Catalogue no. 62F0014MIB

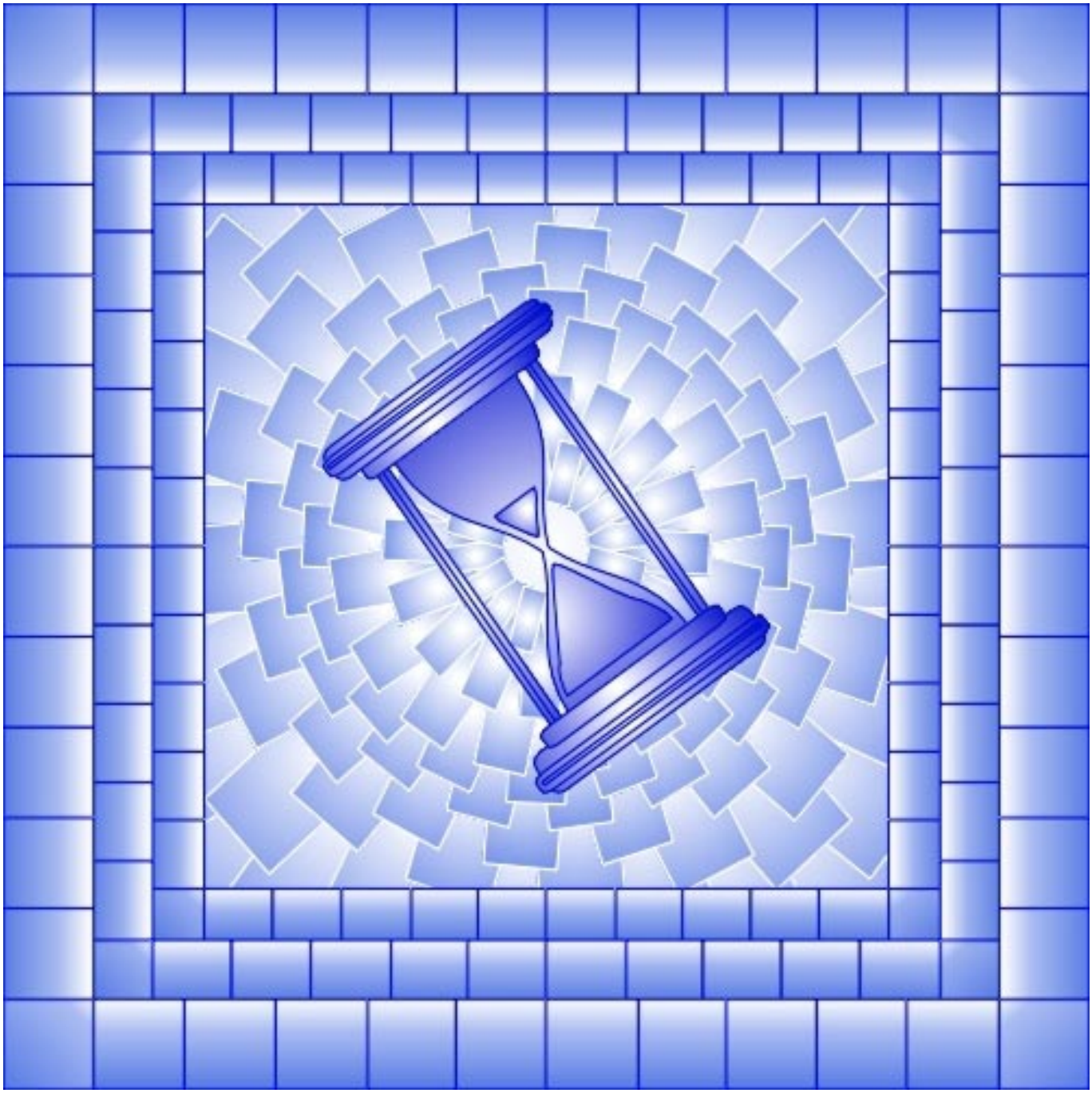
Analytical Series

Prices Division

Televisions: Quality Changes and Scanner Data

By Robin Lowe

No.14



Statistics Canada Statistique Canada

Canada

Televisions: Quality Changes and Scanner Data

By Robin Lowe
Prices Division, Statistics Canada

Catalogue No. 62F0014MIB, Series No. 14
ISBN: 0-660-60629-1
ISSN: 1206-2677

Jean Talon Building, 13-D7, Ottawa, K1A OT6

Facsimile: (613) 951-2848

Telephone: (613) 951-9495

Email: ***lowerob@statcan.ca***

This paper is available at
www.statcan.ca/english/IPS/Data/62F0014MIB.htm

April 2001

Texte français au verso


ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca

Preface

Prices Division first started publishing the Analytical Series in December 1996 as a means to convey conceptual and applied research undertaken by its staff, and at times, by other persons from within or outside Statistics Canada on the subject of price indexes.

All papers are reviewed by a panel of experts from within Statistics Canada or outside the agency. Views expressed in the papers are those of the authors and do not necessarily reflect those of Prices Division or Statistics Canada.

