Assessing the Quality of SPPIs

J. Wood, M. G. Šova and I. Richardson
Office for National Statistics, UK

Abstract

The construction of Service Producer Price Indices (SPPIs) faces problems which are not encountered to the same extent for Goods Producer Price Indices. These problems include: difficulties in defining individual service products on a constant quality basis; difficulty in separating business and retail components; lack of well-developed international standards for the specification and classification of service products. Because of these problems, assessing the quality of SPPIs involves additional concerns: the effects of non-standard data collection methods; the suitability of indices for use in intermediate or final consumption; the reliability of sampling frames for price collection. This paper describes how the UK Office for National Statistics is tackling the questions of measuring quality, setting quality standards and creating procedures to monitor and review the quality of SPPIs.

Keywords: Classifications, Quality Review, Sample Allocation, Sample Surveys, Services Producer Price Indices

1. Introduction

The Office for National Statistics (ONS) in the UK has been publishing Services Producer Prices (SPPIs) since 1996, although we have only recently adopted this internationally standard terminology. Until 2006, ONS published such indices under the name 'Corporate Services Price Indices' (CSPIs). As with most other National Statistical Institutes (NSIs) in the industrialised countries, the development of SPPIs in ONS is relatively recent. This contrasts with Producer Price Indices (PPIs) for goods, which have existed in the UK since 1903. In 1987 the UN-sponsored Voorburg Group began a series of annual, international conferences on official statistics for services, including SPPIs. ONS has been a regular contributor to and beneficiary from the discussions at these conferences.

The development of SPPIs by the ONS began in earnest in 1995 with the creation of the CSPI Branch to develop price indices for a wider range of service products. This development proceeded on a product by product basis, with little attempt to develop an integrated structure for SPPIs. The emphasis was on increasing the SPPI coverage of the service sector, so efforts were concentrated on those products which had a relatively large weight and whose pricing structure allowed the relatively straightforward development of indices on the same basis as the longstanding PPIs. At first, ONS only published indices for separate products. Only in 2003 did an integrated, 'overall' SPPI appear but the coverage of this overall index is still limited.

ONS still classifies all SPPIs as 'experimental', a designation given to statistics which are still in development but are published to allow user feedback on their usefulness and fitness for purpose. In 2004, a review of official statistics in the UK, the so-called 'Allsopp Review' (Allsopp, 2004), recommended major changes to the provision of official statistics. The two main thrusts of the review were the need for more regional statistics (not relevant to this paper) and the need to develop service sector statistics, specifically SPPIs and the National Accounts Index of Services (also an 'experimental' statistic at the time). Recommendations for the latter included increasing the coverage of service sector statistics and improving their quality so that they might be regraded from 'experimental' statistics to high quality National Statistics. After a six-year period of industry reviews, leading to major improvements in its component parts, the Index of Services achieved National Statistics status in March 2007.

Following the successful development and publication of SPPIs for 30 separate products and the Allsopp recommendation to attain National Statistics status for SPPIs, ONS, in 2005, devised a three-year programme of quality reviews for all existing SPPIs, similar to the programme of industry reviews applied to the Index of Services. Unfortunately, resource constraints meant that this programme could not be implemented as planned. Rather than abandon the programme completely, ONS commissioned Dr Markus Šova to conduct a brief, summary review of the 30 existing SPPIs, to identify the main factors affecting their quality and to make recommendations for improvements. Although this did not provide a full programme of review and improvement, as applied for the Index of Services, it maintained some momentum

in the drive to improve the quality of published SPPIs and provided a basis for the future improvement of these indices when resources allow. The review was conducted throughout 2006.

This paper presents the results of this review in the context of a more general discussion of the factors affecting the quality of SPPIs. We hope that our experience will provide useful information to other NSIs for their SPPI quality review programmes and that the paper will stimulate discussion and debate on how best to improve the quality of SPPIs.

