Bulk Discounting and Choice of Package Size

Many grocery items are sold at nonlinear 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, nonsale, price of a 24-pack of soft drinks cans is 2.7 times more than a six-pack, which implies a discount of over 30 percent in the unit price.

A consumer deciding between purchasing a smaller unit and a larger one at a lower per unit price must weigh the benefits of the lower price with the costs of

storing the product longer and any 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.

How Much Do Households Buy in Bulk?

Package size is reported directly in our data. 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. The average household spends 15.8 percent of its total annual expenditure on products with the largest package sizes and 21.2, 21.3, 26.8, and 14.9 percent on the other sizes from largest to smallest quintiles, 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 two quintiles to purchases made in the second-largest size group (which is the most commonly purchased). Households purchase along the entire range of between 0 to nearly 100 percent of their groceries in bulk sizes, spending on average 37 percent of their budget in this size group. At the 10^{th} percentile, the figure is 24 percent, and at the 90^{th} percentile of households, it is 50 percent. Unsurprisingly, single-person households purchase less in bulk than multi-person households: single nonpensioners spend around 34 percent of their budget on the largest pack sizes, compared to 40 percent for couples with children. Single pensioners make even less use of bulk discounts, spending on average 31 percent of their budget on the largest sizes. Households that shop by car spend a slightly higher proportion on bulk items. The relationship with income is non-monotonic—the poorest households with incomes under £10,000 per year spend 36.3 percent on bulk items, those with incomes between £20,000 and £30,000 spend 37.6 percent, and those with incomes above £70,000 spend 33.2 percent. Overall, the behavior of purchasing larger package sizes is similar to purchasing on sale; in fact the two shares are positively correlated at the household level, with a correlation coefficient of 0.23.

How Much Do Consumers Save By Buying In Bulk?

To compute how much households save by buying in bulk, we follow a process similar to that of sales. Our first task is to run a series of regressions, one for each of our product categories, to estimate the magnitude of the savings made from buying in bulk.[3] We then apply these estimated savings to our actual data on household purchases. The dependent variable is the log of the unit price of the good, and the key variables are a set of dummy variables for the quintile of product

[3] There are three food categories in which there is insufficient within-category product size variation to create size quintiles; these categories are dropped from our analysis in this part.

size in which this particular product is found (which can be interpreted relative to the lowest quintile, which is left out). We also include our sale dummy variable from the earlier regression, and a set of controls for time and region. Rather than individual product effects, we control here for brand to capture the product characteristics.[4] The model is:

$$\ln p_{iht} = \beta_j^s s_{it} + \sum_{n=2}^{5} \beta_j^n q_i^n + \eta_{jk} + \tau_j + \rho_j + e_{iht}$$

where, $q_i^p = 1$ if the pack size of product i is in the n^{th} quintile of all products in product group j, and zero otherwise, and where k indexes brand.

Purchasing larger pack sizes results in substantial estimated savings—across all product groups, the average unit price saving from buying in the largest size quintile relative to the second-smallest quintile is almost 37 percent; and from the second-largest size quintile to the second-smallest, is 28 percent.

To compute a household-level savings measure we take the same approach as for sales, and compute the savings from purchasing in the largest two size quintiles in each group using the estimated coefficients, the budget share of each product group for each household and the proportion of spending in each group in the top two size quintiles. 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. However, the smallest size categories are not very popular, and in many cases only a few infrequently purchased brands are available in this size.

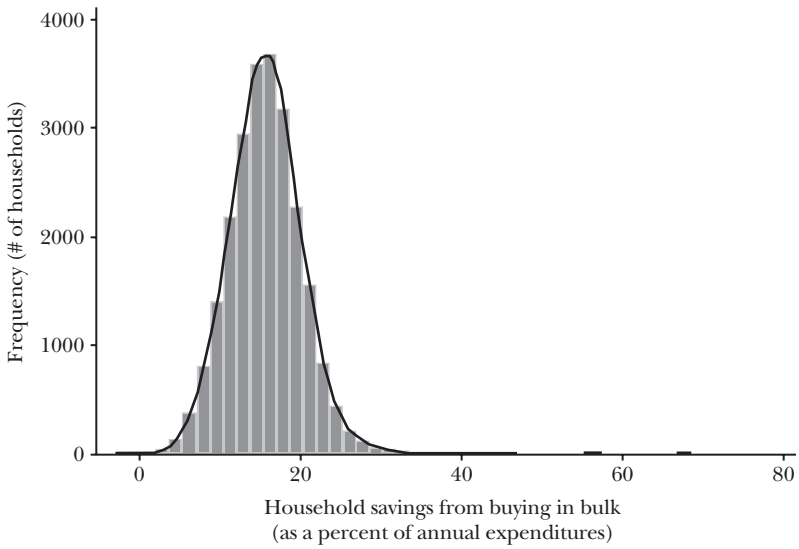
The average household saves 16 percent of its annual expenditure from buying the largest package sizes, which translates into savings of £224 per year. Figure 2 shows the distribution of savings by households, which varies from just less than 0 to almost 70 percent.[5]