The purpose of the series is to disseminate knowledge and stimulate discussion. Questions and comments on any aspect of the papers are welcome and can be forwarded to Louis Marc Ducharme, Director (Internet email: ducharl@statcan.ca; Telephone: 613-951-0688) or to Robin Lowe, Chief, Quality Assurance (Internet email: lowerob@statcan.ca; Telephone: 613-951-9495), Prices Division, Statistics Canada, Ottawa, Ontario, K1A 0T6.


ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca

Abstract

This paper is the first in a series of reports examining the possible use of scanner data for constructing price indexes. This case study focuses on televisions and compares their price behaviour taken from current surveying methods with alternative measures obtained from massaging electronic data records on all sales by a retailer over a comparable period. Examination of the price index history for televisions shows that the recognition and adjustment for quality change in the sample and the impact of shifts in purchasing patterns have similar impact on the index numbers. The advantages of scanner data—that they record actual sales and current purchasing patterns—have to be set against the difficulty of recognising quality change. This analysis shows that while there are substantial gains from using scanner data in monitoring and adjusting for purchasing pattern changes, it is difficult to account for quality changes without micro-editing the data. The scanner data set raises statistical issues, largely questions of what aggregation across time, outlets and products should be done, that have to be answered before using it in index estimation. Future analysis will be aimed at resolving these issues.

Keywords: Price indexes, quality change, sampling, scanner data.


ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca

Table of Contents


ELECTRONIC PUBLICATIONS AVAILABLE AT
www.statcan.ca

1.0 Introduction[1]

The purpose of this paper[2] is to examine some of the problems in the construction of price indexes for televisions using our current procedures and electronic data records[3] that have recently become available. Recent research[4] has documented some of the costs and benefits in the use of scanner data for food items. However, because of product diversity and higher product turnover in the marketplace, these costs and benefits may be quite different for consumer durables such as televisions (Silver et al., 1995, 1997).

The current practice at Statistics Canada consists of collecting prices for a small number of representative items that are monitored over time for changes in price and in quality. When a quality change occurs, it usually signals an item replacement that requires some kind of adjustment. This is where one of the weaknesses of scanner data lies, as the larger number of records may preclude us from giving the same level of care to the quality adjustment process. The critical issue then becomes one of weighing the advantages of additional observations provided by the scanner data against the disadvantages of paying less attention to changes in the quality of the sampled items.

What follows is an analysis of the price indexes that were calculated for televisions using two different approaches—the current practice at Statistics Canada and the use of scanner data. The first part of the paper describes the sample and the quality change evaluation process for televisions over the last seven years. Some assessment of the impact of quality change evaluation is given. The second part of the paper examines the results of scanner data for the period February 1997 to February 1998. Some comparisons are made with the results for the overlapping period from the existing survey, and the reasons for the differences are analysed. A number of issues and questions raised in the paper are discussed in the conclusion. The way in which quality change is handled is important both in the current practice and in treating scanner data. There are other issues concerning aggregation and sample selection that affect the results from the scanner data. However, for those cases where a direct comparison can be made between the prices collected in the current practice and those obtained from the scanner records, the difference between these average movements is greater than the range of results derived from various treatments of the scanner data.

---

[1] *Thanks are due to Candace Ruscher for performing much of the research, and to John Mallon and Hugh Scobie for programming and carrying out the calculations under different scenarios.*

[2] *Prepared for the Ottawa Group meeting in Washington, D.C., April 22-24, 1998.*

[3] *For simplicity, electronic data records will be called scanner data in the rest of the paper.*

[4] *See Reinsdorf (1995), Dalén (1997), Hawkes (1997), de Haan and Opperdoes (1997), and Scobie (1997).*

2.0 CPI index construction for televisions

Televisions exhibit relatively uncommon price behaviour, as their prices have fallen steadily, both in nominal terms and as measured in a price index. The available data set runs from June 1990 through November 1997. There are two representative products: a 20" and a 25" to 27" colour TV. Deviations from the preferred specification are allowed, for example, 19" screens are permitted in the first representative product, and 25" to 28" screens are accepted in the second. Since the most common screen sizes are 20" and 27", they will be featured in the rest of the paper.

Over the study period, the total Canada sample averaged about 220 observations until early 1995, when it dropped to about 140. Altogether there were 337 streams[5] of monthly observations, 89 of which continued throughout the whole period. The others entered and/or exited the sample at various times. The 89 streams that were in the sample for the whole period had 579 quality changes, averaging about one a year, though the number of changes ranged from 2 to 12. The remaining 248 observation streams appear to have had quality changes at about the same rate. There were 1213 quality changes in all.

Most changes in the sample were involuntary. Apart from the sample reduction in 1995, due to budget cuts, deletions from the sample and their subsequent replacement were usually made necessary because the retail outlet became unavailable. Occasionally, new outlets were introduced to replace others that were dropped from the sample. Quality change adjustments were also forced, either because the existing item became unavailable or because the price collector was advised that it was about to become unavailable. There were no changes to the item selection initiated by the statisticians during the study period.

