

This PDF is a selection from a published volume from the
National Bureau of Economic Research

Volume Title: Scanner Data and Price Indexes

Volume Author/Editor: Robert C. Feenstra and Matthew
D. Shapiro, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-23965-9

Volume URL: <http://www.nber.org/books/feen03-1>

Conference Date: September 15-16, 2000

Publication Date: January 2003

Title: Using Scanner Data to Improve the Quality of Measurement
in the Consumer Price Index

Author: William J. Hawkes, Frank W. Piotrowski

URL: <http://www.nber.org/chapters/c9727>

Using Scanner Data to Improve the Quality of Measurement in the Consumer Price Index

William J. Hawkes and Frank W. Piotrowski

“The only emperor is the emperor of ice cream.”
—Wallace Stevens, 1922

1.1 Introduction

Twenty-five years ago, a major technological change swept through U.S. retailing and left the field of marketing research profoundly altered in its wake. Since then, the same tidal wave has moved across most of Europe and the developed countries of Asia and Latin America as well. This technological change involved the source-coding of most fast-moving packaged consumer goods by their manufacturers, using the newly developed Universal Product Code. It also involved the installation, by retailers, of electronic scanning equipment at the checkout counter to “read” and record each item purchased in the store. Subsequently, other kinds of electronic point-of-sale (EPOS) systems were introduced to record transactions of durable goods (e.g., toasters and refrigerators) that had not been source-coded at the manufacturer level.

As a result of this new method of data collection, market research firms were able to obtain and summarize information on consumer sales and retail prices in a much faster, more detailed, and more cost-efficient manner than before. Scanner-based data quickly became the “common language” used by manufacturers, retailers, and marketing research companies to describe and interpret developments in the retail marketplace.

As shown in table 1.1 below, currently in the United States around 10 percent of total consumer expenditures, and around one-quarter of consumer

William J. Hawkes is the former chief statistical officer of ACNielsen and is presently a marketing research consultant. Frank W. Piotrowski is vice-president and chief of the measurement science department of ACNielsen.

The authors thank Bill Cook and Dave Richardson for their comments on an earlier draft of this paper.

Table 1.1 Consumer Expenditures Scannable through Supermarkets, Drugstores, and Mass Merchandisers in CIP (as of December 1999)

CPI Expenditure Categories	Number of CPI Categories	% of Consumer Expenditures Included in These Categories	
		All Expenditures	Expenditures for Goods
Food at home	53	9.6	22.8
Perishables	19	3.2	7.5
Scannable edibles (row 1 less row 2)	34	6.4	15.2
Other supermarket item strata	9	3.8	9.0
Total supermarket scannables (row 3 plus row 4)	43	10.2	24.2

Table 1.2 A Brief Lexical Concordance between Market Research and Consumer Price Index Terminology

Marketing Research	Price Index
Retail selling price	Price (p)
Consumer sales equivalent units	Quantity (q)
Consumer sales dollars	Expenditures (pq)

expenditures on goods (as opposed to services, such as housing services and haircuts), are made in categories that can, in large measure, be represented through scanning data obtained from supermarkets, mass merchandisers, and drugstores.

Price index theorists and practitioners have long been observing and measuring the same consumer behavior as have marketing researchers, even though they sometimes use a slightly different vocabulary to describe the transactions that they are studying, as shown in table 1.2 below. Both price index and marketing research theorists are concerned with the response, or elasticity, of consumer purchases to changes in retail price. Despite this commonality of interest, the public sector involved in producing consumer price indexes has been slower than the private sector of marketing research to utilize and benefit from this new technology. Recently, however, government statistical agencies in many countries have begun to investigate and utilize this new source of data in their consumer price indexes.

This paper, written from the perspective of market researchers who have spent the past twenty-five years working with scanner data, discusses specific ways in which the quality of consumer price information can be improved using this new data source. It also shows how the measurement of product quality can be enhanced through the use of these data. The paper will make use of actual scanner data for a particular product category and

will, we believe for the first time, present and discuss a “total U.S.” simulated price index for a specific Consumer Price Index (CPI) “food-at-home” commodity (item stratum), comparing the results with the corresponding “urban U.S.” figures produced by the U.S. Bureau of Labor Statistics. The paper concludes with a discussion of data aggregation issues in CPI construction.

1.2 How Scanner Data Can Improve the Quality of Consumer Price Indexes’ Measurements

The potential benefits from using scanner or other EPOS data in CPIs can be grouped into three categories: (a) more data and, consequently, less variance; (b) better data and, consequently, less bias; and (c) better methods. We shall consider each in turn.

1.2.1 More Data

In most developed countries, scanner data for supermarket items are based on a number of data points (outlets, items, and weeks) that exceed those currently used in these countries’ CPIs by several orders of magnitude—generally in the range of 1,000 to 1. To cite one example, CPI data for the U.S. breakfast cereals “item stratum” are based on around 675 individual price observations, one observation per month for two or three items per store in a sample of around 300 individual outlets. In contrast, ACNielsen scanner data for supermarkets are based on four or five price observations per month for over 200 cereal items per store in a sample of around 3,000 supermarkets. If the scanner data reporting period, for CPI purposes, is constrained to the first three weeks of each month, then scanner data consist of 1,800,000 price observations per month ($3 \times 200 \times 3000$). This is 2,700 times as many price observations as are currently being obtained for breakfast cereals (1,800,000 divided by 675) in the CPI program. These price records are, in every case, accompanied by actual quantities sold, each week, in each supermarket for each item, in contrast to the implicit quantity weights for each price observation in the CPI, which usually remain unchanged for four or five years.

1.2.2 Better Data

Even within the framework of current CPI designs in most countries, scanner data provide the opportunity for significant quality improvement in terms of bias reduction along several dimensions:

1. *Sample outlet selection.* Scanner retail outlet samples are generally selected from a well-defined frame that lists all universe supermarkets, mass merchandisers, and large drugstores. For the U.S. CPI, sample retail outlets are drawn from a list of “point-of-purchase” outlets obtained from a

sample of around 3,500 households nationally. In many other countries, individual CPI retail outlets are selected from incomplete or geographically restricted frames.