Section 2 provides a brief description of what SPPIs are and what specific factors differentiate them from PPIs for goods. Section 3 summarises the findings and recommendations from the review of UK SPPIs. Section 4 discusses other factors affecting the quality of SPPIs. Section 5 presents ONS's plans for improving and monitoring the quality of UK SPPIs. Section 6 concludes with a general discussion on how the quality of SPPIs may be assessed and improved.

2. What are Services Producer Price Indices?

SPPIs are measures of average price movements for business-to-business services. They are the direct equivalent, in the service sector, of Producer Price Indices for goods. In the UK, this similarity is reflected in almost identical methods of construction: PPIs and SPPIs are all calculated as Laspeyres indices which are rebased every five years. The current base year is 2000 for both PPIs and SPPIs and both sets of indices are undergoing a concurrent programme of rebasing to base year 2005. So great are the similarities, that ONS's prices division has merged its PPI and SPPI teams into combined operational and development teams, each team covering both PPIs and SPPIs, and aims to integrate the PPI and SPPI systems as much as possible, although full integration may take several years.

The two main uses of SPPIs are to estimate price inflation for business-to-business services and to provide deflators for the Index of Services. Clearly, a major customer for SPPIs is ONS's National Accounts Group but there is also a large number of external customers: the Bank of England uses SPPIs for the monitoring of inflation; Eurostat requires SPPIs for a similar purpose and a recent short-term statistics regulation imposes further requirements on member states to produce SPPIs; other Government Departments and commercial organisations use SPPIs for escalation clauses in contracts; academics and researchers use SPPIs for service sector research purposes.

Despite the similarities between PPIs and SPPIs, there are still some major differences. Operationally, in the UK, SPPIs are produced quarterly and PPIs are produced monthly. This partly reflects the relatively recent development of SPPIs in the UK. Because of the need for the limited resources of the then CSPI Branch to pursue both the development of new indices and the regular production of existing indices, production of SPPIs on a quarterly basis ensured the supply of some short-term statistics on business-to-business price inflation, where there was none before, and allowed time, between the quarterly production rounds, for the development of new statistics. This disparity may change as the needs of users change and the integration of PPIs and SPPIs develops.

Another operational difference is in the data sources for index weights and the sampling frame for product selection. PPIs benefit from an annual survey of products required by Eurostat regulation, the 'ProdCom' survey (from Products of the European Community). This provides a data source, in conjunction with other sources, for the estimation of base year index weights. It also provides an annually updated sampling frame for the rotation and replenishment of the sample of products. This ensures that the overall sample size is maintained and that the sampled products reflect changes to products in the market.

There is no annual products survey for services. SPPIs are based on a fixed panel of products, which is replenished and augmented on an ad hoc basis, according to the needs imposed by the attrition of products and respondents. In the early years of development of SPPIs in the UK, the sampling of individual products was mainly purposive and the weights for product categories and individual products were not always reliable, being based on whatever data could be obtained from external and internal sources. In 2001, ONS conducted a survey of service product turnover in 2000 for those product classes then covered by published SPPIs. This led to much improved product category weights for rebasing to base year 2000 and the survey microdata provided a sampling frame for the updating and augmentation of the sample of products. An enhanced survey in 2006 for turnover in 2005, coupled with improvements to estimation methods, will further improve the index weighting for rebasing to base year 2005, which will be implemented in 2008. It will be many years, however, before the annual 'ProdCom' survey is expanded to cover service products and thereby secure for services the same, high quality source of data enjoyed by PPIs for the estimation of weights and provision of a products sampling frame.

Although the two service products turnover surveys mentioned have produced and will produce major improvements to the quality of UK SPPIs, both were subject to one major deficiency. The early development of SPPIs in the UK involved detailed consultations with relevant trade associations to identify appropriate product types and data sources within the relevant industry. This meant that the construction of SPPIs was, and still is, based only on those products whose four-digit product class is the same as the four-digit industry class to which the producing enterprise belongs. The coverage of the SPPIs is therefore not complete, because they exclude products whose class is different from the industry class of the producing enterprise. For the same reason, the two turnover surveys are not fully effective because we ignore potentially useful information for such products. This contrasts with the annual 'ProdCom' survey, which covers all manufacturing classes, both product-based and industry-based, and therefore provides more complete coverage for goods.