Additions to the sample are routinely introduced into the month-to-month matched sample without affecting the index level. However, the index measurement can be affected by the judgement made at the time of a perceived quality change. In total, 1213 quality changes were recorded between June 1990 and November 1997. Of these, 664 (54.7%) were splices.[6] This includes 229 cases where neither the price nor the quality, nor therefore, the index changed. Some of these 229 may have been entries to correct descriptions on the specification, or something equally inconsequential, but it is not possible now to identify them.

[5] *An observation stream is one series of observations over time, adjusted for quality change, if necessary. It begins with an addition to the sample and ends with a deletion from the sample.*

[6] *The word splice is used in this paper to describe the situation where the ratio of the price of the replacement item in the current month to the price of the replaced item in the preceding month is taken to represent their relative qualities. This is different from a true splice, which is based on a comparison of prices at the same time. The additional implicit assumption is that prices have not changed between the two months, which may not be correct. Since televisions are not priced every month, the previous month's price may have been imputed from an earlier period, which makes the assumption more dubious.*

There are several patterns of quality change adjustment. We distinguish between small and large changes—that is, between cases where the price of the replacement item is within 10% of the replaced item, and where it is not. The choice of 10% is arbitrary. It is chosen to separate what may be just small modifications to essentially the same product from cases where one completely different product replaces another. We also distinguish splices from other types of judgement.[7] Within judgements, we differentiate between cases where the quality change is assessed at zero, and where it is not. From 1993 onwards, the archived records contain a fuller description of the items being priced—including, routinely, the manufacturer. So, we can make the distinction between cases where the new item replaces an existing one from the same manufacturer, and where it replaces an item from a different manufacturer (brand change). The results of this analysis are shown in the appendix tables, while the summaries are given below.

2.1 Summary of appendix tables

Of the 1213 quality changes, 732 are small with less than a 10% change in price. Of the 481 larger changes, 233 involved replacing the existing item with a cheaper one, and 248 with a more expensive one.

Except for those 229 cases where nothing changed, splicing is much more prevalent when the price change is large than when it is small. Splicing occurred in only about 37.5% of cases where there was a small price change, as opposed to about 55% for large price changes.

Where judgements are used, there is the same association between the price difference and the resulting index difference as has been found generally (Lowe, 1997). That is, particularly for price changes over 10%, the larger the difference in price between the replacement item and the replaced item, the larger the impact will be on the index. This applies downward as well as upward—cases where the price drops 20% show the index falling more than cases where the price drops 10%. The number of large price increases is slightly higher than the number of large price decreases.

It is noticeable that in a large number of the judgement changes (245 of 549), the quality was not deemed to have changed at all. These cases are concentrated overwhelmingly among the small changes where they account for 71% of both upward and downward adjustments.

There were 750 quality changes since the beginning of 1993. Of these, in 552 cases the change involved replacing an item with another from the same manufacturer, and in 198 there was a change of manufacturer. There were striking differences between them. Splicing was much more common (68%) when there was a change in the manufacturer, than when there was not (50%). To some extent, this difference was due to the higher proportion of large price changes when the manufacturer changed (59%), than when it did not (25%). As seen earlier, splicing is

[7] *This does not imply that splices are not judgements. The label "judgement" may be read as shorthand for judging that, in a given case, market prices do not reflect quality differences.*

more common with large price changes. Splicing, however, was still more common for changed manufacturers right across the board. Similarly, quality is less likely to be judged unchanged when the manufacturer has switched.

2.2 Treating quality change in the field

Patterns of treating quality change coincide with the experience of price evaluators. In most cases they have brochures from the manufacturers to help them assess changes in the specifications of models. The large number of splices with no price change is due to a common practice of manufacturers changing a model number with a new production run, though nothing physically has changed. Also, there are many cases where the quality did not change, but the price and index went up or down slightly. These cases should not really be regarded as quality changes. Many times, a small change in a specification accompanies a model change—in the type of remote control supplied, in the warranty coverage, or in the number or placement of jacks. In these cases, the small adjustment in price can be easily calculated. The price evaluator is not really making a change to the sample, only recording a slight modification to essentially the same model.

In other cases when one item is replaced by another, the price collector is asked to find a similar item if possible, but one that is a volume seller. The volume seller requirement sometimes results in a model that is quite different from the previous one. Price evaluators have adjustment guidelines for different model sizes, from mono sound to stereo sound, and for other common improvements, so they can make reasonable comparisons in many cases. However, the comparison is more difficult when it involves a change of manufacturer. While it is recognised that manufacturers do vary in quality, it is difficult to compare and estimate by how much. That may be why more quality changes are spliced when it includes a change in manufacturer. This is a phenomenon particular to certain retail outlets. Most retailers carry certain manufacturers and change rarely. Some, however, switch manufacturers frequently, making the best buys they can each time.

Some changes that appear to be judgement changes disguise a different assessment. If the replaced item was discounted in previous months, the quality and price of the replacement item are compared to the replaced item at its normal market price, not to the discounted one. If the desired result from that comparison is a splice, the new reference price is calculated to restore the index movement. Thus, there appears to be both a quality and an index change. The true number of judgements therefore is smaller than it appears.