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. This finding is intuitive—shopping by car reduces the cost of transporting larger pack sizes, and purchasing larger packs on average means shopping trips can be more infrequent. Households in lower income categories also save more. The demographic variables, however, are able to explain less than 1 percent of the variation across households in savings from bulk purchasing.

[4] Unlike in the sales case, we cannot use product fixed effects as clearly, as an individual barcode will always be within a given size quintile, but it is likely that other than size, products in the same brand will have similar characteristics. For some product groups, "brand" does not exist as a product characteristic—in these cases we use store fixed effects.

[5] For a small number of product groups the "savings" from buying in the larger pack sizes is negative, which in a few cases translates into a negative household savings. This outcome probably arises because in some groups larger pack sizes are for branded or higher-quality goods and this quality dimension is not adequately captured by the brand and store effects in our regression.

Figure 2
Savings Made by Each Household from Buying Larger Package Sizes



Source: Authors' calculations from 2006 TNS sample.
Note: The histogram shows the savings each household made by purchasing larger pack sizes in 2006. The sample includes 23,877 households.

Generic Brands and Product Choice

For many products, households have the choice of buying a generic or "store" brand. Such generic brands are typically cheaper, and some evidence shows that consumers substitute toward such brands as economic conditions worsen. For example, Gicheva, Hastings, and Villas-Boas (2007) provide evidence that consumers buy more store brands when gas prices rise, and Caronia (2008) finds that the income shock caused by the Argentinean 2002 peso devaluation caused a flight from branded products toward store brands.

How Much Do Households Buy Store Brands?

Many U.K. retailers offer own-brand products, which are targeted at different types of consumers. The most frequently purchased are "standard" own-brand items, which are typically priced more cheaply than national brands. In addition, many retailers offer an "economy" own brand, which is cheaper still, is packaged less attractively, and is clearly aimed at very price-conscious shoppers. These economy versions of store brands are probably closest to store brands in the United States. Some retailers also offer a higher-quality own-brand or "premium" product with a focus on quality, which is priced at a similar level to national brands.

The TNS data allow us to distinguish between regular store brands and economy versions. Economy store brands account for on average about 3.8 percent

of food expenditures in 2006, while standard store brands are much more popular at 41 percent of spending. Families with children spend more on economy store brands (5.3 of total spending for lone parents and 4.3 percent for couples with children), as do households on lower incomes (4.7 percent for those with incomes below £10,000 compared to 2.0 percent for those with incomes over £70,000). However, families with children are less likely to buy standard store brands, suggesting some substitutability between different forms of store own-brands.

How Much Do Households Save by Buying Store Brands?

We follow the same procedure as above to measure the saving from buying store brands; that is, we first run a regression for each separate product category to estimate how much is typically saved by purchasing store brands, and then apply these estimates to actual buying patterns. Again, log unit price is the dependent variable; the key explanatory variables are dummy variables for products that are economy or standard store brands and our usual controls for time and region. We also include a dummy variable for the store brand defined at the level of what is called "fascia," distinguishing different brands within the same chain. (For example, Tesco has three main brands—Tesco Extra, Tesco, and Tesco Metro/Express.)

Thus, the regression is:

$$\ln p_{iht} = \beta_j^e \text{econ}_i + \beta_j^s \text{stan}_i + \eta_{fj} + \tau_j + \rho_j + e_{iht}$$

where f indexes fascia and $\text{econ}_i = 1$ if the product is a economy store brand and $\text{stan}_i = 1$ if the product is a standard store 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 percent cheaper and the standard store brand is 25 percent cheaper.[6]

