Addressing cross-cutting issues arising in the development of SPPI's for use in the measurement of Service Sector GDP

– 2008 Revision

Paper updated for the 23rd Voorburg Group, Aguascalientes, Mexico 22-26 September 2008

Matt Berger
Office for National Statistics, United Kingdom

PREFACE TO 2008 REVISION

The 22nd meeting of the Voorburg Group in Seoul, Korea, 2007, saw a presented paper concerning general methodological issues associated with the development of SPPI's for measurement of service sector GDP[1].   This paper is a revision of the 2007 paper to include key concerns raised by the delegates attending the meeting in Korea in 2007. The key changes incorporated in this revision are summarised in Appendix 1.

1   Introduction to original 2007 version

The Voorburg Group was initiated with the objective of establishing an internationally comparable methodology for measuring the deflated or constant dollar outputs of the service industries. A new approach to this objective commenced with the 21st Voorburg Group meeting in Wiesbaden, 2006, where papers on particular service industries were presented as plenary sessions by both producer price index developers and turnover survey practitioners, and then summarised in a closing session. With the new format trialled for the industries of Road Freight Transport and Management Consultancy Services, a key outcome of the 21st Voorburg was the emerging view that a subset of issues encountered in specific industries was indeed common across many industries within the Service Industry sector. Consequently, recurring themes arose during the meeting of the Group. The strategy session discussion concluded further that these issues would likely arise, again, in future meetings.

2   The problem : Service Industry Statistics[2]

Services are typically produced and delivered in direct contact with the customer, and as a consequence there is a more frequent occurrence of unique service products than in the equivalent case of the production of goods. A unique service product is one that is only provided once to the specifications of an individual customer, and prices cannot be observed over multiple periods. Various business services fall into this category, making price measurement a daunting task. Examples include legal, accounting and consultancy type services, architectural and engineering services, and more frequently "complete logistic solutions" in the transport and storage industries.

Further, service producers often employ *pricing mechanisms* that are complex, at least from the point of view of the statistician attempting to measure price of the service. These mechanisms vary from hourly fees (which are not complex for the service provider to administer but cause considerable complications for statisticians when used in *pricing methods*), to the bundling of services.

---

[1] Berger (2007)

[2] This section borrows heavily from OECD-Eurostat (2005), particularly Chapter 2

National statistical offices are frequently confronted with a combination of both unique services and complex pricing mechanisms. The type of *pricing mechanism* employed by a service provider, together with the particular unique service being delivered, limits the types of *pricing methods* which national statistical agencies may adopt to measure price change. This limitation is further complicated by the compliance cost that some methodologies place on establishments responding to statistical instruments.

A consequence of these limitations is that national statistical offices often choose to employ less preferred pricing methods when constructing SPPI's. For example, time based pricing methods[3] are frequently employed because service providers levy a fee based on hours worked (a pricing mechanism). In many cases the choice of pricing method is undertaken because that method is the only applicable tool that will guarantee survey returns each period. The questions arise, are there adverse implications of using such methods (particularly with respect to measurement of GDP), and can these implications be mitigated?

The business of official statistics in these instances can be thought of as a (very) complicated balancing act. The ongoing constraints of time and agency resources need to be balanced on the one hand against the final use of the statistics, on another hand against the conceptual model which underpins the measurement, and yet on another (!) hand against the compliance cost and response issues associated with adopting a given measurement strategy. Giving too much weight to the conceptual aspect risks delivery of the statistics in any form; yet too little consideration renders the resulting statistics unfit for purpose. This is complicated yet further when there are multiple uses of the statistics.

This paper describes the common difficulties encountered in the measurement of turnover and producer price indexes for services, and to acknowledge the practical compromises adopted by national statistical offices. It further highlights the assumptions of adopting these methodologies, and possible consequences for national accounts when using such measures.