2. *Outlet sample updating.* Scanner samples are usually designed to incorporate new outlets with minimum delay. Consumer price index outlet samples, in most countries, are refreshed or replaced only at periodic intervals, generally once every several years.

3. *Item selection.* Designation of individual items to be priced is carried out, in most countries, in one of two ways:

- Selection of items within a store with probability proportionate to measures of size, in theory based on actual expenditures but often based on shop owners' memory or estimates, or shelf space, or some other "proxy" means. This is the procedure currently used in the United States.
- Purposive or judgmental selection of items by product characteristics, with specific products, or varieties then chosen in the field or, in certain instances, centrally designated. This is the procedure currently used in the United Kingdom and in Canada.

In contrast, scanner data are provided for every item in every category handled and scanned by the store. At the very least, scanner data could be used to check out the validity of the "purposive" item selection methods used in areas such as the United Kingdom and Canada.

4. *Item updating.* Scanner data automatically include all new items appearing in each sampled retail outlet, generally in "real time," or with a delay of a few weeks at most to allow for a full product description to be defined for each new item code. In contrast, CPI new items are brought in only when existing items are discontinued at the individual outlet level or when a complete item reselection is carried out, generally once every several years.¹ Both new items and new outlets are generally linked to the previous price index generated by the old items and old outlets, with no allowance for differences in price levels between new and old, except for item strata where explicit hedonic adjustments can be made. In an important paper, Reinsdorf (1993) showed that the combined effect of new items and new outlets in the U.S. CPI for food-at-home items was to reduce average price levels for food-at-home commodities by 0.25 percent per year.

5. *Better lower-level (within item stratum) expenditure weights.* Lower-level item expenditure or quantity weights used in CPI construction are generally based on estimates made at infrequent intervals from a variety of sources: consumers' recall, consumer purchase diaries, and shopkeepers' estimates. Use of out-of-date expenditure weights is likely to result in an

1. When the CPI's telephone point-of-purchase surveys (POPS) rotation scheme is fully operational, POPS categories will be reselected on an ongoing basis, with a complete outlet or item recycling at four- or five-year intervals.

overstatement of inflation. In contrast, scanner data provide current, up-to-date expenditure weights each week.

1.2.3 Improved CPI Scope, Definitions, or Methods

For many years, price theorists have written of “superlative” or “ideal” price indexes more as a concept than as a reality, since current period quantities have generally not been available. In a scanning environment, this restriction no longer exists. Accordingly, a number of methodological questions come immediately to mind. Scanning data can help provide answers to these important questions:

1. Should expenditure weights be computed for each specific month, or is it better to use more stable weights (e.g., for the most recent year), on the assumption that trading off some temporal “characteristicity” in weights will be more than offset by reducing the greater intransitivity associated with chaining true Fisher or Törnqvist indexes? Triplett (1998) has pointed out the instability and intransitivity that can result from the chaining together of even superlative indexes when quantities and prices change abruptly from month to month. Recent studies at the Bureau of Labor Statistics (BLS) seem to support the desirability of using the most recent annual quantity weights with scanning data, updated each year, rather than using the monthly quantity weights that accompany the monthly prices. Silver (1995) and Diewert (2000b) advocate constructing an annual moving weight to avoid seasonal bounce.

2. To what extent should weekly sales and quantities for individual items in individual outlets be aggregated to construct “unit values” across items, outlets, and time? Although a unit value index fails the identity and proportionality tests and thus, according to Balk (1998), “cannot be called a price index,” it is also true that the unit value index passes the circularity or transitivity test (cited by Balk as a further axiomatic test for a price index), which the Fisher and the Törnqvist, both considered superlative, are guaranteed to fail!² The same is true of the “consistency in aggregation” test: the unit value passes this test, whereas the above superlative indexes fail it. Moreover, the unit value is (or can be) automatically adjusted for new items and new outlets. The question of aggregation is discussed at greater length in section 1.5 of this paper.

3. How can scanner data be used to improve the stratum weights used to aggregate the city-by-category lower-level indexes into higher-level indexes? In a recent paper, Diewert (2000b, 26) has cited the “large measurement errors” in these weights as a serious problem in producing an accurate national price index. Certainly the current need to produce expenditure

2. See discussions in Frisch (1930) and Olt (1996). “Transitivity” simply means that the product of comparative price levels between countries A and B, and between countries B and C, will equal the comparative price level between countries A and C.

weights for the 1,292 “scannable food-at-home” strata (thirty-eight geographic areas times thirty-four item strata) places a high degree of stress on the sample of roughly 10,000 households each providing two weeks of expenditure data in the U.S. Consumer Expenditure Survey.

An obvious alternative is to make use of aggregate outlet-based scanner data for this purpose. As shown in table 1.3 below, Consumer Expenditure Survey data for 1998 and 1997 agree fairly closely at the national level with ACNielsen ScanTrack data for one product category that will be examined extensively in this paper, but this apparent agreement obscures the fact that the Consumer Expenditure data are probably around 16 percent too low when nonscanning outlet types, sales taxes, and Alaska and Hawaii are taken into account. It would seem an easier matter to estimate this missing 16 percent from the BLS point-of-purchase survey or from household-based scanner data than to estimate the entire 100 percent from the Consumer Expenditure Survey.

In addition to using scanner-based data, ACNielsen also measures consumer sales directly through a 55,000 household sample. The sampled households are supplied a hand-held scanner with a downloaded list of retail establishments in their neighborhood. Households are instructed to scan all UPC coded items they purchase and identify the retail establishment where the purchase occurred. The ACNielsen HomeScan service indicates that 89 percent of ice cream-type products are purchased in traditional supermarkets. This would support the claim that the Consumer Expenditure data may underestimate true sales.

In any event, it would be a useful exercise to carry out this comparison for the thirty-three other scannable edible item strata as well, first at the national level and eventually at the geographic area level as well.