As before, we compute a household-level savings measure by weighting the category-level savings using household-specific expenditure weights according to the share of household spending in each category and the share of own-brand products in each category. Our finding is that households save on average 2 percent of their annual expenditure by buying store economy brands, with households who buy standard store brands saving on average 3.7 percent. This translates into an average saving of £25 for economy own brands and £50 for standard own brands on average. The savings from standard store brands are larger, despite the lower discounts, because the share of expenditure is much higher. There is substantial variation across households in these savings, as for sales and bulk purchases. The 10th percentile of economy savings is zero, suggesting a considerable minority of households *never* purchase economy brands; for standard store brands, the 10th percentile of savings is just 0.1 percent. At the 90th percentile, savings for economy brands is 4.9 percent, and for standard own brands, it is 7.6 percent. Interestingly, at least 5 percent of households realize negative "savings" from buying standard

[6] We consider "premium" store brands to be comparable to branded goods, so we do not consider them.

store brands, suggesting they buy generic items that are more expensive than their branded alternatives. This finding illustrates that the quality distinction between own-brand and branded items in the U.K. can sometimes be quite small. No household realizes a negative savings from economy brands, suggesting a more obvious quality differential between economy generic items and branded items.

Outlet Choice

Probably the largest change over the last decade in food retailing in the United States and the United Kingdom is the increased market share of a single firm: Walmart in U.S. food retailing and Tesco in U.K. food retailing. Walmart is the largest food retailer in the United States, with sales higher than either Kroger, Supervalu-Albertsons, or Safeway, which are the largest supermarket chains. In the United Kingdom, Tesco has similarly gained market share rapidly. The Competition Commission (2008) reports the Tesco share of grocery sales increasing from 20.2 percent in 2002 to 27.6 percent in 2007. Based on slightly different data, TNS (2008) reported a Tesco grocery market share for August 2008 of 31.6 percent. However, the fraction of expenditure in Tesco varies considerably across households and over products.

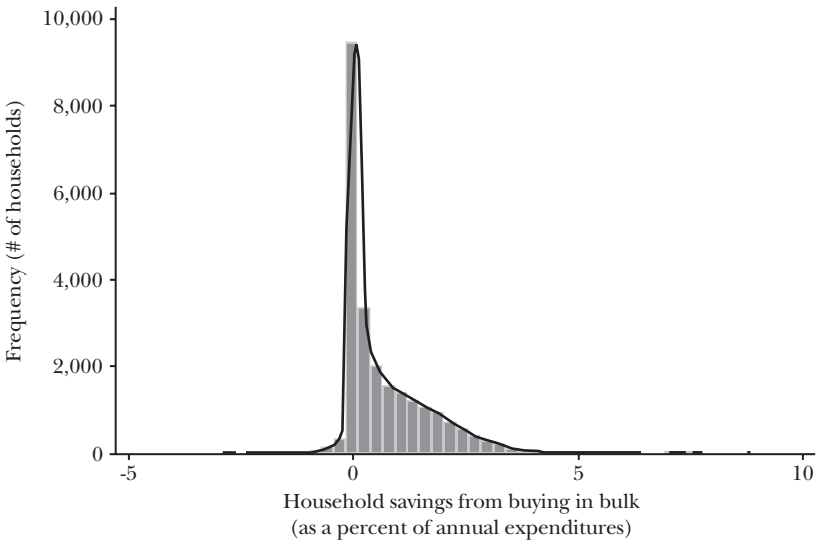
In our data, information on stores is collected via the households, who record the store of purchase for each shopping trip. For large or chain stores, the precise store is known (that is, the address is known); for corner and local stores, the specific store location is typically not known. Households vary substantially in the share of their purchases made at Tesco. The average household spends 32 percent of total annual food expenditure at Tesco. However, nearly 20 percent of households spend nothing at Tesco, and 1.7 percent of households spend all of their food budget there.[7] 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 Tesco, and 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.

To discover how much households save by shopping at Tesco, we again first estimate a series of regressions, one for each product category. The dependent variable is the log of the unit price paid by each household for a specific good at a specific time. The key explanatory variable is a dummy variable for whether the item was purchased at Tesco, and we also include the other control variables we have been including throughout. Thus, the regression takes the form:

$$\ln p_{iht} = \beta_j \text{Tesco}_{iht} + \eta_{kj} + \tau_j + \rho_j + e_{iht},$$

[7] As a comparison, using similar consumer-level data from the United States, we find that the average household spends 16 percent of its food expenditure at Walmart, while 23 percent of households do not purchase any food at Walmart.