3   Key findings from Wiesbaden 2006

The meeting in Wiesbaden adopted the process of a two-part cycle for establishing methodology for particular service industries. The first year would see the sharing of experiences of individual national statistical offices with the Group (for both producer price indexes and turnover), with a summary session indicating the types of issues to be included in a unifying "sector paper" to be presented in the second year. When both Road Freight Transport and Management Consultancy Services were presented, the following themes arose in common:

   a) *Measurement of product versus measurement of industry*
   b) *Determining scope*: the need for clarity regarding international services versus exports of services, and the need for guidelines when confronted with subcontracting, outsourcing and offshoring.

With particular regard to Management Consultancy Services, the Group discussions also indicated that several issues encountered thus far were known to be (or considered likely to be) issues for other types of business services, in particular:
   c) *Use of time based methods for PPIs*
   d) *Impact of changes in production techniques*

With regard to Road Freight transport, the Group discussions also indicated that one group of issues encountered thus far were known to be (or considered likely to be) issues for other types of transport services, in particular:

---

[3] Time based methods and their limitations are discussed in further detail in later sections.

e) *Bundled services, and Multi-modal solutions*: where the pricing mechanism groups together a range of separate services, or where the production of a service incorporates several other services (e.g. road freight travelling via ferry).

These issues or themes will be discussed from a conceptual viewpoint as well as consideration of the practical compromises that national statistical offices are often forced to adopt.

4 Measurement of product versus measurement of industry

In an ideal[4] world, economic activity, real world business units, classifications and statistical units would align perfectly. Establishments[5] would only undertake primary and ancillary activities, doing away with the need for the definition of secondary activities (and thus secondary products). Industry and product classifications would align so thoroughly that all industry statistics could be considered aggregation of product level statistics. In such a statistician's utopia, each business unit would be split into homogeneous statistical units, each of which undertook one (and only one) type of activity, and hence only produced one type of product. The measurement of product would be equivalent to the measurement of industry.

Moving on from this abstract ideal we see that

- business units are defined according to business needs, not statistical standards;

- certain economic activities produce more than one product simultaneously;

- the same products may be produced by using different techniques of production;

- otherwise identical outputs with different uses or different economic destinations need to be treated as separate products[6];

- product classifications provide information beyond merely usual industry of origin;

- production functions evolve and businesses both horizontally and vertically diversify over time; and

- under the most rigorous statistical units model, even the most homogeneously defined statistical units inevitably undertake secondary production.

Thus it is not possible to consider product statistics as simply a disaggregation of industry statistics, nor industry statistics as some aggregate of product data.

---

[4] "Ideal" here solely taking the context of "ideal for service industry statisticians to measure in practice." See 1993 SNA 5.40 through 5.47 for a discussion of industries and homogeneous statistical units.

[5] As per the 1993 SNA definition:

  5.21 The establishment combines both the kind-of-activity dimension and the locality dimension. An establishment is defined as an enterprise, or part of an enterprise, that is situated in a single location and in which only a single (non-ancillary) productive activity is carried out or in which the principal productive activity accounts for most of the value added.

  5.22 Although the definition of an establishment allows for the possibility that there may be one or more secondary activities carried out, they should be on a small scale compared with the principal activity. If a secondary activity within an enterprise is as important, or nearly as important, as the principal activity, then that activity should be treated as taking place within a separate establishment from that in which the principal activity takes place. The definition of an establishment does not permit an ancillary activity to constitute an establishment on its own.

[6] SNA93 Paragraph 6.107 "Differences in quality that are attributable to differences in the physical characteristics of the goods or services concerned are easily recognized, but not all differences in quality are of this kind. There are other factors which can give rise to differences in quality For example, goods or services delivered in different locations, or at different times, must be treated as different qualities even if they are otherwise physically identical. These differences stem from the fact that the marginal utility of a particular kind of good or service for purchasers or consumers situated in one location may be very different from that for purchasers in other locations, while the costs of delivering goods or services in different locations also vary. Transporting a good to a location in which it is in greater demand is a process of production in its own right in which the good is transformed into a higher quality good."