4. How should product categories be defined and subdivided in such a

Table 1.3 Total U.S. Sales of Ice Cream and Related Products, 1997 and 1998 (based on Consumer Expenditure Survey and ScanTrack; sales in \$millions)

Year	CES	ScanTrack	ScanTrack vs. CES
1997	5586	5580	-0.1%
1998	5682	5810	+2.3%
1997-98 % change	+1.7	+4.1	+2.4
Adjustments to make 1998 ScanTrack comparable to CES			
Outlets other than supermarkets	14%		
Alaska and Hawaii	1%		
Sales taxes	2%		
<i>Subtotal</i>	17%		
Reported difference	2%		
Adjusted difference, ScanTrack vs. CES	19%		

way as to maximize the temporal and geographic transitivity of the resulting indexes and to avoid item “churn” (i.e., excessive item turnover) while enhancing the comparability of items across time and across geography?

5. The foregoing leads to the next question: how many item strata should there be in the CPI, and how should they be defined? The current fifty-three food-at-home item strata have changed very little in the past twenty years. Essentially, the number of item strata has been determined by the limitations of the 10,000-household Consumer Expenditure Survey sample, the need for continuity and for seasonal adjustment, and the data collection budget. If we were freed from these constraints, how might we want to proceed?

An orthogonal view, or at least an alternative view, of the food-at-home universe is shown in table 1.4. Note that the ACNielsen item structure tends to reflect the “department” or physical layout of the typical supermarket, whereas the BLS structure is driven by classification of item complements and substitutes. Thus, ACNielsen classifies Ice Cream as a frozen food, rather than as a dairy product; ACNielsen classifies vegetables in three different places: frozen, canned, and perishable, just as they are found in the supermarket, whereas BLS groups them together under the fruits and vegetables heading and then divides them into subgroups by form.

Obviously, there is no right or wrong in these taxonomies, but there are differences. Of greater interest is that ACNielsen has further subdivided its 64 categories into 603 separate modules. To some extent BLS also carries out further subdivisions of its 34 item strata, both explicitly into its 44 “entry-level items” (ELIs), and implicitly (for sampling purposes of individual items to be priced in specified outlets) through a “disaggregation” procedure that partitions “entry-level items” down into successively smaller subgroupings based on criteria such as form and package size. Through this

Table 1.4 How ACNielsen and BLS View the World of Food-at-Home

	ACNielsen	BLS
Division of food-at-home items	6 departments (dry grocery, frozen foods, dairy, deli, packaged meats, perishables); 64 categories (excludes fresh meat and produce); 603 modules (excludes fresh meat and produce)	6 commodity groups (cereals & bakery products, meat, poultry, eggs & fish, dairy, fruits & vegetables, nonalcoholic beverages, other); 18 expenditure classes; 53 item strata (of which 34 exclude fresh meat and produce); 65 entry-level items (of which 44 exclude fresh meat and produce)
Division of “ice cream and related products”	4 modules in 2 categories under “frozen foods”	1 entry-level item in 1 item stratum under “dairy”
Classification of ice cream	frozen food	dairy product

process, the 34 “scannable edible” item strata are thus further subdivided into 79 “clusters” or mutually exclusive and exhaustive building blocks. These 79 clusters might be a natural next level for category detail to be used as input to higher-level index aggregates in the CPI if resources allowed.

However, the conceptual distinction between what should be a cluster and what should merely be a disaggregation criterion is not always clear. In the CPI example above, fresh fruits and vegetables are divided into eight separate item strata, whereas canned vegetables share an item stratum with canned fruits, although the item stratum is subdivided two separate clusters, one for fruits and one for vegetables. ACNielsen has sixty-nine separate modules for canned fruits and vegetables. Is sixty-nine too many? Is one too few? Is two?

With scanner data, alternative partitionings of the entire product space can be carried out on an experimental basis to determine what the optimal clustering rules should be, in order to produce the most reliable and efficient overall index at various geographical levels.

6. There has been a blizzard of papers written over the past decade, both within BLS and from interested observers, on the subject of within-stratum price elasticity. These papers have been written in the context of two very real and important issues:

- Did the Laspeyres assumption of a Leontief preference function lead to a serious upward bias in the CPI prior to 1999?
- Has the geometric mean formula used in lower-level index construction starting in 1999, which assumes Cobb-Douglas preferences and unit elasticities, fully corrected the problem?

With few exceptions, such as a reference by Moulton (1993) to one earlier article by Tellis (1988), most of the literature on this subject has been informed more by opinions or “a prioristic” arguments than by factual evidence. Until recent years, in the absence of good information on quantities to accompany price information, economists had little choice but to speculate about these elasticities. However, scanner data provide a veritable library of information pertaining to individual item elasticities and cross-elasticities. In the United States and many other countries, manufacturers and retailers conduct and commission literally hundreds of studies each year in which outlet-level scanner data are used to measure these elasticities for the purpose of planning their pricing and promotional strategies. Certainly the public sector (e.g., the BLS and its counterpart agencies in other countries) could and should make use of the same data source in order to answer these critical questions on price elasticities, functional forms, and biases.

7. A final opportunity for improving the quality of measurement is one that also improves the measurement of quality. By “measurement of quality” we are referring to the measurement of product quality for use in price index construction. This can be done by relating bundles of product char-

acteristics to the price of each product in such a way as to generate hedonic quality coefficients on each characteristic and thereby enable new items to be linked together with old or existing items in a price index without loss of continuity. This will be the topic of the next section.

1.3 How Scanner Data Can Improve the Measurement of (Product) Quality in the Consumer Price Index

There are (at least) four different methods³ that have been proposed or used to adjust for quality changes in the CPI framework, or for quality differences in an interarea price comparison framework.

1.3.1 Direct Adjustment

This procedure is normally used when the new product differs from the old only by a factor of size—generally package size. For example, if a manufacturer downsizes from a 220-gram size to a 200 gram size but leaves the price unchanged, this would be treated as a 10 percent price increase under a “direct adjustment” procedure. This method assumes a linear, nonzero intercept pricing model. This assumption is generally a good one when the size change is small, but it tends to break down when substantially different package sizes are being compared (see table 1.5 for some real examples).