Figure 3
Savings Made by Each Household from Shopping at Tesco



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

where $\text{Tesco}_{iht} = 1$ if household h bought good i at time t in a Tesco store. As previously, τ and ρ are time and region dummies, and η_k are brand dummies. Relative to the other dimensions, examined above, the potential savings seem much more modest. The average discount in Tesco is 1.6 percent, and the median is 1.0 percent. There also 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 before. Figure 3 shows the distribution of these savings. Again, the savings made from Tesco shopping are smaller than those from some of the other channels discussed so far: just 0.7 percent on average, or around £10 per year in cash terms; the median saving is 0.2 percent. Even those households that save the most through Tesco shopping save less than 10 percent by doing so. Indeed, around 7.5 percent of households make negative "savings" through shopping at Tesco—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 U.K. and U.S. experiences of a single retailer gaining a substantial market share in retail food sales. In the United States, Walmart's growth was largely attributable to having lower prices. In the United Kingdom, Tesco is not on average a consistently or extremely low-priced store. Across the 189 product groups in our analysis, the average savings

at Tesco is 1.6 percent; nevertheless, average savings is negative (that is, Tesco prices are higher) in 79 categories.[8]

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.

Comparing the Savings

We have described four different ways in which households can seek out lower prices when buying food: sales, bulk quantities, store brands, and outlet substitutions. We now assess the relative importance of these different dimensions.

We start with a comparison of potential savings that can be made (in terms of lower prices) based on the four sets of regressions described above. Recall that each set of regressions had 189 equations, one for each food category, to give estimates of the potential savings in each food category from sales, large package sizes, generic brands, and outlet choice, controlling as far as possible for quality differences across products. Panel A of Table 2 shows the distribution of the potential savings implied by the β coefficients in these regressions. This distribution is based entirely on differences in prices; that is, these are not weighted by the quantity of goods sold or by expenditure on different goods. The thought experiment is "how much would a household save if it switched from buying none of its purchases on sale/in bulk/on generics/at Tesco to buying all of its purchases on sale/in bulk/on generics/at Tesco?"

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

Panel B of Table 2 shows the extent to which households actually make use of these different savings channels, summarizing the distribution of household spending (as a share of total spending) on each. Expenditure shares do not track the potential savings on each share: for example, economy generic brands offer the greatest potential for savings, yet expenditure shares on these brands are relatively small. Similarly, Tesco offers small potential savings yet attracts a large share of total spending. These findings reflect the fact that potential price savings are not the only motivation for shopping choices—economy generics may well be lower quality than standard generics or branded items and are not always available in all store types; Tesco stores may compete on nonprice grounds or may simply be more or less conveniently located for different households.

[8] In fact, the U.K. retailer that is closest to Walmart is Asda—which was taken over by Walmart in 1999. Asda advertises itself as a low-priced store. However, it has not experienced the same growth in market share as Walmart has in the United States.

Table 2

Potential Savings, Expenditures on Methods of Saving, and Actual Saving

Panel A: Potential Savings from Each Channel

Savings channel	*Mean*	*Std. dev.*	*10th percentile of saving*	*Median saving*	*90th percentile of saving*
Sale	21.7%	6.0%	14.1%	22.4%	38.8%
4th size quintile	28.1%	27.7%	−2.2%	27.1%	65.2%
5th 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%

Panel B: Household Expenditure Shares on Various Savings Channels

Savings channel	*Mean saving*	*Std. dev.*	*10th percentile of saving*	*Median saving*	*90th percentile of saving*
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%

Panel C: Household Savings from Various Channels as % of Expenditure

Savings channel	*Mean saving*	*Std. dev.*	*10th percentile of saving*	*Median saving*	*90th percentile of saving*
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%

Panel D: Household Savings from Various Channels in £ per Year

Savings channel	*Mean saving*	*Std. dev.*	*10th percentile of saving*	*Median percentile of saving*	*90th percentile of saving*
Sale	£96	£73	£20	£79	£193
Largest two size quintiles	£224	£160	£52	£192	£437
Economy generic brand	£25	£36	£1	£12	£64
Standard generic brand	£50	£70	£1	£42	£126
Tesco purchase	£10	£17	£0	£2	£30

Source: Authors' calculations from 2006 TNS sample.

Notes: In Panel A, potential savings are based on the coefficients obtained from the earlier four regressions in the text; 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.

Looking more closely at this data, households 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 goods (a similar finding for U.S. data is in Leibtag and Kaufman, 2003). Similarly, households that buy economy generic also tend to purchase in bulk.