Noting this, the Methodological Guide for SPPI's states[7]:

> "This guide assumes that the main use of SPPI is deflation in national accounts, and therefore suggests that SPPI be product-based rather than industry-based. Following recommendations of the SNA/ESA, the GDP compilation in countries is increasingly based on the framework of supply and use tables, which means that data on output of industries are available by product groups and are not shown only as a sum of various types of products. In this situation, it is preferable to deflate output by product groups and to separate primary from secondary production rather than use a single deflator for the total output of an industry. In other words, by using product-based PPIs national accountants may themselves estimate price and volume development for industry outputs where changes in the composition of output are taken frequently into account.
>
> Having product SPPI available for deflation purposes is more practical rather than using an industry-based price index, which would require frequent re-weighting to match with the industry output in national accounts."

In the measurement of service industry GDP, the aim of the SPPI should be a *product level index*; that is, it should measure changes in the prices of products regardless of the industry of the establishment producing them.

In addition to the requirements for the use of SPPI's in the national accounts, development of SPPI's at the product level improves the SPPI's themselves, since compilation of product level indexes results in groupings of more homogeneous activities, better ensuring that period-to-period price measurement is undertaken to constant quality.

The development of product level SPPI's is frequently constrained by competing uses of price indexes (inflationary analysis and contract escalation, in addition to the aforementioned product level deflation). But the key constraint to development of detailed product level SPPI's is *data availability* at the product level.

National statistical offices appropriately give high priority to measurement of turnover (revenue) by *industry*, as such measures are vital economic indicators in their own regard, and key components of many other aggregate economic indicators such as the national accounts. Accordingly, much of the infrastructure of official business statistics has been developed as a consequence of the needs of industry turnover measurement. Business registers and other sampling frames classify establishments according to industry, and do not typically feature information on service industry *products*. Subsequently any sampling schemes to identify product level data are multistage samples[8]. This approach is typically employed for manufacturing PPI's (for example, the Products in the European Community inquiry, or PRODCOM, forms the second stage frame for manufacturing PPI's for many national statistical offices in Europe); yet such detailed data sources are not readily available for service industry products.

The nature of service industries means measurement of product level revenue is a complex task – the same issues that face price index practitioners are encountered when attempting to measure turnover at the product level. In particular, the intangible nature of many services, and the bundling of service products, mean that both the identification and measurement of product level data is complex. Surveys undertaken to measure service products will be of higher burden (compliance cost) than their equivalents in the manufacturing industries, and higher burden than those concerned with measurement of service industry outputs alone.

---

[7] OECD-Eurostat (2005), paragraph 13

[8] First stage samples establishments within industry, second stage samples products within establishments.

Compared with the manufacturing statistics, limited existing data are available from which to draw service product samples, or to use for service product weighting; development of product level SPPI's require other data sources. Consequently, the SPPI's developed by national statistical offices are frequently a mixture of industry and product SPPI's.

Whereas product level SPPI's are the desired tool for use in the measurement of service sector GDP, the use of industry SPPI's is possible under the following strict assumptions. Such assumptions need to be considered on an industry-by-industry (and indeed product-by-product) basis, the validity of which may vary between national economies and, within national economies, may vary by region or market.

The assumptions that permit the substitution of product SPPI's by industry SPPI's are:

a) product and industry almost directly align - the industry has negligible secondary production, and almost all production of the product occurs within the one industry; *or*

b) that both

   i. price movements align - secondary products within the industry have the same price movements as the primary product[9], *and*

   ii. products have the same quality *(including *same use)* regardless of the industry that produces them.

The extent to which these assumptions hold indicates the suitability of substituting an industry level index for a product level index. If these assumptions do not hold, the use of industry level SPPI's to deflate product level data in the national accounts results in biased measures of GDP.