1.3.2 Manufacturer’s Cost Adjustment

When a new product attribute, or a change in product attributes, has been introduced, a value on this new or changed attribute can sometimes be obtained from manufacturer’s cost information. The accompanying change in retail price can then be partitioned into a pure price increase component and a quality change component, based on the manufacturer’s cost information. This method has been used, for example, in making some quality change adjustments for new automobiles.

1.3.3 Hedonic Adjustment

In this method, which dates back at least to Court (1939), implicit price components for each of a bundle of product characteristics are determined by a regression procedure that expresses the price of a product as a function of the coefficients associated with each characteristic. The price of a new product (or different product) can then be compared with that of the previously existing product when one utilizes these coefficients.

There are three potential difficulties with this method. One is that the he-

3. Jack Triplett tells me (WJH) that there are actually a lot more than four, and he is probably right. Most of the other candidates strike me as variants of the four that we have listed. (see, for example, Wynne and Sigalla 1994). Methods have also been developed to partition an observed difference into a pure price change component and a quality change component. Greenlees (2000) discusses these issues in detail, as does Triplett (1997).

Table 1.5 Examples of Nonlinear Pricing: Organization “Y”

Product and Size	Price (\$)	Price per Unit (\$)
Brand “A” Corn Oil		
16 oz.	1.53	3.06 (per quart)
24 oz.	1.95	2.60
32 oz.	2.45	2.45
48 oz.	3.15	2.10
64 oz.	4.13	2.07
128 oz.	7.87	1.97
Brand “B” Coffee		
13 oz.	2.71	3.34 (per pound)
26 oz.	5.23	3.22
39 oz.	7.59	3.11
Brand “C” Rice		
1 lb.	0.63	0.63 (per pound)
2 lb.	1.19	0.59
5 lb.	2.75	0.55
10 lb.	4.79	0.48
Brand “D” Peanut Butter		
12 oz.	1.65	2.20 (per pound)
18 oz.	1.89	1.68
28 oz.	3.09	1.77
40 oz.	4.35	1.74
64 oz.	6.85	1.71
Brand “E” Sugar		
1 lb.	0.63	0.63 (per pound)
2 lb.	1.01	0.51
5 lb.	1.99	0.40
10 lb.	3.95	0.39
Brand “F” Large Eggs		
6	0.79	0.131 (per egg)
12	1.29	0.107
18	1.95	0.108
Brand “G” 2% Milk		
1 qt.	0.95	0.95 (per quart)
1/2 gal.	1.69	0.84
1 gal.	2.69	0.67
Brand “H” Catsup		
14 oz.	0.91	2.08 (per quart)
20 oz.	1.12	1.79
28 oz.	1.95	2.23
40 oz.	2.59	2.07
60 oz.	3.59	1.80

donic coefficients can be unstable across time and across geography. Another is that the hedonic variables can be intercorrelated, leading to all the usual problems of out-of-sample forecasts based on least squares. Finally, reliable calculation of the hedonic factors requires a large number of observations—often larger than the number of CPI quotes available to estimate them.

Scanning or EPOS data offer a way around many of these obstacles, especially for certain Durable goods items where product turnover is rapid, the number of relevant characteristics is large, and the CPI sample size is small. For this reason, hedonic adjustment research using EPOS data in the United States, Canada, and the United Kingdom has tended to focus on durable goods, as evidenced by a number of papers on this topic presented at this and other recent conferences (see, e.g., Lowe 1998; Ioannides and Silver 1997).

Hedonic adjustment procedures, however, can also be used for fast-moving consumer goods, such as packaged foods and household products. For example, a few years ago the packaged laundry detergent industry switched to a new “high-density” formulation that made direct comparison of new and old product prices impossible without taking this product reformulation into account. A similar discontinuity arose with the introduction of “high yield” extraction coffee roasting techniques fifteen or twenty years ago. Other examples could easily be cited.

The richness of the scanner-based data sets, not only in number of observations but also in the systematic collection of detailed product characteristics, makes hedonic adjustment for new items in these fast-moving consumer goods much easier to carry out. Beyond that, however, it can also greatly facilitate the process of disaggregation in the current CPI system, in which individual items are selected for pricing in the store based on a hierarchical list of price-determining characteristics. This process of identifying and ranking the relevant price-determining characteristics for each cluster, ELI, or item stratum can be greatly refined with the use of scanning data.

1.3.4 Characteristic-Based Subgroups

Finally, a robust alternative to making explicit hedonic adjustments in certain cases involves using the hedonic coefficients to identify price-determining characteristics and then combining items into subgroups based on these characteristics, with a unit value price computed across the items in each subgroup. Such a procedure would make it possible to compare prices across cities, regions, and even countries where the set of brands handled differs from place to place. It would also provide a mechanism for handling new items (e.g., flavors) within existing brands or even the introduction of an existing brand into a new region of the United States.

The number of detailed product characteristics available for such analyses in the ACNielsen ScanTrack product reference file is substantial. Currently, ACNielsen has over 3 million unique UPCs coded in its item dictionary, with over 3,000 new entries added on a weekly basis. The product reference file contains nearly 600 unique characteristics. Each specific product category has a set of mandatory characteristics that are coded with product in hand by ACNielsen associates. Many of these characteristics could be considered price-determining for their category. Appendix C pro-

vides a listing of characteristics associated with the ice cream category. Values for these characteristics are only populated if present on product packaging.

1.4 An Extended Example Using One Item Stratum

One of the thirty-four “scannable edible” item strata for which separate price indexes are computed in the CPI is “ice cream and related products.” This item stratum corresponds to a total of four ACNielsen ScanTrack “product modules”: bulk ice cream, ice milk and sherbet, frozen yogurt, and frozen novelties. Each module can be approximated by a combination of characteristics used in the CPI disaggregation procedure for the “ice cream and related products” item stratum. One of the modules—bulk ice cream—is identical to a commodity that is included in the CPI “average price” series.

A profile of this product category is shown on tables 1.6 and 1.7. Note

Table 1.6 How Many Different Kinds of Ice Cream and Related Products Are There, and How Many of Them Are New?

