Meeting of the Ottawa Group
25 – 27 August 1999, Reykjavik, Iceland

Introduction

The fifth meeting of the International Working Group on Price Indices, the Ottawa Group was held in Reykjavík, Iceland, 25-27 August 1999.

These proceedings include the final versions of the papers presented, summary of the meeting by Robin Lowe and Louis-Marc Ducharme and other meeting information. The proceedings are also available on a CD ROM.

The topics under discussion were two: Treatment of quality change in price indices for durable goods and different concepts of price indices serving different purposes.

The participants were 36 and there were 18 papers on these topics available at the meeting.

Many people were active in the preparation of the meeting both at Statistics Iceland and on behalf of the Ottawa Group Steering Committee and I would like to thank them all for their efforts.

The next meeting of the Ottawa Group will be held in Canberra, Australia 2-6 April 2001 and the preparations for that meeting are well on the way now.

Summary

Treatment of quality change in price indexes for durable goods
Chaired by Bert Balk, Mick Silver and Jack E. Triplett, 26 August 1999.
By Robin Lowe.

In this session there were nine papers that addressed various aspects of the issue. One group described, with some evaluation, various methods used to deal with perceived quality changes in the indexes (usually the CPI) in their countries: The Netherlands, Norway, Germany and Canada, (Hoven, Sandberg, Hoffman and Schultz, respectively). Another group dealt with the actual or potential application of so-called hedonic methods in their indexes: U.S., France, Canada, and U.K., (Greenlees, Lacroix, Lowe and Silver, respectively). The other paper by Sellwood, recently of Eurostat, examined some problems arising from the apparent conflict between economic theory and its application in practice. There was animated and useful discussion of the issues raised in the papers, and if this report concentrates on the drawbacks and failures of methods used it is because this points the way to future work on the subject.

Comparison of the papers and discussion showed that there were wide variations in practice and in philosophical approach. Resolution of these differences was hindered by the lack of a common vocabulary, and a failure to put problems into a proper context. Sellwood pointed out that the problems of representing the universe are different from the problems of keeping the sample doing what it was designed for. In his view, representing the universe, (until possibly the application of scanner data,) is the task of sampling, and the treatment of new goods, including the evaluation of any consumer surplus added by increased variety is its domain. Quality change evaluation is concerned with maintaining a sample to its specifications, until that sample should be replaced. The problems of evaluating quality change within a sample are not assisted by the theory of index numbers that deals with

comparisons of universes. This view, however, is not universally accepted; some would expand the scope of quality change evaluation to include changes in the universe.

All countries use some of the following methods to make adjustments for perceived quality change: a) find a replacement of equal quality, b) find a replacement eventually but in the meantime estimate the price movement (implicitly) as the movement of the remainder of the sample, c) as b) but the explicit estimate is on the movement of some other aggregation, d) find a replacement available at the same time as the disappearing product and show no change in the index movement due to the replacement, e) when the replacement is only available in the next period stretch the assumption in d) to show no index movement between the two periods, f) when the change is that an option becomes standard accept all or part of the prior option price as quality, g) when the difference can be expressed wholly as a difference in quantity adjust the reported price movement by the change in that quantity.

Most organisations have names for some of these methods; they vary, and they suggest contrary things to others. Adopting a common vocabulary would help intelligent discussion and understanding. Ralph Turvey has been working on a glossary for Eurostat, and it was agreed to support this work and its results.

Each of the methods on this list share the property that once the conditions for its use have been determined, the application and result is automatic. (This is also true for the hedonic method, once the model has been designed.) However, although the adjustment is automatic once it has been triggered, it is not devoid of judgement, which has to be used to determine whether to recognise a change and which method to use. Different organisations have widely varying preferences for the use of different methods and impose different conditions for their application. Many of these have arisen from operational practices, and aspects of sample design and product selection. For example, an organisation that relies on casual employees for price collection may use only a) or b), as only the simplest conditions can be put on finding replacements. Similar concerns arise when the price collection is very decentralised – it is difficult to control, so only simple options may be used. Whether the specification of the surveyed product is tightly or loosely drawn also has an effect both on the incidence of perceived quality change and on its treatment.

It is thus impossible to advocate particular adjustments without understanding the operational and sampling system embedded in the system. Eurostat has discovered this in its work on harmonised indexes. The only agreement so far has been to ban e) as a routine adjustment method. Some warnings may be given, however: i) Methods b), c), d), and e) produce similar results, as they all depend on the assumption that relative prices reflect relative qualities. There were references to evidence that systematic use of any of these can lead to consistent errors in one direction or another as replacements are made whose prices are not both representative of the market; ii) method a) can lead to traps –replacements may be made without documentation (as there is no quality change), and therefore no verification, and there will be a tendency to accept a replacement item at the same price as being of equal quality though its features may have changed. Whatever methods are used, there is still the need to monitor price collection closely, both to recognise quality change when it occurs, and to assess whether the usual adjustment technique is satisfactory.