For the service sector, the following observations are of particular note in assessing these assumptions:

- many production functions are driven by labour inputs – where different types of service can be produced by the same individuals, it is reasonable to expect price movements to move similarly (e.g. different types of management consultancy products); but

- many services require specialist labour inputs (e.g. different types of legal professional), and as a consequence there may exist variations in supply (and hence price) at the product level within a given industry; and

- the bespoke nature of many services means than many producers target niche or specialist markets; and

- the inability to retrade services[10] means that price discrimination is a key issue in measurement for SPPI's.

Use of industry SPPI's in the measurement of service sector GDP results in robust GDP measures to the extent that the assumptions of equivalence of service and industry SPPI's are met. Put another

---

[9] There is a particular example that illustrates the need for this assumption: industries offering individual secondary products at non-market prices as incentives to capture segments of the market for their primary product; for example, if banks change prices for insurance policies so that they are below market rates – essentially offering insurance as a "loss-leader" product, a banking *industry* SPPI would show a fall, whereas *product* level indices would only show a fall in the SPPI for insurance. If the banking industry SPPI is then used for banking products, the resulting GDP is an overestimate, and demonstrates upward bias.
[10] See 1993 SNA 16.113

way, the greater the divergence from the assumption of equivalence, the greater the likelihood of bias in resulting measures of service sector GDP.

5   Scope: international services versus exports of services

A recurring theme in Voorburg meetings arises regarding definition of scope, particularly for exports. This issue is frequently encountered in consideration of transport services but also arises in discussions regarding business services. Differences of opinion have been expressed as to when particular services are out of scope[11] of SPPI's and turnover surveys. For example:

- Can legal services be exported (apart from cases heard in international courts of law)?

- Can road freight transport services be exported when vehicles do not cross international boundaries?

- Are road freight transport services in scope if the entire transport activity takes place on foreign soil?

OECD-Eurostat (2005) clarifies this issue as follows[12]:

Note that GDP is not intended to measure the production taking place within the geographical boundary of the economic territory. Some of the production of a resident producer may take place abroad, while some of the production taking place within the geographical boundary of the economy may be carried out by non-resident producer units. For example, a resident producer may have teams of employees working abroad temporarily providing management consultancy services. This output is an export of a resident producer and the productive activity does not contribute to the GDP of the country in which it takes places. Thus, the distinction between resident and non-resident institutional units is crucial to the definition and coverage of GDP. In practice, of course, most of the productive activity of resident producers takes place within the country in which they are resident. However, producers in service industries which typically have to deliver their outputs directly to their clients wherever they are located are increasingly tending to engage in production in more than one country, a practice which is encouraged by rapid transportation and instantaneous communication facilities. Geographical boundaries between adjacent countries are becoming less significant for mobile service producers, especially in small countries bordered by several other countries.

To simplify[13]: an export is determined by "who pays for the service" and not "where the service takes place". So considering the types of examples above,

- legal services are exported if the purchaser is a non-resident

- road freight exports are determined by the residency of the purchaser, and not the geographic region where the activity takes place

---

[11] It is important to distinguish the difference between scope and coverage at this point. Scope here is used to refer to the population of interest ("the service production of all resident producers"). Coverage is used here to refer to that that part of the population that is a survey is attempting to measure; there may be a myriad of practical reasons why some activity is conceptually in-scope but considered out-on-coverage (for example, negligible impact on final results).

[12] OECD-Eurostat (2005), glossary, with some slight amendment to give contemporary examples

[13] This is intentionally an oversimplification of the issues surrounding the measurement of international trade in services, with the aim of illustrating that a service physically delivered within the boundaries of one country can be an export for that country. The subtleties and complexities associated with measurement of trade in services are considered, at length, in UN et al (2002). In particular, the comparable and coherent measurement of services involving natural persons delivering services abroad (for example, temporary employees) is separately addressed in Annex I of UN et al (2002)

The need for separate consideration of exports of services can be considered from two perspectives: first, exports are part of domestic production and need to be included as part of total output. Second, services that have different economic destinations should be considered as different products.