Expanding the 'ProdCom' survey to include services would not only allow us to increase the coverage of SPPIs and increase the efficiency of the survey (by using more of the data provided), it would also have concomitant benefits for PPIs. Many enterprises, especially the larger ones, produce both goods and services. Their assignment to a manufacturing or service industry class is determined by the relative preponderance of the products sold. The exclusion of service industries from the 'ProdCom' survey therefore, to some extent, reduces the coverage of PPIs through omission of goods produced by enterprises assigned to service industries.

As well as these operational differences between SPPIs and PPIs, there are also conceptual differences which make assessing the quality of SPPIs particularly difficult. For example, international standards for the classification of service products are poorly defined compared to goods. The recently revised, European Classification of Products by Activity, CPA (2008), used as the basis for classifying PPIs, contains 84 pages of definitions for the production industries but only 40 pages for service industries, although services account for about 80% of GDP in the UK. The following extract is typical of the level of detail provided for service products:

56.3 Beverage serving services
56.30 Beverage serving services
56.30.1 Beverage serving services
56.30.10 Beverage serving services

Contrast this with the much more detailed prescriptions for the equivalent goods:

11.0 Beverages
11.01 Distilled alcoholic beverages
11.01.1 Distilled alcoholic beverages
11.01.10 Distilled alcoholic beverages
11.01.9 Sub-contracted operations as part of manufacturing of distilled alcoholic beverages
11.01.99 Sub-contracted operations as part of manufacturing of distilled alcoholic beverages
11.02 Wine from grape
11.02.1 Wine of fresh grapes; grape must
11.02.11 Sparkling wine of fresh grapes
11.02.12 Wine of fresh grapes, except sparkling wine; grape must
11.02.2 Wine lees; argol
11.02.20 Wine lees; argol
11.02.9 Sub-contracted operations as part of manufacturing of wine from grape
11.02.99 Sub-contracted operations as part of manufacturing of wine from grape
11.03 Cider and other fruit wines
11.03.1 Other fermented beverages (e.g., cider, perry, mead); mixed beverages containing alcohol
11.03.10 Other fermented beverages (e.g., cider, perry, mead); mixed beverages containing alcohol
11.03.9 Sub-contracted operations as part of manufacturing of cider and other fruit wines
11.03.99 Sub-contracted operations as part of manufacturing of cider and other fruit wines
11.04 Other non-distilled fermented beverages
11.04.1 Vermouth and other flavoured wine of fresh grapes
11.04.10 Vermouth and other flavoured wine of fresh grapes
11.04.9 Sub-contracted operations as part of manufacturing of other non-distilled fermented beverages
11.04.99 Sub-contracted operations as part of manufacturing of other non-distilled fermented beverages
11.05 Beer
11.05.1 Beer, except dregs from brewing
11.05.10 Beer, except dregs from brewing
11.05.2 Brewing or distilling dregs
11.05.20 Brewing or distilling dregs
11.05.9 Sub-contracted operations as part of manufacturing of beer
11.05.99 Sub-contracted operations as part of manufacturing of beer
11.06 Malt
11.06.1 Malt
11.06.10 Malt
11.06.9 Sub-contracted operations as part of manufacturing of malt

11.06.99 Sub-contracted operations as part of manufacturing of malt
11.07 Soft drinks; mineral waters and other bottled waters
11.07.1 Mineral waters and soft drinks
11.07.11 Mineral waters and aerated waters, not sweetened nor flavoured
11.07.19 Other non alcoholic beverages
11.07.9 Sub-contracted operations as part of manufacturing of mineral waters and soft drinks
11.07.99 Sub-contracted operations as part of manufacturing of mineral waters and soft drinks

Although ONS has followed the CPA definitions of four-digit product classes, the inadequacy of the CPA at lower levels of detail for service products has forced ONS to develop in-house definitions for lower level categories and subcategories through detailed discussions with trade associations. This is an expensive duplication of effort by ONS and other NSIs and the resultant sets of different, non-standard classifications hinder international comparisons of services producer price inflation.