ScanTrack Module	Number of Different Items in 1999			% of 1999 \$ Sales from New Items
	Total	Old	New	
Bulk ice cream	8,056	7,422	634	4.2
Ice milk and sherbet	788	756	32	4.1
Frozen yogurt	900	865	35	3.0
Frozen novelties	3,106	2,785	321	4.9
Total	12,850	11,828	1,022	4.4

Note: An item is defined as a particular flavor of a particular size of a particular type of a particular brand. This generally, but not always, corresponds to one UPC. New items are items sold in 1999 but not in 1998.

Table 1.7 Profile of “Ice Cream and Related Products” Item Stratum Based on ScanTrack Data for 1999

Type	1999 Sales (\$Millions)	% of Item Stratum 1999 \$ Sales	% of Sales on Promotion	
			Dollars	Quantities
Bulk ice cream	3,895	64.0	11.8	15.0
Ice milk and sherbet	133	2.2	9.1	12.6
Frozen yogurt	264	4.4	12.6	19.2
Frozen novelties	1,788	29.4	22.7	30.3
Total	6,082	100.0	15.0	n.a.

Note: n.a. = not available.

that there were nearly 13,000 separate items of “ice cream and related products” sold somewhere in the United States at some time during 1999. More than 1,000 of these items were new (i.e., not found in 1998), and these new items accounted for more than 4 percent of U.S. ice cream expenditures in 1999 (this could imply an annual rate of 8 percent at year end). Fifteen percent of total expenditures on ice cream and related products were for items that were on promotion during the week in which they were purchased (“on promotion” means that they were on display, were advertised by the retailer, or had a temporary price reduction). The typical U.S. supermarket, incidentally, stocks more than 500 of these items.

Because of the close correspondence between the ACNielsen and CPI category definitions, it is possible to compare CPI with ScanTrack “price index” data for the item stratum as a whole, and CPI with ScanTrack “average price” data for the “bulk ice cream” module. The results of these comparisons are set forth on tables 1.8 and 1.9.

In table 1.8, note that the two “average price” series are in very close agreement. Note also, however, that the ScanTrack “average price” in 1999

Table 1.8 Average U.S. Price of Bulk Ice Cream per 64 oz. Equivalent Unit

Source	1998	1999	Average Price Ratio \times 100
CPI ^a	\$3.127	\$3.296	105.42
ScanTrack ^b	\$3.171	\$3.360	105.96
ScanTrack vs. CPI	+1.4%	+1.9%	
ScanTrack ^b less 2 new brands	\$3.171	\$3.330	105.01
ScanTrack less 2 new brands vs. CPI	+1.4%	+1.0%	

^aBased on linear average of monthly average price series.

^bAnnual dollar sales divided by annual equivalent unit sales.

Table 1.9 1998–99 U.S. Ice Cream Price Trends Reported by ScanTrack and CPI

Source	Index Type	1999 Index (1998 = 100) by Module				Total
		Bulk Ice Cream	Ice Milk and Sherbet	Frozen Yogurt	Frozen Novelties	
CPI	Geomean	n.a.	n.a.	n.a.	n.a.	<i>1.0400</i>
ScanTrack	Geomean	1.050	1.043	1.036	1.028	<i>1.0422</i>
ScanTrack	Törnqvist	1.049	1.046	1.035	1.027	1.0420
ScanTrack	Laspeyres	1.051	1.044	1.037	1.029	1.0433
ScanTrack	Fisher	1.049	1.046	1.036	1.028	1.0420
ScanTrack	Unit value ^a	<i>1.060</i>	1.027	1.023	1.029	n.a.
CPI	Unit value ^b	<i>1.054</i>				

Notes: n.a. = not available. Figures in italics indicate agreement between CPI and ScanTrack values.

^aPrice based on annual dollar sales divided by annual equivalent sales.

^bPrices based on linear average of monthly “average price” series.

was increased, by 1 full percent, by the appearance in the market of two new “luxury” brands that could not have been included in the CPI “average price” series in 1999 because of the delay from the time that a new item comes into the marketplace to the time that it has a chance of selection in the CPI. To some extent, this calls into question the belief that the “average price” series correctly reflects only the downward price pressures that Reinsdorf (1993) characterized as resulting from “Schumpeterian creative destruction.” We now see that new items can raise average price levels as well as lower them. Other recent examples of this phenomenon are “boutique brewery” brands of beer and “gourmet” brands of coffee. In theory, the right hedonics should be able to identify and correct for these instances, but it would not be an easy job.

In table 1.9, the relationship between CPI and ScanTrack year-to-year price index trends is surprisingly close, although the high degree of correspondence among Laspeyres, Fisher, Geomean, and Törnqvist indexes suggests that inter-item price elasticities in this category may indeed be close to unity. For ScanTrack Geomean and Laspeyres indexes, 1998 was used as the base year, in contrast to the implicit weights generated by the CPI outlet and item selection process, which probably reflect an average product mix dating back to around 1996.

It is obvious that dividing “ice cream and related items” into four separate modules would result in a moderate reduction in variance in measuring year-to-year change at the item stratum level, since the subgroups differ among themselves in their year-to-year change.

Finally, it is interesting to note that ice cream is a category with moderately strong seasonal variation. The weekly seasonal expenditure peak (4 July) is roughly twice as high as the seasonal trough (New Year’s week).

1.5 A Coda (or Cadenza) on Aggregation

Henri Theil gives a useful characterization of data aggregation issues in his *Linear Aggregation of Economic Relations* (Theil 1955) and elaborates on it in his *Economics and Information Theory* (Theil 1967). In the earlier book, Theil identifies three dimensions of aggregation: aggregation over individuals, aggregation over commodities, and aggregation over time. In the later book Theil goes on to discuss application of his ideas to price index theory.

For convenience, I have shown in table 1.10 five potential levels of aggregation for each dimension with scanner data. Current CPI practice is generally to choose level 1 for time, level 2 for space, and level 3 for entity. For our simulated CPI using scanning data, we have aggregated time and space to level 5 but used level 2 for entity.