Given this careful, conservative programme of quality assessment, the following questions arise: what is its impact on the index, and what would be the result if different treatments were used? We have recalculated the index numbers for the study period using a number of scenarios reflecting different treatments of quality change.

2.3 Applying different assessments of quality change

The first scenario replicates current practice.[8] The second scenario splices all quality changes—an option that is easy to adopt, and which underlies some uses of scanner data. The third scenario splices large price changes only, but keeps the existing results for small changes. The fourth scenario is the reverse of the third—adjustments for large price changes are kept, but small changes are spliced. The purpose of scenarios 3 and 4 is to see the relative impacts of large and small changes on the index. The fifth scenario, a simplistic one, calculates the index with all quality changes ignored and all price changes accepted as pure price change. Finally, there is a variant, scenario 1a, which does not reflect an alternate method of assessing quality changes, but employs a different computational handling of splices. In this scenario, spliced quality changes are excluded from the matched sample for the month in which the quality change occurs.

Table 1 shows the separate results for both common sizes of televisions, and for the two combined, for the whole study period. A large part of the drop in prices occurred between 1990 and 1992. The prices used in this study were exclusive of retail sales taxes, so the replacement of the manufacturers' sales tax by the Goods and Services Tax probably accounted for some of this drop.

TABLE 1

Percentage change in TV index, Canada, June 1990 to November 1997

Scenario		All TVs	20" TV	27" TV	27" TV less 20" TV
1	Current practice	-23.6	-20.5	-27.3	-6.8
2	All quality changes spliced	-20.7	-18.6	-23.3	-4.7
3	Large price changes spliced	-21.7	-19.1	-24.8	-5.7
4	Small price changes spliced	-22.6	-19.9	-25.9	-6.0
5	All price changes treated as pure price changes	-23.5	-20.2	-27.6	-7.4
1a	*Current practice, most splices excluded*[9]	*-24.5*	*-21.5*	*-28.2*	*-6.7*
	Range of results Scenarios 1 to 4	**2.9**	**1.9**	**4.0**	

[8] *The first scenario does not recreate the historical indexes exactly, for a number of reasons, including that the regional strata were simplified for these calculations.*

[9] *Those cases where neither the price nor the quality changed were not regarded as quality changes, so they were not taken out of the sample for the calculation.*

Three things are apparent in Table 1. First, however treated, prices for 20" TVs have fallen substantially less than prices for 27" TVs. This is true whatever method of handling quality change is used. The differences for scenarios 1 to 4 (5 would not be used in practice) range from 4.7% to 6.8%. By contrast, the range of results from these different scenarios applied to one size of television is only 1.9% for 20" televisions and 4.0% for 27" televisions. This suggests that making sure the sample selection is representative is more important in this case than choosing the best quality adjustment technique.

Second, the range of results from different scenarios (1.9% to 4.0%) is small compared to the overall price movement (20% or more). This is confirmed indirectly by the behaviour of the 89 streams of observations that were in the sample for the whole study period. The greatest amount of price change, about $\frac{5}{6}$, occurred in months where there was no quality change. To some extent this is due to the low inflation during this period. When there were quality changes, the number of downward replacements offset the number of upward ones, so their impacts tended to cancel out. This suggests that if the replacements reflect market changes, many people traded up to higher-priced items instead of benefiting from lower prices.

As should be expected, the largest difference among the scenarios 1 to 4 is between scenario 1 – current practice, and scenario 2 – splicing all quality changes. This difference, for all televisions together, is 2.9%. If we regard splicing as the default treatment we see that applying other judgements lowered the index by 2.9%. The impact of the judgements on large price changes was about twice as much (1.9%) as applying it to small changes (1.0%), even though splicing was more common for large changes. It is clear from looking at the patterns of quality change adjustment in the appendix tables that the net impact of assessing small price changes was to lower the index, but it was not obvious that the judgements on the larger price changes would have more impact. There were approximately equal numbers of upward and downward adjustments, and one must wonder how many of the splices on large price changes should be replaced by judgements.

It is curious that the simplistic approach, scenario 5, produces a result close to the official index (scenario 1). This closeness is an accident of the time period. Between 1990 and the end of 1991, the index under scenario 5 fell sharply compared to the official index, then rose from the end of 1995. These periods correspond to the periods of greatest weakness and recovery in the Canadian economy, and the result is consistent with consumers trading down, then up, accordingly. This provides some validation of the changes in item selection that have occurred over the period, despite the limitations imposed by the specifications.

Third, the computational practice of keeping the spliced observation in the matched sample for the month in which it is spliced has a relatively significant impact. The impact is about one-third the impact of applying judgements. Scenario 1a shows that the drag on indexes by splicing was about 1% over the period, while the impact of quality adjustment was almost 3%. As prices were falling for this commodity, splicing has kept the index higher. Preliminary testing on other commodities suggests that this may be a general result, particularly for durable goods whose prices are tending to decline.