In the construction of SPPIs, as for PPIs, it is essential that prices are supplied on a like-for-like basis every quarter. That is, there should be no specification or quality changes in the services which contribute to the index. This ensures that any differences over consecutive quarters in the supplied prices, and in the index, arise solely because of genuine price movements. A conceptual difficulty therefore arises in those cases, such as occur in computer programming and other product classes, when each service provided is unique and pricing on a like-for-like basis is not possible. In extreme cases, it is difficult to define the service product at all. Because of these conceptual difficulties, ONS use non-standard methods of price collection for certain service product classes. These methods are discussed in sub-section 4.1 below.

Another conceptual difficulty is in differentiating between prices for business services as against retail services. This differentiation is important for National Accounts purposes, in order to distinguish between price movements for intermediate consumption and price movements for final consumption. Failure to identify differences in price movements between these two modes of consumption in National Accounts deflators will lead to biased estimates of Gross Domestic Product.

3. Quality Assessment of ONS SPPIs

A convenient framework for the assessment of output quality is the set of Eurostat output quality dimensions (Eurostat, 2002 and Mazzi *et al.*, 2005). These are used as part of the standard quality assessment framework within ONS (ONS, 2006). ONS also use the prototype set of process quality dimensions defined in Eurostat (2002). The table below presents a brief list of both sets of quality dimensions.

Output Quality	**Process Quality**
Relevance (including Coverage)	Efficiency
Accuracy	Effectiveness
Timeliness and Punctuality	Robustness
Accessibility and Clarity	Flexibility
Coherence	Transparency
Comparability	Integration

Although the review had these quality dimensions in mind, its results were presented according to different aspects of index number construction, rather than according to these quality dimensions. This is because the review's purpose was to identify weaknesses in the construction of SPPIs and to recommend improvements. It was therefore more important to structure the report according to the stages in the process of construction, which the business area could then work on, than according to these quality dimensions. For example, improving sample selection could affect accuracy (bias and variance), timeliness (when the sampled data are available) and robustness (the reliability of data sources) but all these quality dimensions would be affected by a single change (or collection of changes) to the sample selection process. This brief report of the review's findings follows its categorisation.

The review inspected 30 published indices, which covered 29 four-digit CPA product classes (class 74.83 is published as two separate indices) and 198 product categories (the lowest level of aggregation in ONS's in-house classifications). The complete list of indices reviewed is as follows, classified according to an ONS modification of the CPA (2002).

50.2 Maintenance and Repair of Motor Vehicles
55.1 Hotels
55.5 Canteens and Catering

60.10.1	Business Rail Fares
60.10.9	Rail Freight
60.23.1	Bus and Coach Hire
60.24.9	Freight Transport by Road
61.101	Commercial Vehicle Ferries
61.102	Sea and Coastal Water Freight
63.4	Freight Forwarding
64.11	National Post Parcels
64.12	Courier Services
64.2	Business Telecoms Services
65.121	Banking Services
70.2	Property Rental Payments
70.3	Real Estate Agency
71.32	Construction Plant Hire
74.13	Market Research
74.3	Technical Testing
74.5	Employment Agencies
74.602	Security Services
74.7	Industrial Cleaning
74.819	Commercial Film Processing
74.82	Contract Packaging
74.83a	Direct Marketing & Secretarial Services
74.83b	Translation & Interpretation Services
80.42	Adult Education
90.001	Sewerage Services
90.002	Waste Disposal
93.01	Commercial Washing & Dry Cleaning

The review made 134 recommendations for improvement but many of these were duplicates, in the sense that the same or similar recommendations were made for several indices. Because of the similarity of many recommendations, it is possible to classify the review's findings into three main topics: weighting, sampling, other aspects.