Using somewhat different terminology, Parsons and Schultz (1976) apply these notions of aggregation to the field of Marketing Research in their

Table 1.10 Potential Aggregation Levels for Price Index Construction Using Scanner Data

Aggregation Level	Dimension		
	Time	Space	Entity
5 (highest)	Annual	U.S.	Category
4	Monthly	“City”	Segment
3	Weekly	Organization within city	Type by size within brand
2	Price point within week	Outlet	Item
1 (lowest)	Daily ^a	Household	UPC

^aIn Japan, for example, where many price promotions are in effect for one, two, or three days, scanning data are supplied by retailers on a daily basis.

book *Marketing Models and Economic Research*. The relevant chapter in their book was actually written by Dick Wittink, who has subsequently written several papers on the subject using scanning data. In these articles Wittink (see, e.g., Foekens, Leeflang, and Wittnik 1997) finds that it is difficult (and in certain circumstances impossible) to estimate price elasticities reliably using data that have been aggregated across outlets, especially when different prices have been collapsed in the aggregation.

Is this also true for price indexes? Consider the following two questions: (a) Do eggs cost more in stores east of the Mississippi this year than they did last year? (b) Do eggs cost more in stores east of the Mississippi this year than they do in stores west of the Mississippi?

Are these equivalent questions? More precisely, should we use equivalent methods in comparing prices across time as we do in comparing prices across geography?

According to Peter Hill, the answer might well be “yes.” In chapter 16 of *Systems of National Accounts*, Hill (1993) writes that “A price index is an average of the proportionate changes in the prices of a specified set of goods and services between two periods in time.”

Hill goes on to state that “It is possible to compare prices and volumes between countries using the same general methodology as for intertemporal comparisons within a single country.”

Thus, Hill’s two statements might be combined as follows: “A price index is the average of the proportional differences in the prices of a specified set of goods and services between two price regimes,” where “regime” is used to denote a particular time and place.

In trying to apply this principle to our egg price question, however, we run into an immediate problem: in comparing prices across time, we normally make use of outlet-level data, lining up or “lacing together” the individual outlet prices across time and then combining the resulting price ratios using a Laspeyres, Geomean, Törnqvist, or some other aggregation formula.

When we try to compare egg prices across geography, however, we find

that lacing across outlets won't work, because the eyelets on one side of the shoe (or outlets on one side of the river) don't match up with those on the other side. Thus, in making interspatial comparisons, we have no choice but to aggregate outlets all the way up to the regional (or, in the case of purchasing power parities, national) level. We have no hesitation about doing this for interspatial comparisons, but we are reluctant to do so for intertemporal ones. Why is this?

One of the attractive properties of the Geomean is that, with it, the operational problem goes away, since the product of ratios is the same as the ratio of products so long as the weights stay constant. However, the conceptual problem stays. At the very least, obtaining the price ratios from an identical set of outlets has substantial variance implications because of the correlation between item prices across time at the individual outlet level. Aggregated data can also take into account new outlets and items. The question is whether they should do so.

There has been surprisingly little discussion of this issue in the literature, although a few authors have come close. Dalen (1992) has pointed out the difference in concept between the two methods as they apply in a time-series context, but he avoids taking sides. Empirically, Reinsdorf (1998), examining a scanner data set for coffee in two geographic markets, finds little difference between time-series indexes resulting from the use of data aggregated or disaggregated across outlets, but he expresses no preference between them. Kokoski, Moulton, and Zieschang (1996) have written a lengthy article describing a method for comparing prices across cities using hedonic procedures, in which they mention that the procedure can be adapted to time-series comparisons and for achieving transitivity between temporal and spatial measures, but they do not consider the effect of different across-outlet aggregation methods in achieving this transitivity.

Diewert (1995a) suggests that "if individual outlet data or transactions were not available or were considered to be too detailed, then unit values for a homogenous commodity over all outlets might form the lowest level of aggregation." Saglio (1994) starts with data aggregated across all outlets in France by six different shop types and never even considers whether the data should be kept disaggregated by outlet. Magnien and Pougard (1999) use a similar procedure on French scanner data. Dalen (1997), presenting the results of Swedish scanner data both ways (aggregated vs. disaggregated across outlets), merely argues "for the use of unit values at least over time and *perhaps* also over outlets in a market area" (emphasis added). De Haan and Opperdoes (1997), presenting scanner data for the Netherlands, look at the data both ways but do not reach a conclusion, either. Pollock (1995) briefly considers the issue but simply concludes that "there are opportunities here for empirical work." No one seems to have devised a general rule or even considered the difference between time-series and geographic approaches to aggregation across outlets.

In the price indexes that we have constructed for the ice cream and related products item stratum, we have aggregated the outlet data into a single unit value all the way up to the national level but have kept the item detail fully disaggregated for all 12,000 items common to both years. We have done this largely as a matter of computational ease rather than methodological preference. It seems to us unlikely that we would get a substantially different answer if we had kept the item detail disaggregated by outlet, but we cannot guarantee it.

This is, in our opinion, a methodological issue worthy of further theoretical and empirical consideration. Obviously the concept of interspatial price comparisons has little meaning when viewed from the perspective of the consumer, except for the rare consumer who may be trying to decide on which side of the Mississippi he should buy his eggs.

There are related issues in regard to aggregation across time, especially for products with strong seasonal variations in sales. These have been discussed by Diewert (1996, 2000b) and by Turvey (1998). Scanner data provide an opportunity to examine these issues as well. The effect of various temporal aggregation procedures is shown in table 1.11. Echoing Turvey's

Table 1.11 1988–99 Annual Comparison of Price Index Trends for “Ice Cream and Related Products” Using Various Methods for Aggregating Monthly Price Indexes