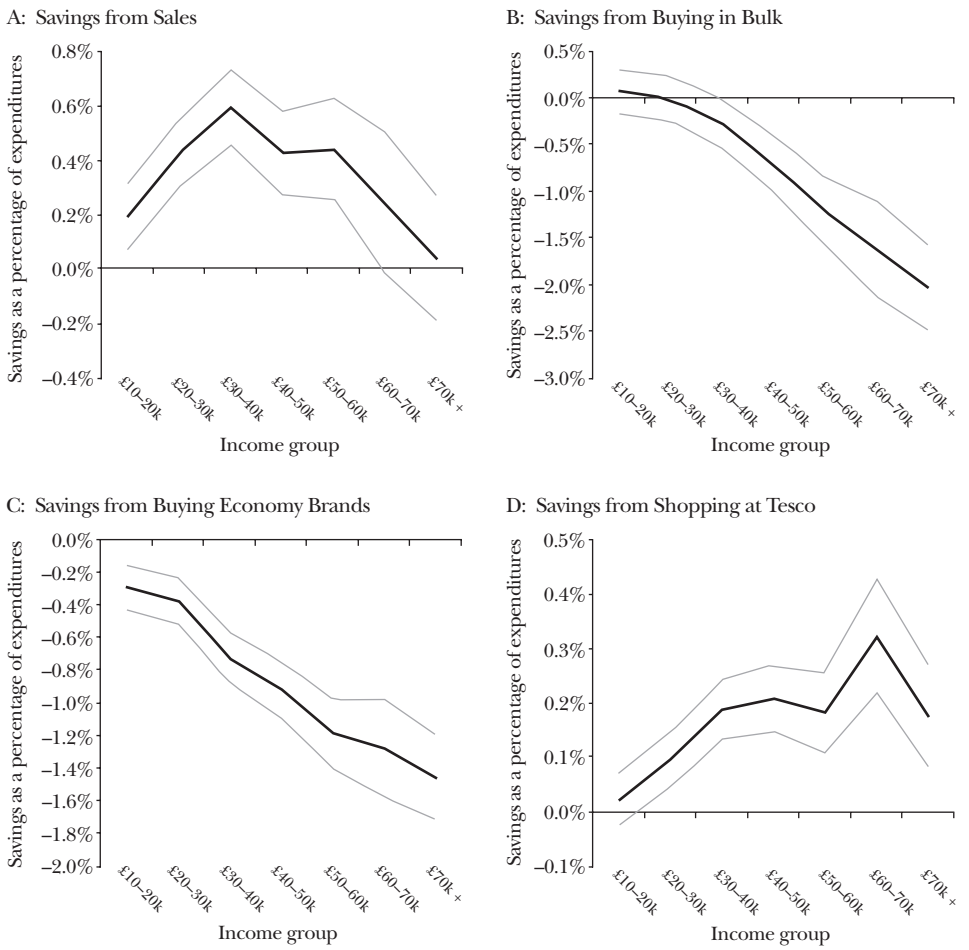
Panels C and D of Table 2 combine the potential savings in Panel A with the household expenditure choices in Panel B to estimate the *actual* savings that households obtain from these different choices. In Panel C, we express this savings in terms of percentage of expenditure, while in Panel D, we present the savings in British pounds per year. The savings from bulk purchasing are the largest, followed by savings from sales, purchases of generics (standard then economy), and shopping at Tesco. The relatively large savings from bulk purchases reflect both their large potential saving (Panel A) and their high expenditure share (Panel B). For standard generic brands, both the potential savings and expenditure shares were relatively high but the actual savings are quite low. This outcome occurs because the item categories where the generic expenditure share is higher are also those where the potential savings are relatively low (which may imply that the difference between generic brands and branded products is also quite low, meaning households are more willing to substitute toward generic brands).

Some warnings should be issued about interpreting these estimates. First, they do not account for the *costs* of savings. In the case of sales and bulk purchases, these costs involve storage and transport, and potential depreciation in the quality of the product over time. For generic brands, the cost could include quality differences. For outlet choice, travel costs as well as potential quality differences between stores could be important. These savings could therefore be interpreted as "gross" savings, but the "net" savings from the different channels may not be of the same magnitude or even rank. An additional warning is that the savings measures overlap. For example, if products sold in Tesco are more frequently on sale, the corresponding savings are potentially counted twice, both in Tesco and in the sales measures; one could think of estimating these savings jointly in a single model, but it is then harder to interpret the individual effects.