Several practical issues arise when considering their measurement. Foremost among these is the issue of identification. National statistical offices frequently utilise establishment surveys for the measurement of price and turnover data, collecting data from the producer of the services. The focus on characteristics of the producer and not the purchaser may mean that the measurement instruments themselves may not differentiate between delivery of a service for domestic use, and services provided for non-resident purchasers. Determination of such details is frequently better achieved through instruments dedicated to the measurement of international trade in services[14]. One implication here is that when service providers cannot explicitly identify exports, it is unclear whether the resulting statistics are measuring all output (as desired) or only output consumed domestically. The proper resolution of issues is beyond the scope of this paper.

For any given service type, national statistical offices need to make an assessment as to the implications of *not* directly measuring exports of the service. This assessment should consider first the relative size of exports, and second the market in which such services are priced. Clearly, if service providers make no distinction regarding residency of purchasers when setting prices, then the argument for separate measurement of exports of the particular service is weakened. If however there is substantial price discrimination for the export market, the use of a "domestic only" price index for the purposes of deflation will frequently yield biased measurement of GDP. The likelihood of such price discrimination will vary from service to service, and from country to country. An additional consideration is whether exported services have the same quality as those produced for domestic consumption. Examples here would include low-cost (low-feature) airlines as compared with traditional carriers, particularly in the case where the low-cost carriers target non-resident consumers. Different qualities are different products and require separate measurement.

6 Time based methods[15]

The subject of methods based on working time is perhaps the most controversial of all issues discussed at the Voorburg Group, for it is the use of this class of pricing methods which highlights most definitively the practical compromises that national statistical offices are forced to make. Further, time based methods comprehensively illustrate the difficulties surrounding the very definition of services produced, and also the consequences of adopting methodologies which are not sufficiently integrated for the purposes of measurement of service sector GDP.

Methods based on working time do not measure the price received for delivery of service but instead attempt to measure the prices of time spent in service provision. This approach therefore carries the assumption that services correspond directly or predominantly to different types of chargeable hours actually worked for a client.

Time based methods arise for several types of reason:

a) *The service being produced cannot be adequately defined* and as a consequence an input-cost type measure is adopted;

b) *The service is well defined but is tailored uniquely to customers*, and cannot be repeatedly priced to constant quality, with the resulting method adopted measuring components based on working time;

---

[14] As noted, a full treatment of measurement of international trade in services is given in UN et al (2002).

[15] This section uses the OECD-Eurostat definition of time-based methods – that is, a model price method that relies in part on price per hour (or day etc.) is also considered a time-based method. This section borrows heavily from OECD-Eurostat (2005) Section 2.9

c) *Regardless of the service definition, the pricing mechanism adopted by responding establishments is related to time worked*, and the pricing method reflects the book-keeping practices; or

d) *The compliance cost of other methods is high,* and establishments refuse to participate in a timely manner unless presented with a method based on time worked.

When confronted with unique services the most appropriate alternative is to use model pricing[16]. However, for many service industries, or for groups of establishments within these industries, the model pricing approach is burdensome. Burden is applied when the model is defined, when the national statistical office monitors the representativity of model, when the model is updated, revised or changed, and – more frequently – whenever key staff changes occur within responding establishments. The consequence of such burden ranges from, at least, dissatisfied or recalcitrant respondents, to, at the other extreme, high levels of non-response, or data of questionable quality[17]. The impact is therefore potential for bias to enter the measures of price change, and subsequently into derived measures of GDP. Accordingly, national statistical offices frequently depart from model pricing in these cases, avoiding the known problems associated with the application of this particular method.

However, while robust collection of data based on working-time mitigates the known biases surrounding non-response and burdensome compliance cost, the use of such methods themselves frequently results in *biased estimates of GDP*. If a price index is assumed to be equal to the development of hourly rates, the resulting volume measures the time devoted to service provision (that is, the number of hours worked), rather than in the volume of services themselves[18]. More explicitly, use of a time based method for SPPI's tends to result in biased estimates of GDP[19].