Month	CPI (1982-84 = 100)			ACNielsen ScanTrack Share of Annual			
	Year		1999 ÷ 1998	Expenditures		Quantities	
	1998	1999		1998	1999	1998	1999
January	153.5	165.2	1.0762	0.0617	0.0623	0.0632	0.0606
February	150.8	163.5	1.0842	0.0681	0.0704	0.0708	0.0704
March	150.0	160.2	1.0680	0.0712	0.0728	0.0747	0.0741
April	152.4	162.4	1.0656	0.0773	0.0782	0.0787	0.0787
May	150.9	160.0	1.0603	0.0959	0.0939	0.0997	0.0966
June	153.2	161.7	1.0555	0.1078	0.1133	0.1108	0.1127
July	153.1	158.8	1.0372	0.1186	0.1180	0.1179	0.1198
August	155.2	159.2	1.0258	0.1041	0.1011	0.1031	0.1022
September	157.9	159.8	1.0120	0.0896	0.0863	0.0876	0.0865
October	162.3	163.8	1.0092	0.0762	0.0767	0.0726	0.0745
November	163.9	162.0	0.9984	0.0682	0.0671	0.0630	0.0661
December	162.8	164.1	1.0080	0.0614	0.0600	0.0579	0.0579
Annual	155.5	161.725	n.a.	1.0000	1.0000	1.0000	1.0000
1999 ÷ 1998							
“Official” CPI (Dutot aggregation)			1.0400				
Unweighted CPI (Jevons aggregation)			1.0413				
Unweighted CPI (Carli aggregation)			1.0417				
Weighted CPI (Törnqvist-Theil aggregation)			1.0412				
Weighted CPI (unit value aggregation)			1.0410				

Note: n.a. = not applicable.

conclusion on a different set of data, we find that the differences “are very small, but they do exist.”

Appendix A

Formulas Used for Table 1.9

CPI Geomean. See description in “The Experimental CPI using Geometric Means” (U.S. Bureau of Labor Statistics Research Paper, 1997).

$$\text{ScanTrack Geomean} = \prod_i \left(\frac{p_i^{99}}{p_i^{98}} \right)^{S_i^{98}}$$

Where p_i is an annual total U.S. unit value price for the i th item with sales in both 1998 and 1999.

$$S_i^{98} = \frac{E_i^{98}}{\sum E_i^{98}}, \text{ and } E = p_i q_i = \text{price} \times \text{quantity for the item}$$

$$\text{ScanTrack Törnqvist} = \prod_i \left(\frac{p_i^{99}}{p_i^{98}} \right)^{(S_i^{98} + S_i^{99})/2}$$

Where S_i^{99} is defined as for S_i^{98} , above.

$$\text{ScanTrack Laspeyres} = \frac{\sum_i p_i^{99} q_i^{98}}{\sum_i p_i^{98} q_i^{98}}$$

$$\text{ScanTrack Fisher} = \left(\frac{\sum_i p_i^{99} q_i^{98}}{\sum_i p_i^{98} q_i^{98}} \times \frac{\sum_i p_i^{99} q_i^{99}}{\sum_i p_i^{98} q_i^{99}} \right)^{1/2}$$

Appendix B

Formulas Used for Table 1.11

Here I have borrowed the notation of Diewert (2000b).

$$\text{Carli: } P_C(p_n^0, p_n^t) \equiv \sum_{m=1}^{12} \left(\frac{1}{12} \right) \left(\frac{p_n^{t,m}}{p_n^{0,m}} \right); \quad n = 1, \dots, N; \quad t = 1, \dots, T$$

where $p_n^{t,m}$ denotes the price for commodity n in month m in year t .

$$\text{Jevons: } P_j(p_n^0, p_n^t) \equiv \sum_{m=1}^{12} \left(\frac{p_n^{t,m}}{p_n^{0,m}} \right)^{1/12}; \quad n = 1, \dots, N; \quad t = 1, \dots, T$$

$$\text{Dutot: } P_D(p_n^0, p_n^t) \equiv \left[\sum_{m=1}^{12} \left(\frac{1}{12} \right) p_n^{t,m} \right] / \left[\sum_{m=1}^{12} \left(\frac{1}{12} \right) p_n^{0,m} \right];$$

$$n = 1, \dots, N; \quad t = 1, \dots, T,$$

where the year-to-price vector for commodity n is

$$\mathbf{p}_n^t \equiv (\mathbf{p}_n^{t,1}, \mathbf{p}_n^{t,2}, \dots, \mathbf{p}_n^{t,12}); \quad n = 1, \dots, N; \quad t = 0, 1, \dots, T.$$

$$\text{Törnqvist-Theil: } P_T(p_n^0, p_n^t, s_n^0, s_n^t) \equiv \sum_{m=1}^{12} \left(\frac{1}{2} \right) [s_n^{0,m} + s_n^{t,m}] \ln \left(\frac{p_n^{t,m}}{p_n^{0,m}} \right);$$

$$n = 1, \dots, N; \quad t = 1, \dots, T,$$

where the year t vector of monthly expenditure shares for commodity n is

$$\mathbf{s}_n^t \equiv (\mathbf{s}_n^{t,1}, \mathbf{s}_n^{t,2}, \dots, \mathbf{s}_n^{t,12}); \quad n = 1, \dots, N; \quad t = 0, 1, \dots, T.$$

$$\text{Unit value: } \frac{P_n^t}{P_n^0} \equiv \frac{\sum_{m=1}^{12} \sigma_n^{t,m} p_n^{t,m}}{\sum_{m=1}^{12} \sigma_n^{0,m} p_n^{0,m}}; \quad n = 1, \dots, N; \quad t = 0, 1, \dots, T.$$

where $\sigma_n^{t,m}$ is defined as

$$\sigma_n^{t,m} \equiv \frac{q_n^{t,m}}{Q_n^t}; \quad n = 1, \dots, N; \quad t = 0, 1, \dots, T.$$

Appendix C

Ice Cream Category Item Characteristics (from ACNielsen product reference)

Characteristic Typename

Artificial color flavor presence claim

Bonus pack

Branded component or flavor

Calcium presence claim

Calorie Claim

Calories per serving size

Cholesterol

Cholesterol presence claim

Claim

Commodity group
Common consumer name
Dietary fiber
Endorsement
Enrobing flavor
Fat calorie per serving size
Fat presence claim
Fat substitute type
Flavor
Form
Imported or domestic
Lactose presence claim
Licensed trademark
Manufacturer suggested price claim
Manufacturing process
Milk fat
Monounsaturated fat gram
Naked product source
Natural or artificial ingredient claim
Origin
Product storage as stated
Product weight
Protein gram
Saturated fat gram
Season
Serving per container
Serving size household
Serving size metric
Sodium
Sodium presence claim
Sorbitol
Strategic ingredient presence claim
Sugar alcohol
Sugar gram
Sweetener presence claim
Sweetener type
Target group condition
Total fat gram
Vitamin presence claim
Package general shape
Polyunsaturated fat gram
Potassium
Preparation method
Prepriced