3.1 Weighting

A major finding here was that many weights, for both categories and units, were excessively large or excessively small. Some of the largest unit weights were potentially disclosive because of their dominant impact on the published index. Many of the large unit weights could be attributed to apparent under-sampling of the relevant category (we consider sample allocation in the sub-section *Sampling* below) and the review recommended a review of sample allocation to determine whether increasing the sample sizes for these categories would be appropriate. The very small unit weights are a waste of resources because much time and money are spent in obtaining and using the required price data but such units have negligible impact on the published indices. The wide disparity of weights within a category also increases the variance, and so reduces the accuracy, of the index. The review recommended discontinuing the units with unacceptably small weights and using the resources so freed to supplement the sample where the greatest benefits would be obtained.

Problems identified regarding class and category weights were incomplete and unclear category definitions, doubts regarding the derivations of the weights and the limited coverage of the data used to calculate the weights. These problems are important. Data from the 2000 turnover survey were used to recalculate class and category weights for rebasing to base year 2000 and this recalculation produced major changes, and improvements, to the weighting. Despite these improvements, there were still limitations to the data available, especially at the category level, where data were extracted from a wide variety of ad hoc sources, including some guessing. Coverage problems included the lack of a comprehensive product-based sampling frame, as discussed in section 2 above, and the often unknown coverage of the disparate sources used to calculate the weights. The review recommended the use of a more coherent method of calculating the weights for rebasing, making full use of the 2005 turnover survey data and data from National Accounts Supply-Use Tables.

3.2 Sampling

A major sampling problem was the prevalence of purposive selection of respondents to supply price data. This arose historically from the early years of developing SPPIs, when no suitable product-based sampling frame existed. Early sampling frames were based on membership lists of trade associations and purposive selection was often adopted to ensure inclusion in the sample of the most important suppliers of the services that needed to be covered. Some randomised selection was applied for replenishment and renewal of the SPPI sample, using the sample for the 2000 turnover survey as a sampling frame. This mixture of randomised and purposive selection of respondents complicates the estimation of design-based weights. The review recommended the implementation of randomised rotation of the sample of respondents when resources allow but it is likely to be several years before this recommendation is implemented fully. In fact, it is only since 1999 that the PPI sample has been selected randomly, with annual rotation, rather than as a fixed panel with occasional replenishment.

The second major sampling problem was the evident imbalance in sample sizes for some categories, relative to the category weights. Excessively large or small sample sizes, relative to the category weight, often accounted for the excessively small or large unit

weights discussed in the subsection on *Weighting*. The review recommended that the sample allocations within each class should be reviewed when SPPI standard errors became available to allow an optimal allocation to be calculated. Although the review's brief was to consider each product class index separately, substantial misallocation between product classes had previously been identified. Action taken to correct this misallocation was limited because standard errors were not then available and an accurate, optimal allocation was not yet possible.

3.3 Other aspects

a. Descriptions of product units were often vague. It is important that the specification of these units is precise so that any changes in price caused by changes in specification are recognised and identified as quality changes, not price movements. The review recommended that such descriptions be made more precise.

b. Some SPPIs (employment agencies, real estate agencies) use percentage fees applied to indices obtained from other institutions. Because the percentage fees applied change infrequently, movements in these SPPIs are largely determined by movements in these other indices. The review recommended reviews of these other indices to ensure that they are fit for this purpose.

c. Other SPPIs (rail fares, property rental, maintenance and repair of motor vehicles) use indices obtained from other institutions directly as components of the SPPIs. For these also, the review recommended reviews of these other indices to ensure that they are fit for this purpose.

d. Definitions of 'Europe' were often badly specified. Some categories, mainly in transport and communications, have a geographical scope and, for the UK, it is often useful to distinguish 'Europe' from other destinations. However, depending on the context, the meaning of 'Europe' may vary (for example: the continent of Europe, the European Union, member states of the European Monetary Union). Recent changes to the composition of the European Union and the European Monetary Union also complicate the application of these definitions, even when they are clear. The review found one instance of a product unit in the 'non-Europe' category which should have been in the 'Europe' category because of the accession of new member states of the European Union. As with product unit definitions, the review recommended that definitions of 'Europe' be made more precise.

e. The review also identified a variety of processing risks in the compilation of the indices. There is extensive use of spreadsheets, often with manual transcription of data between spreadsheets, giving rise to a considerable risk of transcription error. The review also commented on the lack of a single, unified system for the calculation of SPPIs. Such a system would help to reduce the processing risks identified.