Explicit intervention to make quality adjustments that are based on an evaluation of each individual case is relatively unusual. Most organisations, if they cannot apply one of the automatic adjustment methods, reject the comparison (which implicitly means adopting b)

above). What intervention that is done is based on information from manufacturers' costs (e.g. automobiles), or from their brochures (appliances, electronic products), or by a knowledgeable price collector examining the products directly (clothing). These are all used primarily in situations where a product and its replacement are rarely in the market at the same time, and either the older product is discounted before its disappearance, or the new model is introduced at a higher price because of its novelty value (or both). These methods are only as good as the market intelligence obtained and they are subject to a consistent error if always misapplied. There have been arguments that where these interventions have been used automobile and clothing indexes have been too high, as too many "improvements" have been recognised.

An interventionist method that can be applied to individual adjustments objectively has been the use of regression on characteristics, or hedonic methods. The four papers on experience with this also showed much diversity. Although attached to consumer theory the method has similar problems in practice as other methods of adjustment for quality change. Sellwood's question – how do we tell whether to accept or reject a regression study - elicited only that it depends on knowing how the technology and market for the product works, which begs the question. Scientific method is to try to replicate other people's results, but we are not in the position to do that yet. There has been too much variance in products selected, time periods and periodicity and no agreement on the characteristics to collect in different countries even for the same products. There was agreement at this meeting by several countries working in this area to seek out common products and common or joint research activity in these areas. It will concentrate on home computers, major appliances and home entertainment products.

It has been observed that most hedonic studies produce results lowering the measure of price change. This may have something to do with the choice of subjects. In areas where there is not much technological change – clothing and large appliances – results are sometimes higher. They are typically lower, sometimes much lower, for those products where there is rapid technical change, and these are areas recognised to be particularly difficult. It is not clear in some studies whether the lower results are due to improved quality change evaluation, the effect of product substitution, or a better data set of prices. The hedonic method depends on the assumption that relative prices observed reflect relative qualities, and it is in these areas that that assumption is rejected in conventional quality change adjustment. Also, it is observed that the method cannot deal with new characteristics, and while new in this context has not been clearly determined, the distinguishing characteristics in many advanced products gradually filter from the most luxurious models to the standard ones, and it is not clear that the valuation of these characteristics is handled properly. Hedonic methods appear to work best where the product variation is among a set of alternatives that does not change much over time (clothing), or where the shift in products is in characteristics that are enduringly important (computers).

Almost everyone has commented on the resources necessary to produce hedonic indexes. The models need to be run frequently, apparently, and they need a lot of data. This has still to be tested. There are no studies yet on how much data is needed to estimate 'good' models, nor on how robust quality change adjustment is, given alternative models. But assuming that large data sets are required, then other alternative approaches to measurement, extending beyond maintaining a constant sample, become possible. The use of electronic data records ('scanner data') may eliminate the need to evaluate one-to-one replacements in small samples, though the appearance and disappearance of items from the scanner data set introduces other problems. There is a little evidence that just matching subsets of these

records may produce similar results to the hedonic results run on them. This will be useful if it means that it will not be necessary to collect the range of characteristics. This question is another that remains to be examined.

One paper quoted Adam Smith's comment that on quality all information is uncertain. That remains the case. Error is endemic to the problem, not only because there will be no agreement on the right answer. Because the situations are so varied and the methods of adjustment differ in so many ways it is not possible to describe the effect of "quality change adjustment". There are different potential errors and biases in every area, not just in consumer durables which were considered in this session. In all organisations there is a lack of an analytical framework to test and evaluate different approaches. This is one of the reasons that the group agreed to build on the Ottawa Group website a meta database where research on quality change can be stored and results of particular studies in different organisations can be posted and compared. The purpose of this, apart to improve the flow of information, is to support the statisticians in index making offices, upon whose constant attention to the data, in specific cases, and design of samples, the accuracy of indexes depends.

Discussion about the future working plans of the Ottawa Group
Chaired by Louis Marc Ducharme, 25 August 1999.
By Louis Marc Ducharme.

The purpose of the session was twofold: First, to report to the Group on the progress made on a number of issues raised at the last meeting in Washington on the objectives, organisation and output of the Group. And second, to seek the views on the future agenda of the Group as proposed by the Steering Committee.

Summary:
In Washington, a number of concerns were raised about the Ottawa Group, including:

- There was a feeling that the Group lacked a mid and long term agenda.
- That the benefit of the Group's deliberation was limited mainly to the National Statistical Offices of participants, due to the lack of clearly defined output and wide dissemination of the papers and the outcome discussion.

In order to examine some of these concerns, it was agreed to create a Steering Committee. The current committee is composed of Bert Balk, Louis Marc Ducharme, John Greenlees, Rósmundur Guðnason, and Keith Woolford. Thierry Lacroix has recently agreed to join the Steering Committee. The Committee had two tasks: the first one was to re-state the terms of reference and formalise the organisation of the Group and the second one was to outline the work plan for the near future.