The savings vary with household demographics. Retired households tend to be less likely to save on any of the four dimensions relative to other household groups. Childless couples and single adults are next, with an intermediate willingness to take advantage of these methods of saving. Families with children tend to save more through each of the savings channels than those without children, with the highest relative gains for bulk discounting and sales. This pattern makes sense: larger families have larger consumption needs and so they are more likely to benefit from sales and bulk purchasing for a given storage cost.

Figure 4 shows the estimated savings from each channel according to household income relative to households with low incomes below £10,000 per year. Bulk purchases and economy brands follow the same pattern, with savings decreasing monotonically with income, showing that these channels are used most by low-income households. By contrast, saving from sales (and, to a lesser extent, from

Figure 4
Savings by Income Group
(relative to households with incomes under £10,000)



Source: Calculated from the 2006 TNS sample.
Note: The dark line in the middle of each graph represents the estimated average savings and the lighter lines represent the 95% confidence intervals.

Tesco purchases) are non-monotonic with income: low-income households do not seem to save with these methods, maybe because they are unable, and neither do high-income households, maybe because it is not worth their time.

Implications for Price Indexes

It is well known that consumer choices can generate bias in standard price indexes. The most common example is the well-known substitution bias, which is

easiest to illustrate in its pure form. Suppose a consumer consumes two products, rice and potatoes. Initially both products cost $1 per pound, and the consumer chooses one pound of each. In the second period, the price of rice increases to $2, but the price of potatoes is unchanged. Using a fixed bundle consisting of one pound of each product, the measured increase in the price index is 50 percent. If rice and potatoes are perfect substitutes, then the consumer will simply substitute potatoes instead of the rice with no loss of utility, and so the price index will considerably overstate the true change in the cost of living. However, assume that the two goods are not perfect substitutes, and as relative prices change, the consumer substitutes from rice to potatoes and consumes, for example, 1.5 pounds of potatoes and only 0.5 pounds of rice. The actual increase in expenditure is now only 25 percent, which is also the measured price increase using the second-period bundle—measured this way, the possibility for substitution would overestimate the price increase. However, it should be remembered that the consumer has lost utility being pushed to substitute from rice to potatoes, and the measured change in expenditure does not take this utility loss into account, so the 25 percent increase in expenditure underestimates the loss in utility that has occurred.

Similar logic suggests that a price index based on a fixed basket of goods will miss a shift toward the cheaper own brand, especially if store branded products are under-sampled in the bundle used to compute the price index. (We do not know whether branded products are over- or under-sampled in the U.K. food price index; our point here is a general methodological one that if they are, then this sort of bias will result.)

For the purposes of constructing price indexes, the implications of "outlet substitution"—that is, the effect of consumers shifting their purchases to different and less-expensive retail outlets—have been well documented (for example, Reinsdorf, 1994; Boskin, Dulberger, Gordon, Grilliches, and Jorgenson, 1998; Hausman and Leibtag, 2004). In the United States, the way that prices have traditionally been collected may not fully capture price changes in retailers like Walmart, or the approach treats the price difference between Walmart and others as a reflection of quality, rather than a genuine lower price. The situation is somewhat different in the United Kingdom; the national statistical agency, the Office for National Statistics, does attempt to reflect the annual changes in supermarket share when calculating food and grocery price indices.

Consumer stockpiling in reaction to sales also has implications for the standard measurement of price indices. It implies a separation between a standard price index based on purchase prices and a consumption-based price index that should account for the ability to store the product. 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. When product B goes on sale, the consumer purchases four units and

consumes one unit each week over the next few weeks. We assume the consumer does not increase consumption in response to the sale price; we also ignore the consumer's discount factor. The consumer saves on each of the units that is purchased and stored; for the last unit, for example, the consumer pays $1 and stores it for three periods at a total cost of $0.75, for a total saving of $0.25 relative to buying the product at the regular $2 price. The consumer, however, will not save from buying additional units above the four purchased because of the storage costs. If the consumer bought a fifth unit on sale, the storage costs for the unit for four periods will exactly equal the savings (and so 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 price during nonsale periods, so when consuming one unit of each product, the base is $4. The consumption-based price index for the period of the sale and the following weeks is $(2.00 + 1.00)/4.00 = 0.75$ for the sale period; $(2.00 + 1.25)/4.00 = 0.8125$ for the next week; $(2.00+1.50)/4.00 = 0.875$ for the next week; $(2.00+1.75)/4.00 = 0.9375$ for the next week; and then 1.00 for every following week when the sale has ended and no further consumption of stored products occurs. A standard price index will capture some price reduction in the week of the sale, depending on the quantity weight used to compute the index. However, a standard price 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. Aggregation across weeks to construct a monthly price will also not solve the problem even if the timing is captured properly. In this case, the aggregation will overestimate the benefits from purchasing on sale because it will ignore the storage cost.