3.0 Calculations based on scanner data

A large retailer, with many stores across Canada, provided the data used for these calculations. The data contain the number sold and average price for each identified product code by month and by store. The price is the actual transaction price before taxes. Data from the stores were aggregated to create one average price and one total quantity for each product code for each month. The product codes distinguish models to approximately the same level of detail as our official CPI survey—for example, a new production run under a different model number will carry a different product code in this database. The number of product codes reporting sales in any month is about 200 for all stores. The company carries only a few manufacturers, but a full range of products from those manufacturers. The product code description provides enough information to identify the make and model, so by using brochures or consulting manufacturers, the characteristics of each can be obtained. That has not been done yet; however, the description routinely includes the screen size, which has been useful in distinguishing subsets of the range of models.

The range of models comprises five groups: 20" (19" to 21"); 27" (25" to 29"); 32/35" (31" to 36"); 13" (including 9"); and projection TVs. Indexes are calculated for these specifications separately and grouped together.

It would be useful to trim the data of insignificant and unreliable records. As usual, most of the revenue is concentrated in a few products, though the concentration curve is relatively flat. The best selling product only accounted for about 8% of sales usually, and there were rarely more than two with even 5% of the total revenue. Over the whole thirteen months 93 product codes accounted for 80% of total sales.

If the data are to be used for index production, we need criteria to determine what data should be accepted in the current month. As a first step, we considered only those product codes that contributed to the top 80% of sales in each month, and calculated monthly estimates on matched data. The results of the chained series over the twelve monthly comparisons from February 1997 to February 1998 are given for a variety of formulae in Table 2.

TABLE 2

February 1998 TV index (February 1997=100), chaining monthly matched samples for products accounting for 80% of each monthly total sale

					Details within "Other"		
	All	**20"**	**27"**	**Other**	**32/35"**	**13"**	**Proj. TV**
Laspeyres	94.7	98.9	94.8	90.2	87.8	103.6	87.2
Paasche	91.7	95.9	89.9	89.6	89.1	98.4	88.1
Fisher	93.2	97.4	92.3	89.9	88.5	100.9	87.6
Tornqvist	93.1	97.5	92.4	89.8	88.4	101.0	87.9
Geometric (base-weighted)	94.2	98.6	94.3	89.5	87.3	103.2	86.7
Geometric (current-weighted)	92.3	96.3	90.4	90.3	89.7	98.8	88.6

Most patterns are as expected. The base-weighted indexes are higher than the current-weighted ones, except for the 32/35" televisions and the projection TVs. The Fisher and Tornqvist indexes are almost identical. The 27" set declined in price more than the 20" set as it did in the CPI survey. But the most striking result is a comparison with the official CPI. Between February and November 1997, the CPI for televisions rose slightly, while all the indexes based on scanner data fell substantially. (During this period, the indexes fell more than what is shown in Table 2 because from November 1997 to February 1998, the scanner indexes all rose.[10]) How do we account for these differences? Some possible explanations are examined below.

1. ***The items listed as "Other" in this database are not included in the CPI, and their indexes have fallen the most.***

   This is true, but the indexes for 20" and 27" televisions are also significantly lower than in the official CPI survey.

2. ***There was a shift in 1997 towards the higher-priced products—the largest televisions and projection TVs. The monthly chained index reflects the substitution between representative commodities that is not reflected in the CPI.***

   This shift is shown in Table 3.

[10] *When this paper was written, the CPI was not available for the whole period February 1997 to February 1998.*

TABLE 3

Percentage distribution of TV revenue, by product group, February 1997 to February 1998

	20"	27"	32/35"	13"	Proj. TV	Other*
February to May 1997	21.4	44.3	18.3	7.2	8.3	0.4
June to August 1997	20.2	40.3	24.4	8.0	6.8	0.3
September to November 1997	16.7	42.5	24.2	6.3	9.8	0.5
December 1997 to February 1998	14.8	36.1	27.3	5.7	14.2	1.9

***The main sales under this category were screens for projection televisions.*

As the data are only for one year we cannot be sure that these figures are not just seasonal variations. However, comparison of the distributions for February 1997 and February 1998 shows a similar shift.

Nevertheless, the assumption that the monthly chained index is lower because of substitution among televisions of different sizes does not explain the results. Taking the indexes of the five product groups separately and weighting them with the pattern of the first four months, the Fisher indexes produce an overall index of 91.8 instead of 93.2.

The reason for this inconsistency has to do with how representative the scanner sample was. Because the sample was restricted to a subset of the top 80% of total sales, a higher proportion of 20" and 27" televisions got included. For projection televisions, in particular, there were a wide variety of models sold, few of which had sufficient weight to be counted among the top 80%. On average over the year, 72% of sales of 20" TVs and 83% of sales of 27" TVs were included in the matched samples. For the other groups the percentages were 69% for 32/35" TVs, 60% for 13" TVs, and 46% for projection TVs. For both the 32/35" models and projection TVs, the representation was lowest early in the period when prices were falling fastest. So, although the chaining reflected the shifting retail patterns, the larger televisions were routinely under-represented, in what was a self-weighted sample.