Preservative presence claim
Product claim
Product count
Product size

References

- Balk, B. 1998. On the use of unit value indices as consumer price subindices. *Proceedings of the fourth meeting of the International Working Group on Price Indices*, 112–120. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics.
- Court, A. J. 1939. Hedonic price indexes with automotive examples. In *The Dynamics of Automobile Demand*, 95–119. New York: General Motors.
- Dalen, J. 1992. Computing elementary aggregates in the Swedish consumer price index. *Journal of Official Statistics* 8:129–47.
- . 1997. Experiments with Swedish scanner data. In *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. Bert Balk, 163–68. Research Paper no. 9806. Voorburg, the Netherlands: Statistics Netherlands, Division of Research and Development, Department of Statistical Methods.
- Diewert, W. E. 1995a. Axiomatic and Economic approaches to international comparisons. Discussion Paper no. 95-01. Vancouver, Canada: University of British Columbia, Department of Economics.
- . 1995b. Price and volume measures in the system of national accounts. Discussion Paper no. 95-02. Vancouver, Canada: University of British Columbia, Department of Economics.
- . 1996. Seasonal commodities, high inflation, and index number theory. Discussion Paper no. 96-06. Vancouver, Canada: University of British Columbia, Department of Economics.
- . 2000a. The Consumer Price Index and index number purpose. Discussion Paper no. 00-02. Vancouver, Canada: University of British Columbia, Department of Economics.
- . 2000b. Notes on producing an annual superlative index using monthly price data. Discussion Paper no. 00-08. Vancouver, Canada: University of British Columbia, Department of Economics.
- Foekens, E. W., P. Leeflang, and D. Wittnik. 1997. Hierarchical versus other market share models for markets with many items. *International Journal of Research in Marketing* 14:359–78.
- Frisch, R. 1930. The problem of index numbers. *Journal of the American Statistical Association* 25 (December): 397–406.
- Greenlees, J. S. 2000. Consumer price indexes: Methods for quality and variety change. *Statistical Journal of the United Nations* 17:59–74.
- Haan, J. de, and E. Opperdoes. 1997. Estimation of the coffee price index using scanner data: The choice of the micro index. In *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. Bert Balk, 191–202. Research Paper no. 9806. Voorburg, the Netherlands: Statistics Netherlands, Division of Research and Development, Department of Statistical Methods.
- Hill, P. 1993. Price and volume measures. In *System of National Accounts 1993*, Organization for Economic Cooperation and Development staff, 379–406. N.p.: United Nations Publications.

- Ioannides, C. and M. Silver. 1997. Chained, exact, and superlative hedonic price changes: Estimates from micro data. In *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. Balk, 215–224. Research Paper no. 9806. Voorburg, the Netherlands: Statistics Netherlands, Division of Research and Development, Department of Statistical Methods.
- Kokoski, M. F., B. Moulton, and K. Zieschang. 1996. Interarea price comparisons for heterogeneous goods and several levels of commodity aggregation. Working Paper no. 291. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics.
- Lowe, R. 1998. Televisions: Quality changes and scanner data. In *Proceedings of the fourth meeting of the International Working Group on Price Indices*, 5–20. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics.
- Magnien, F., and J. Pougard. 1999. Non-parametric approach to the cost-of-living index. In *Proceedings of the Measurement of Inflation Conference*, ed. Mick Silver and David Fenwick, 345–76. Cardiff, Wales: Cardiff University.
- Moulton, B. 1993. Basic components of the CPI: Estimation of price changes. *Monthly Labor Review* (December): 13–24. Washington, D.C.: Bureau of Labor Statistics.
- Olt, B. 1996. *Axiom und Struktur in der Statistischen Preisindextheorie* (Axiom and structure in statistical price index theory). Frankfurt, Germany: Peter Lang.
- Parsons, L., and R. Schultz. 1976. *Marketing models and econometric research*. Amsterdam: North-Holland Publishing Company.
- Pollock, R. A. 1995. Elementary aggregates in the CPI. Paper prepared for U.S. Bureau of Labor Statistics. University of Washington, Department of Economics.
- Reinsdorf, M. 1993. The effect of outlet price differentials on the U.S. Consumer Price Index. In *Price Measurements and their Uses*, ed. M. F. Foss, M. E. Manser, and A. H. Young, 227–54. Chicago: University of Chicago Press.
- . 1998. Constructing basic components for the U.S. CPI from scanner data: A test using data on coffee. Working Paper no. 277. Washington, D.C.: U.S. Bureau of Labor Statistics.
- Saglio, A. 1994. Comparative changes in average price and a price index: Two case studies. Paper presented at the International Conference on Price Indices. October 31–November 2, Ottawa, Canada.
- Silver, M. 1995. Elementary aggregates, micro-indices, and scanner data: Some issues in the compilation of consumer price indices. *Review of Income and Wealth* 41 (4): 427–35. Cardiff, Wales: Cardiff University, Cardiff Business School.
- Tellis, G. 1988. The price elasticity of selective demand: A meta-analysis of econometric models to sales. *Journal of Marketing Research* 25:331–41.
- Theil, H. 1955. *Economics and information theory*. Amsterdam: North-Holland Publishing Company.
- . 1967. *Linear aggregation of economic relations*. Amsterdam: North-Holland Publishing Company.
- Triplett, J. 1997. Measuring consumption: The post-1973 slowdown and the research issues. *St. Louis Federal Reserve Bank Review* 79 (3): 9–46.
- . 1998. Elementary indexes for a consumer price index. In *Proceedings of the fourth meeting of the International Working Group on Price Indices*, 176–97. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics.
- Turvey, R. 1998. Months versus years. In *Proceedings of the fourth meeting of the International Working Group on Price Indices*, 198–202. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics.
- Wynne, M., and F. Sigalla. 1994. The Consumer Price Index. *Economic Review* (second quarter): 1–22.