The issues in measurement of a price index with bulk goods and varying unit prices are very similar to those that arise in the case of sales. First, there is an issue of whether statistical agencies correctly sample all the relevant prices. In the United Kingdom, the specification of food items to be collected as part of the basket of items used to calculate inflation rates typically contains an exact size that must be priced. It is unusual for different sizes of the same product to be sampled and 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. Without sampling different sizes of goods, statistical agencies will miss this effect and compute an index that overestimates price increases.

Occasionally, statistical agencies will sample different package sizes (when products change, for example, and firms replace a pack of one size with another size). A common response by statistical agencies is to link the price from different sizes using unit values. For example, if a pack of 100g priced normally at £1 is not available, but instead the item is only available as a 200g pack priced at £1.80, the agency will "pretend" the price is £0.90 per 100g. However, this practice is only justified if prices are linear in package size, which is rarely the case. Otherwise, as pointed out by Triplett (2004), this approach will underestimate the price index.

The savings measures we presented above do not allow us to quantify directly

the effect of the various choices made by consumers on the measurement of the price index. To measure these effects we need to compute a standard price index, which ignores the substitution effects, and compare it to a price index that accounts for the different forms of substitution. Previous work has estimated the effect of substitution bias as well as outlet bias, and has found it to be significant. Our results suggest that the bias due to stockpiling and bulk purchases have the potential to be of the same order of magnitude.

Concluding Comments

In this paper, we have documented the potential and actual savings that consumers in the U.K. realize from various forms of substitution. The results suggest that savings from sales and bulk buying are of a similar order of magnitude to those due to economy brands and outlet choice. The data in this paper come from the United Kingdom; an interesting question is how these findings are comparable to other countries. Preliminary analysis of similar data from the United States suggests similar findings, although with some interesting differences. Sales and bulk purchases, and the savings they entail, seem to be even more significant in the United States. This pattern should not be surprising. On average, U.S. homes are larger and consumers tend to shop more with cars, so transport and storage costs are lower. As a result, larger sizes not only tend to be more often purchased in the United States but also *offered* more by stores; for example, a gallon-size ice-cream pack is quite popular in the United States but is typically not available in the United Kingdom. Store brands are also different in the United States and typically less important than in the United Kingdom. As discussed earlier, the savings from Walmart seem to be more significant than the savings from shopping at Tesco. Taking these factors together, our conclusion that savings from sales and bulk purchases are important probably holds with even greater force in the U.S. economy. Further cross-national comparisons of how these savings vary across households would be interesting; Aguiar and Hurst (2007), for example, used scanner data similar to the data used in this study, and found that poorer and older households typically paid lower prices for identical products than younger, richer households and that families with children paid the highest prices. Given that sales, bulk purchasing, store choice, and branding all influence the price paid for essentially identical items, these channels of saving will be crucial in determining who pays more and whether the pattern of relative prices differs across countries.

Data on household consumption that include barcodes for specific products, details about stores, and a range of household information are a relatively recent development for academic economists. Economists are still learning how to explore and exploit their possibilities. Future research may, for example, take the costs of transportation and storage into account, and eventually be able to estimate a full demand system incorporating these different elements of purchasing behaviors. Other research will doubtless look at these data from the perspective of nutritional

choices, or from the perspective of how much flexibility households have to adjust their costs by changing their shopping or purchasing behavior.

This work also has implications for improving the consumer price index. A standard price index based on a fixed basket of goods will overstate the rise in the true cost of living because it does not properly consider sales and bulk purchasing. According to our measures, the extent of this bias might be of the same or even greater magnitude than the better-known substitution and outlet biases. Yet biases due to sales and bulk purchasing have received much less attention in the price index literature. We second Triplett (2003) in suggesting that this is a fruitful area for future research.

■ *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. We wish to thank the editors of this journal and seminar participants at the U.S. Bureau of Economic Analysis and the 2008 World Congress on National Accounts and Economic Performance Measures for Nations for their comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Department of Agriculture.*

References

Aguiar, Mark, and Erik Hurst. 2007. "Life-Cycle Prices and Production." *American Economic Review,* 97(5): 1533–59.