About the terms of reference:
The Committee agreed to the original terms of reference, which in a nut shell says:

"… to provide a forum for specialists and practitioners who work for, or serve as advisors to national statistical agencies or international organisations to exchange their experiences and thoughts on crucial problems of measuring price change…"

"… the focus of the Group is on applied research particularly, thought not exclusively, in the area of consumer prices indices."

For the Steering Committee, this meant that the Group could not be regarded as an official body able to bless or establish standards, however, it would be appropriate for the Group to make recommendation on best practices to national or international bodies.

Recommendation:
In terms of its scope, the Group will continue to focus mainly on CPI issues, but recognised that some issues may be addressed from other perspectives (i.e. PPI).

About participation:
The Committee agreed that only specialists actively working on problems under discussion would be invited to the Group meeting as active participant. For the sake of efficient discussion, it is necessary to limit the number of participants. The host country can, however, invite observers to the meeting as has been the tradition so far.

Recommendation:
Given that some issues may be of interest to experts from other fields of expertise (i.e. PPI), the Group agreed that it would welcome a limited participation of these experts when the topics of proposed papers correspond to the agenda of the meeting and may contribute to the discussions. It is strongly recommended that CPI and other national experts co-ordinate their participation to the meeting. In order to avoid augmenting the number of participants to the meeting, preference, if necessary, will be given to CPI specialists. If there will be over-participation, the Steering Committee reserves the right to determine who will participate.

About the agenda:
The Steering Committee has proposed to focus the Iceland and Australia meetings on problems of quality adjustments with some secondary topics. At this Iceland meeting, two themes were discussed: I) Treatment of quality change in price indices for durable goods and ii) Different concepts of price indices serving different purposes. For the meeting in Australia, in March 2001, three main topics are proposed:

1- The first topic focus on the construction (conceptual and practical problems) of price index numbers for services, particularly services of telecommunication and utilities. These are often characterised by complex ways of charging (subscription, all inclusive prices, pre-paid cards, packaging services, etc.), rapid market change (ex: in the case of telecommunication: interlacing and merging tendency of various technologies and media) and the difficulties for price statisticians to obtaining detailed information in a quickly evolving market.

2- The second one is related to the treatment of quality change for durable goods, in continuity with the work plan undertaken in Iceland and;

3- The third one relates to various effects on prices: Subsidized consumption, bonus cards, rebates and promotion on retail prices and products and services becoming free of charges. Submission on a limited number of contributed papers, not falling under the three topics mentioned above will be considered by the Steering Committee when finalizing the agenda.

About the relationship between the Ottawa Group and the TEG
The Steering Committee proposed to keep an agenda distinct from the agenda of the two Technical Expert Groups (TEG-CPI and TEG-PPI) as they attempt to fulfil different objectives.

Recommendation:
The Steering Committee is prepared to allocate a session to its future agenda to provide comments and advice to the TEGs on the preparation of the manuals. In fact, as the membership of these TEGs shows, there is already a good deal of collaboration through the participation of the same individuals in these activities.

About output and the work plan
At last year's meeting in Washington, the Ottawa Group decided to devote a large part of its work programme on the production of guidelines and outlines of best practices for quality adjustments. The agenda of both the Iceland and Australia meetings reflect this priority.

This year, the Committee proposed that for the Iceland's meeting we not only disseminate electronic proceedings of the papers, but also a summary of discussions.

Two proposals were made on the future work plan for the Ottawa Group:

1. First, Statistics Iceland proposed the creation of a meta-database on the treatment of quality changes, where all available information about quality adjustments would be maintained in one place. The database would be divided into three parts. The first part would contain the collection of all available and known written information about the methodology of quality adjustment indicating where it is available. The second part would extend to all available information on techniques of quality adjustments. The third part would contain voluntary information from different countries with descriptions of practical quality adjustment procedures carried out (See proposal of Guðnason, R. Ottawa Group Quality Database).

Decision:
Statistics Canada has agreed to host the database on the Ottawa Web-site and Statistics Iceland agreed to participate in the working group to develop its organisation and content. Eurostat has agreed to share information on their database on quality.

2. Second, it was proposed that some countries get together and apply the same type of quality adjustments (hedonic) to similar goods in order to test if a single model could be applied on different markets. A number of hedonic projects already exist in a number of countries for television, washing machine, personal computer. The first step will be to collaborate and establish comparable results and compare the treatment of data for similar goods. This will include testing hedonic models to see whether they are similar across different markets, and when scanner data is available, comparing its behaviour. These results could be presented at the next meeting. The longer term objective would be to share the experience of hedonic model over a number of Agencies.

Decision:
A number of Statistical Agencies (The Bureau of Labor Statistics, INSEE, the Office of National Statistics and Statistics Canada) expressed interest in the project.