4. Other Factors affecting the Quality of SPPIs

Of the 30 ONS indices reviewed, 23 are based on samples of price data for individual product units (although three of these are purposively selected in their entirety), five are based on indices supplied by other institutions (three directly and two by the application of percentage fees for sampled units) and two are unit value indices. There is much literature on the use of non-standard methods of constructing SPPIs, so it is perhaps surprising that there are only two such indices in the 30 indices reviewed. One reason for this is that the indices ONS developed initially were those with fewer of the conceptual difficulties which give rise to alternative methods of construction, such as unit value indices and the use of model contracts and hourly rates. It is likely that the development of new SPPIs within ONS will have to make greater use of such non-standard methods.

4.1 Non-standard Methods of Price Collection

Non-standard methods complicate quality assessment and quality comparisons between different indices because the most important aspects of quality are different for different methods. For example, unit value indices offer complete coverage of the population and, unlike the standard indices based on sampled prices, are therefore not subject to sampling error. However, if movements in unit values arise from a changing mix of products rather than price movements, the unit value index will be biased. In extreme cases, this bias can be financially significant. Assessing the relative merits of such different indices is not easy.

The following paragraphs consider the most important quality aspects of these non-standard methods.

4.1.1 Unit Value Indices

These indices provide complete coverage of the population with no sampling error but there is a potential bias, which can be financially significant.

They are relatively cheap and easy to implement because the aggregate data required are usually provided by external organisations at little or no cost. However, the dependency on other organisations may mean that receipt of data is not sufficiently timely. The lack of any quality adjustment in the data is also a concern. As an example of the possible problems arising from such indices, the ONS had to withdraw its unit value business telecommunications index in 2006 because a serious bias was identified. The index was re-instated in early 2007 after improvements to data supply and a reconstruction of the index. This example demonstrates the need for constant, careful monitoring of these indices.

4.1.2 Hedonic Modelling

This is appropriate where products are multifarious or unique but may be defined by a finite set of identifying characteristics. It requires a timely and reliable data supply for model fitting and the model needs to include the most important characteristics. Problems may arise if the model doesn't fit well or important characteristics are omitted. ONS considered this approach for professional services but the data available were inadequate.

4.1.3 Model Contracts

These are time-consuming and labour-intensive to construct and impose a heavy burden on respondents. Prices supplied are not actual prices charged, so there is a potential bias, which is exacerbated if supplied prices are not updated in line with actual prices. Because of the heavy respondent burden, sample sizes tend to be small, which further reduces the quality of this type of index.

4.1.4 Hourly Rates

These are relatively cheap and easy to develop but the cost of production is similar to that for indices based on sampled prices. There is potential bias from inadequate quality adjustment which, if it exists, is likely to be limited and subjective. The appropriate weighting to use for the hourly rates collected may be difficult to assess. ONS are considering using this approach for the development of new indices because of its relative ease of implementation and the limited staff resources available.

4.2 Other Factors

4.2.1 Scope and Coverage

As discussed in section 2, current price collection for ONS's SPPIs is limited to those industries which share the same 4-digit class code as the relevant product. There is a need to increase the scope of price collection to obtain product prices from industries with different class codes, thereby also increasing the sample's coverage of the product population. Ideally, this would be done by expanding the ProdCom survey to include services. This would have the added advantage of covering goods produced by service providers and services provided by goods manufacturers. Current ONS resources do not allow this for the near future. The best that we might hope for is that the products survey for the next rebasing in 2010 will be designed with an extended scope and coverage.