Baye, Michael R. 1985. "Price Dispersion and Functional Price Indices." *Econometrica,* 53(1): 213–23.

Boizot, Christine, Jean-Marc Robin, and Michael Visser. 2001. "The Demand for Food Products: An Analysis of Interpurchase Times and Purchased Quantities." *The Economic Journal,* 111(470): 391–419.

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson. 1998. "Consumer Prices, the Consumer Price Index, and the Cost of Living." *Journal of Economic Perspectives,* 12(1): 3–26.

Caronia, Mateo. 2008. "The Microeconomics of Large Devaluations." Unpublished paper, Northwestern University.

Competition Commission. 2008. Appendix 3.1 of "Market Investigation into the Supply of Groceries in the UK." http://www.competition-commission.org.uk/rep_pub/reports/2008/fulltext/538_3_1.pdf.

Crawford, Ian. 1996. "UK Household Cost-of-Living Indices, 1979–92." In *New Inequalities: The Changing Distribution of Income and Wealth in the United Kingdom,* ed. John Hills, chap. 4. Cambridge University Press.

Crawford, Ian, and Zoë Smith. 2002. "Distributional Aspects of Inflation." IFS Commentary 90. London: Institute for Fiscal Studies. http://www.ifs.org.uk/comms/comm90.pdf.

Deaton, Angus. 1998. "Getting Price Right: What Should Be Done?" *The Journal of Economic Perspectives,* 12(1): 37–46.

Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior.* Cambridge University Press.

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. 1993. "The Early History of Price Index Research." In *Essays in Index Number Theory,* Volume 1, Contributions to Economic Analysis 217, ed. W. E. Diewert and

A. O. Nakamura, 33–65. Amsterdam: North Holland.

Feenstra, Robert C., and Matthew D. Shapiro. 2003. "High-Frequency Substitution and the Measurement of Price Indexes." In *Scanner Data and Price Indexes,* 123–150. University of Chicago Press.

Gicheva, Dora, Justine Hastings, and Sofia B. Villas-Boas. 2007. "Revisiting the Income Effect: Gasoline Prices and Grocery Purchases." NBER Working Paper 13614.

▶**Hausman, Jerry.** 2003. "Sources of Bias and Solutions to Bias in the Consumer Price Index." *The Journal of Economic Perspectives,* 17(1): 23–44.

Hausman, Jerry, and Ephraim Leibtag. 2004. "CPI Bias from Supercenters: Does the BLS Know That Wal-Mart Exists?" NBER Working Paper 10712.

▶**Hausman, Jerry, and Ephriam Leibtag.** 2007. "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart." *Journal of Applied Econometrics,* 22(7): 1157–77.

Hendel, Igal, and Aviv Nevo. 2006a. "Sales and Consumer Inventory." *The RAND Journal of Economics,* 37(3): 543–61.

▶**Hendel, Igal, and Aviv Nevo.** 2006b. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica,* 74(6): 1637–73.

Leibtag, Ephram, and Phil R. 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.

Leicester, Andrew, Cormac O'Dea, and Zoë Oldfield. 2008. "The Inflation Experience of Older Households." IFS Commentary 106. London: Institute for Fiscal Studies. http://www.ifs.org.uk/comms/comm106.pdf.

Office for National Statistics. 2007. "Consumer Prices Index Technical Manual." http://www.statistics.gov.uk/downloads/theme_economy/CPI_Technical_Manual.pdf.

Pollak, Robert A. 1998. "The Consumer Price Index: A Research Agenda and Three Proposals." *The Journal of Economic Perspectives,* 12(1): 69–78.

▶**Reinsdorf, Marshall.** 1994. "The Effect of Price Dispersion on Cost of Living Indexes." *International Economic Review,* 35(1): 137–49.

▶**Schultze, Charles L.** 2003. "The Consumer Price Index: Conceptual Issues and Practical Suggestions." *Journal of Economic Perspectives,* 17(1): 3–22.

TNS. 2008. "Price Now Clearly Driving Market Shares." Press Release, August, 19. http://www.tnsglobal.com/_assets/files/TNS_Market_Research_Supermarket_Share_data_mmentary.pdf.

Triplett, Jack E. 2003. "Using Scanner Data in Consumer Price Indexes: Some Neglected Conceptual Considerations." In *Scanner Data and Price Indexes,* ed R. C. Feenstra and M. D. Shapiro, p 151–162. University of Chicago Press.

Triplett, Jack E. 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.