3. The selection criteria produce something akin to the bouncing effect. If a product drops below the 80% sales level for even one month, it drops out of two months' comparisons. If this drop in sales is due to a relatively high price, that product will be excluded when prices are rising.

In the calculations, this appeared to be a problem. We computed indexes on different criteria—under which the condition for initial inclusion in the sample was the same—but once in the sample, a product stayed for as long as it was reporting sales. However, under these conditions the indexes came out virtually the same—Fisher 93.1 instead of 93.2. We believe this is because the criterion keeps old models in the sample longer when their sales are small, but are at greatly discounted prices.

4. When new items are linked into the sample, there is a ratchet effect if they are introduced at relatively high prices, and then the prices are cut. The drop shows up, but not the initial increase.

This is demonstrated in Table 4, where Model 2 differs from Model 1 only in its model number.

TABLE 4

Monthly chained index calculations when one TV model replaces an identical one in the market

	Model 1		Model 2		Combined average price of models 1 and 2 $	Indexes (February 1997=100)			
1997-98	Units sold	Average price $	Units sold	Average price $		Laspeyres	Paasche	Fisher	Based on combined average price
Feb.	91	846			846	100.0	100.0	100.0	100.0
Mar.	99	850			850	100.5	100.5	100.5	100.5
Apr.	66	850			850	100.4	100.4	100.4	100.4
May	73	845			845	99.9	99.9	99.9	99.9
June	68	844			844	99.8	99.8	99.8	99.8
July	53	828			828	97.9	97.9	97.9	97.9
Aug.	85	778	15	883	794	92.0	92.0	92.0	93.8
Sept.	63	732	73	874	808	87.3	89.1	88.2	95.5
Oct.	17	697	79	863	834	84.9	87.5	86.2	98.5
Nov.			87	852	852	83.8	86.4	85.1	100.7
Dec.			114	845	845	83.2	85.7	84.4	99.9
Jan.			68	884	884	87.0	89.6	88.3	104.5
Feb.			55	904	904	88.9	91.7	90.3	106.8

The first three index computations are based on the assumption that Models 1 and 2 are not directly comparable. From February to August, and after October 1997, the movements of the three indexes—Laspeyres, Paasche and Fisher—are identical. That is because the prices of one model only are driving them. Even in August, a true Paasche index cannot be calculated because there is no observed price for Model 2 in July, and in November, a true Laspeyres index cannot be calculated because there is no observed price for Model 1 in that month. These three measures only differ between August and October.

The last column in Table 4 is based on recognising that the two product codes describe identical models, so the sales data can be combined. The index is based on the weighted average price each month. The difference is remarkable. It is not obvious why there should be such a large difference between the prices of the two models in the same stores at the same time. Most sales were made in stores where both models were available, although the price of Model 2 was lower on average where Model 1 was not available, and the price of Model 1 was higher where Model 2 was not available. If this situation is widespread it clearly has a large downward impact on the index measurement. We know that there are many substitutions of this kind, where the new model is identical to the old, or differs only slightly. However, it is hard to say what the overall impact of this ratchet effect will be. In this case, it appears that the introduction of the newer model at the higher price—together with the reductions on the older model—may have been designed to clear old stock. We do not know whether this is true generally.

Table 5 shows the results of direct comparisons between February 1997 and February 1998 for those models that were available in both periods.

TABLE 5

Index (February 1997=100) based on direct comparisons from February 1997 to February 1998 for available TV models

	All	20"	27"	32/35"	13"	Proj. TV
Laspeyres	89.3	93.7	88.7	83.3	90.0	84.6
Paasche	91.6	96.1	92.8	81.1	94.1	88.6
Fisher	90.4	94.9	90.7	82.2	92.0	86.6

These results are lower than, usually much lower than, the monthly chained indexes. Of course, at the beginning of the study period, these models could have been recent introductions to the sample, and their prices could have been in the process of falling from initially high levels. We cannot confirm this. Furthermore, while the models included account for 75% of sales in February 1997 (90% of 20" and 27" TVs), their share had fallen to 20% by 1998 (nearly 30% for 20" and 27" TVs). This reinforces the necessity to update the sample selection promptly. Nevertheless, these models, whose prices were falling steadily, were available, albeit with a declining share of the market. Prices in the rest of the market could not have been rising relative to this subset unless consumers were very uninformed, or put a high premium on novelty.

5. The price behaviour of the company supplying data for this study is not typical of the market as a whole.

We cannot ascertain whether the shift to higher-priced products in 1997 was experienced generally throughout Canada. We can compare the price behaviour of the respondent company with others in the CPI sample because some of the CPI data are collected from this company. Two simple comparisons were made: between the subset of CPI prices from the respondent company and the CPI generally, and between the scanner results for those models observed in the CPI and all models for the same size of TV in the scanner database.