4.2.2 Annual Chain-linking

Rapidly changing markets (telecommunications is a prime example) create methodological and practical problems in the construction of price indices. Changes in the quality of products, with the continual appearance of new and disappearance of old products, mean that samples have to be continually updated and replenished, and product weights become out of date. This can lead to serious bias in the indices with regard to short-term price movements, as was found at the last rebasing to base year 2000, when rebased estimates for SPPIs and goods PPIs showed substantially lower inflation than the previously published estimates. The solution to these problems is to update the product weights and to replenish the samples annually. This also helps to counteract any bias in unit value indices. For annual chain-linking to be done effectively requires extension of the annual ProdCom survey to cover services, as discussed in the previous paragraph, but lack of resources makes this option a long-term aspiration rather than a near-term aim.

4.2.3 Timeliness

Use of non-standard data sources (such as administrative data, unit value data or indices produced by other institutions) can lead to delays in the receipt of data. This can require imputation of the most recent index numbers (such as by forecasting), which reduces their value as an up-to-date indicator. In extreme circumstances, the very reliability of the data source may be in doubt, with some risk of price data not being available at all.

4.2.4 Cost

Different methods of price collection and different data sources have different associated costs. This is relevant to sample allocation, if the aim is to produce the most accurate estimate of SPPI inflation possible for a given cost. For example, prices obtained directly from a website are relatively cheap to obtain and may justify a larger sample size for the relevant product class. ONS does not currently allow for differences in cost when deciding on sample allocation but the wide disparity in costs for different methods of price collection would justify investigation of this when resources allow.

5. ONS Plans to improve and monitor the Quality of SPPIs

ONS have started a programme of work to implement the more tractable of the review's recommendations. This programme includes the immediate re-specification of vague descriptions for product units and of 'Europe', the gradual reduction of processing risks and careful monitoring of the unit value business telecommunications index. ONS have deferred implementation of the more resource-intensive recommendations until sufficient resources are available.

ONS also plan to incease the accuracy of the top-level SPPI by improving the sample allocation. Standard deviations of price movements for all product classes are now available to inform optimisation of the allocation (see Bucknall *et al.*, 2005, for a description of the standard error methodology). Optimisation has identified some large disparities between the current and optimal allocations. These may have arisen because in the early years of development of SPPIs in the UK, the main concern was to investigate and develop indices for individual product classes, with little regard for their relative importance and variance. Only now, when we have a substantial body of product class indices and standard deviations of price movements are available, is it possible to investigate the question of optimal allocation.

The optimal allocation suggests that some rationalisation of classifications is required. Some categories with small weights should be merged to produce a viable combined category while some categories with large weights should perhaps be split into more homogeneous, smaller categories. However, there are also concerns about the reliability of the source data for some classes, casting doubt on the reliability of some aspects of the optimisation. For example, in employment agencies, it may be large implicit standard errors in the earnings data used which give rise to the relatively large standard deviations of price movements observed for this class. It may be more effective and efficient to improve the accuracy of the source data than to implement the large, suggested increase in allocation for this class. ONS shall investigate this and other anomalies before applying large-scale re-allocations.

Because of the limited resources available, ONS will, for the time being, confine allocation changes to major misallocations. Detailed re-allocation will be deferred until after rebasing to 2005. This will also allow a re-optimisation of the allocation using the more up-to-date 2005 weights, which will also have benefited from proposed improvements in estimation methods.

ONS intend to review again the quality of SPPIs after rebasing, re-allocation and other recommendations have been implemented. This further review will include new indices not covered in the review described in this paper. Currently, however, resource constraints are a major hindrance to the timely implementation of quality improvements and are likely to remain so for the foreseeable future.

6. Discussion and Conclusion

The focus of the quality review described in this paper has been on making recommendations to enable the practical implementation of quality improvements. The recommendations made may be classified into three main types: those that are outside the control of the production team; those that are within the control of the production team but are too resource-intensive for immediate implementation; and those that can be implemented immediately.

Recommendations that are outside the control of the production team include: extension of the annual ProdCom survey to cover services; annual chain-linking; and rationalisation of the index production system to reduce processing risks. All of those require major resources, whose allocation is decided by senior managers, who also have to consider the relative merits of demands for resources from many other business areas. Of these recommendations, annual chain-linking is likely to produce the greatest benefits to output quality but it can only be implemented in conjunction with the other two recommendations. All three recommendations should therefore be treated as a single requirement and only the commitment of senior managers to provide adequate resources for the implementation of these recommendations will allow this major improvement to the quality of UK SPPIs.

Recommendations that are within the control of the production team but are too resource-intensive for immediate implementation include a complete sample re-allocation and more detailed review of classes identified as having specific problems. The production team will allocate staff resources to the implementation of these recommendations when staffing levels have increased.

Recommendations that can be implemented immediately include: better product unit and geographical definitions; the most important changes to sample allocation; and rebasing. The limited sample re-allocation to be implemented is deliberately rough-and-ready because of residual uncertainties regarding some product class allocations and to accommodate practical constraints. Nonetheless, it should produce a noticeable improvement in the accuracy of the top-level SPPI. Because rebasing is a planned, five-yearly occurrence, resources are automatically allocated for it but its implementation will produce improvements to the quality of UK SPPIs in addition to those generated from the updating of weights. Estimation of the weights has been the subject of careful investigation and estimation methods will be improved since the previous rebasing. Rebasing also provides an opportunity for the rationalisation of classifications, in the light of the optimised allocation. Detailed sample re-allocation after rebasing will benefit from the updated weights and their improved reliability. However, several years will be required to implement these recommendations fully and implementation will be applied in a controlled manner to match available resources and business priorities.

The quality review presented here is only a first step in a difficult and complicated area. The review has identified several shortcomings in existing ONS SPPIs and has prepared the way for important improvements to these indices. Many of these shortcomings have arisen because SPPIs are relatively new and the major focus since ONS starting developing them in 1995 has been to develop relevant product classifications and set up index production on a regular basis. Because of the limited resources available for these developments, they have necessarily been applied on a piecemeal basis but ONS staff involved in these developments have made significant progress over the last dozen years. Only now that a substantial body of published indices has been set up has it been possible to consider an integrated approach to assessing quality and preparing plans for further improvements. Many years will pass before these further improvements are implemented fully, so this review is by no means a definitive statement on the quality of ONS SPPIs. Rather, it is better regarded as the latest stage in the continuing improvement of ONS SPPIs.

References

Allsopp, C. (2004). *Review of Statistics for Economic Policymaking*. HMSO. http://www.hm-treasury.gov.uk/consultations_and_legislation/allsop_review/consult_allsopp_index.cfm

Bucknall, R., Sova, M., and Wood, J. (2005). Estimating standard errors of movements in Producer Price Indices. *Tenth GSS Methodology Conference, 27 June 2005*. http://www.statistics.gov.uk/events/gss2005/agenda.asp

CPA (2002). *Classification of Products by Activity 2002*, Eurostat. http://ec.europa.eu/eurostat/ramon

CPA (2008). *Classification of Products by Activity 2008*, Eurostat. http://circa.europa.eu/irc/dsis/nacecpacon/info/data/en/index.htm

Eurostat (2002). *Quality in the European Statistical system- the way forward*, Office for Official Publications of the European Communities. http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1073,46587259&_dad=portal&_schema=PORTAL&p_product_code=KS-45-02-814

Mazzi, G.L., Museux, J.-M., and Savio, G. (2005). *Quality measures for economic indicators*. Office for Official Publications of the European Communities. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-DT-05-003/EN/KS-DT-05-003-EN.PDF

ONS (2006). *Guidelines for Measuring Statistical Quality*, Office for National Statistics. http://www.statistics.gov.uk/qualitymeasures