The results for the period February to November 1997 (putting February 1997=100) for TVs are: CPI data from the respondent company 102.2 for 20" TVs and 104.1 for 27" TVs. The official CPI index for TVs (20" and 27" together) was 101.2, slightly though not significantly lower. For scanner data for those models selected in the CPI, the Fisher indexes were: 98.0 for 20" TVs (94.9 for all 20" models) and 87.7 for 27" TVs (88.9 for all 27" models). While there are variations, particularly for the 20" sets, these results do not suggest the selection is particularly unusual.

Another comparison that can be made is to use the scanner prices that match the models selected in the CPI and use the CPI evaluation of quality differences to adjust the prices to a constant quality equivalent. When the models are the same the scanner data for the models are averaged together, as illustrated in Table 4. As seen in one case, when a model is replaced by another valued at $\frac{8}{9}$ the quality, the prices of the second model, devalued by $\frac{1}{9}$, are averaged with the first. Using these adjusted values the Fisher index for 27" TVs became 91.3—closing about $\frac{1}{6}$ the gap between the scanner and CPI index movements. On the other hand the index for 20" TVs became 98.1, virtually unchanged from the unadjusted series. This is, of course, much too small a test for the result to be extrapolated.

4.0 Conclusion

The recitation of the methods and alternatives in treating quality change in the CPI survey demonstrates the limited impact that they have on the index calculation, even compared to the impact of sample selection at the detailed level.

The differences between these indexes and ones obtained from scanner data are striking and difficult to explain. The scanner data have to be managed carefully. There are challenging questions to answer concerning the choice of the subset of data to be included in the calculations, whether to group together different products or to weight together their separate indexes, and how to introduce new items into the calculation.

Of these, the problem of introducing new products into the index may be the most important. It is clear that the assumption underlying any use of matched samples—that relative prices reflect relative qualities—does not hold here. At the very least, replacements that are really continuations of the same product under a different brand name must be recognised. More generally, the kind of transformation grid for different characteristics, derived from regression analysis described in Silver et al. (1997), should be applied if the characteristics can be obtained quickly enough.[11] However, the few comparisons that have been made using prices of constant quality suggest that the range of impact between doing quality adjustment well or badly is still small compared to the difference between the scanner data and the official CPI survey results.

The benefits of using scanner data may be greater in the area of electronics (compared with food items for example) because of the wider dispersion of products and the more rapid turnover of models. However, analysis of these data throws some doubt on the assumption that it is not possible to give enough attention to individual scanner data reports. There are 10 streams of observations from the respondent company for televisions in the CPI survey. In the first set of matched samples, there is only an average of 35 streams of data, which covers a broader range of products (large and smaller TVs and projection TVs). For 20" and 27" TVs, there is only an average of 20 streams. This is not a number of a different order of magnitude, and it may be practical to reduce the scanner sample further. Of course, the number of new products which have to be monitored in case they reach the threshold of being included in the index, is rather higher. Nevertheless, a possibility may be to combine the constant updating of weighting data and actual transaction prices with the statistical care applied to the evaluation of quality changes.

[11] *While there has not yet been time to collect and code the characteristics of models in the scanner data, there is sufficient description on the current CPI sample to test some characteristics. The regression run on November 1995 data produced dummy variables for manufacturers quite similar to Silver's study (viz. all adjustments referenced to Sony).* ***Silver****: Panasonic .944, Toshiba .931, Hitachi .905, JVC .844, Sharp .834, Sanyo .826, Samsung .786;* ***CPI****: Panasonic .909, Hitachi .907, JVC .861, Sharp .812, Toshiba .793, Sanyo .779, (Zenith .773), Samsung .756, (RCA .738). Except for Toshiba, which has few models in the CPI sample, the rankings are identical and the factors are similar.*

References

Dalén, J. (1997) "Experiments with Swedish scanner data". Proceedings of the Third Meeting of the Ottawa Group, International Conference on Price Indices, pp. 163-168, April 16-18, 1997, Voorburg, Statistics Netherlands.

de Haan, J. and E. Opperdoes (1997) "Estimation of the coffee price index using scanner data: Simulation of official practices" and "Estimation of the coffee price index using scanner data: The choice of the microindex". Proceedings of the Third Meeting of the Ottawa Group, International Conference on Price Indices, pp. 183-202, April 16-18, 1997, Voorburg, Statistics Netherlands.

Hawkes, W.J. (1997) "Reconciliation of consumer price index trends with corresponding trends in average prices for quasi-homogeneous goods using scanner data". Proceedings of the Third Meeting of the Ottawa Group, International Conference on Price Indices, pp. 145-162, April 16-18, 1997, Voorburg, Statistics Netherlands.

Lowe, R. (1997) "Item selection and quality change in the Canadian CPI". Proceedings of the Third Meeting of the Ottawa Group, International Conference on Price Indices, pp. 275-290, April 16-18, 1997, Voorburg, Statistics Netherlands.

Reinsdorf, M. (1995) "Constructing basic component indexes for the U.S. CPI from scanner data: A test using data on coffee". Presented at the National Bureau of Economic Review (NBER) Conference on Productivity, Working Paper 277, July 1995, Cambridge, Massachusetts, U.S. Bureau of Statistics.

Scobie, H. (1997) "Use of scanner data – The impact of new goods". Unpublished paper prepared for the meeting of the Price Measurement Advisory Committee, May 1997, Ottawa, Statistics 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*, Series 41, Number 4, December 1995, Cardiff Business School, University of Wales College of Cardiff.

Silver, M., C. Ioannides, and M. Haworth (1997) "Hedonic quality adjustments for non-comparable items for consumer price indices". Proceedings of the Third Meeting of the Ottawa Group, International Conference on Price Indices, pp. 203-214, April 16-18, 1997, Voorburg, Statistics Netherlands.

Appendix tables

TABLE I

Distribution of quality changes for televisions, July 1990 to November 1997, by ratio of replacement item to replaced item

Price ratio	Total quality changes	Non-splices (judgements)	Splices	Percentage of total changes that are spliced	Index ratio for judgements[1]
0.5 to 0.69	15	5	10	66.7	0.917
0.7 to 0.79	65	31	34	52.3	0.882
0.8 to 0.89	153	67	86	56.2	0.904
0.9 to 0.99	271	173	98	36.2	0.948
1	274	45	229	83.6	0.968
1.01 to 1.09	187	114	73	39.0	1.043
1.1 to 1.19	136	64	72	52.9	1.073
1.2 to 1.29	66	32	34	51.5	1.112
1.3 or more	46	18	28	60.9	1.083
All changes	**1213**	**549**	**664**	**54.7**	
Large changes (ratio <0.9 or >1.09)	481	217	264	54.9	
Small changes (excl. ratio=1)	458	287	171	37.3	

[1] This is the average calculated pure price ratio for those judgements (i.e. non-splices) where the ratio of nominal reported prices falls into each category. As can be seen, the estimated pure price ratio rises with the nominal price ratio.

TABLE II

Distribution of quality changes for televisions, January 1993 to November 1997, by price ratio of replacement item to replaced item, by whether brand changes with the quality change

	Brand does not change				Brand changes			
Price ratio	**Total quality changes**	Judgements	**Splices**	**% of total spliced**	**Total quality changes**	Judgements	**Splices**	**% of total spliced**
0.5 to 0.69	4	1	3	75.0	3	1	2	66.7
0.7 to 0.79	8	3	5	62.5	15	3	12	80.0
0.8 to 0.89	51	25	26	51.0	32	3	29	90.6
0.9 to 0.99	139	98	41	29.5	21	8	13	61.9
1	170	29	141	82.9	28	3	25	89.3
1.01 to 1.09	107	80	27	25.2	32	18	14	43.8
1.1 to 1.19	50	30	20	40.0	24	11	13	54.2
1.2 to 1.29	14	8	6	42.9	19	8	11	57.9
1.3 or more	9	3	6	66.7	24	8	16	66.7
All changes	**552**	**277**	**275**	**49.8**	**198**	**63**	**135**	**68.2**
Large changes (ratio <0.9 or >1.09)	136	70	66	48.5	117	34	83	70.9
Small changes (excl. ratio=1)	246	178	68	27.6	53	26	27	50.9

Splices are more likely to occur when the brand (manufacturer) changes. This is partly because brand changes are more common with large price changes—major replacements—but the rate of splicing when the brand changes is higher for every price ratio.

TABLE III

Distribution of judgement quality change assessments for televisions, January 1993 to November 1997, by price ratio, by whether brand changes, and by whether quality is judged to be changed or not

	Brand does not change				Brand changes			
Price ratio	**Total assess-ments**	**Quality changed**	**Quality not changed**	**% quality unchanged**	**Total assess-ments**	**Quality changed**	**Quality not changed**	**% quality unchanged**
0.5 to 0.69	1	1	0	0.0	1	1	0	0.0
0.7 to 0.79	3	3	0	0.0	3	3	0	0.0
0.8 to 0.89	26	15	11	42.3	3	2	1	33.3
0.9 to 0.99	97	22	75	77.3	8	3	5	62.5
1	29	29	0	0.0	3	3	0	0.0
1.01 to 1.09	80	15	65	81.3	18	6	12	66.7
1.1 to 1.19	30	25	5	16.7	11	9	2	18.2
1.2 to 1.29	8	6	2	25.0	8	8	0	0.0
1.3 or more	3	3	0	0.0	8	8	0	0.0
All changes	**277**	**119**	**158**	**57.0**	**63**	**43**	**20**	**31.7**
Large changes (ratio <0.9 or >1.09)	71	53	18	25.4	34	31	3	8.8
Small changes (excl. ratio=1)	177	37	140	79.1	26	9	17	65.4

Again, brand (manufacturer) changes occur more often when there are large price changes. Nevertheless, when the brand changes, a replacement item is less likely to be judged of equal quality to the replaced item, whatever the circumstances. Most quality change assessments when the price change is small turn out to find no difference in quality. Many of these assessments probably involve nothing more than a change in the label of the same model—though clearly not when the brand changes.