Two questions then arise: how likely and important are these biases, and what can be done to mitigate them?

Biases arise in time based SPPI's because they fail to measure changes in labour productivity[20]. No change in the product (service delivered) is determinable apart from changes in staffing structure, and even in such cases the underlying assumptions are that labour productivity does not change "within staffing levels". The extent to which this assumption is not valid is an indication of the degree to which any resulting measure of GDP is biased. In particular, there is a strong argument that suggests that changes of staffing structures themselves are strongly correlated with changes of roles and duties within a "staffing level"; and a change of roles and duties is explicitly a change in the quality (utility) of the labour provided. Thus, contrary to the aims of such a measure, SPPI's that assume that labour productivity does not change within staffing levels are failing to price to constant quality, arguably including times of change between "staffing levels".

The key issue is that time-based SPPI's are a *compromise*, and that such a compromise is sensitive to changes in the scope of billable working hours. For example, consider the situation where law firms adopt electronic searches of case law, rather than sending clerks to manually search law libraries. Charge-out rates for legal partners might increase to cover the introduction of new

---

[16] OECD-Eurostat (2005), 2.8.1

[17] Typical examples of poor quality data include long series of nil price changes (in the presence of dynamic labour conditions), or changes where all the components of the model move by exactly the same amount.

[18] Time-based methods are often used for price indexes when it is difficult to define the service being provided; it should come as no surprise that the resulting measure of volume of services is also ill-defined. This is a characteristic of the difficulty of measuring services. See the appendix for an example of the relationship between volume change and hours worked.

[19] There is one exception to this statement: measures based on working time do not result in biased estimates of GDP when the product is "provision of hours of labour" – that is, "supply of personnel services" Group 8512 of CPC1.1. This does not extend to products where the service provider charges "price per hour", but *only* when the *output of the industry* is "hours placed".

[20] More completely, it might be better to say "changes in labour productivity, and also changes arising because of changes in multi-factor productivity due to, for example, a change in the mixture of labour and technological capital". The impact of technology is discussed later, but a thorough discussion of multi-factor productivity, its definition and measurement are well beyond the scope of this paper.

technology. Such technology clearly allows them to complete tasks more quickly, in turn reducing their billable hours per client, whilst providing the same[21] service to their clients. In such a case the more efficient delivery of the service is not reflected in the price index, and the price index would be upwardly biased; that is, the price index is measuring an increase in the *hourly fee* when the actual *price per service delivered* has not increased. Subsequently, any resultant volume measure (GDP) would be downwardly biased[22].

In attempting to mitigate these potential biases, the implications for national statistical offices are as follows:

a) The key feature to consider is to ensure a consistency between prices and labour input;

b) Prices measured should be realised rates (e.g. realised hourly rates or fee income per grade of worker) – the use of salary or other input costs alone is not considered appropriate, and should not be used unless non-wage labour costs, profit margins, discounting, insurance, rent, and other overheads are also included as part of a model based approach;

c) The use of time based methods should be applied in "bands of expertise", which should be defined on the basis of
   - staffing levels
   - qualification
   - type of activity
   - type of project;

d) Procedures should be adopted to capture changes in duties, roles and responsibilities within "bands of expertise" and where possible quantify them;

e) Mechanisms should be employed to identify changes in coverage of billable hours;

f) Mechanisms should be employed to determine whether realised rates have been influenced by changing technology, administrative reorganisation, or changes to billing structures;

g) Mechanisms should be employed to detect and measure changes in productivity in individual service industries.

Regarding mechanisms to detect changes in productivity[23], this could include specifically designed survey instruments, a series of visits to key or representative respondents, or additional questions as part of SPPI processing. Alternatively this information may be obtainable from government agencies, from industry associations or from elsewhere in the national statistical system[24]. Regardless of the actual mechanisms employed, use of time-based methods for SPPI's necessitates use of tools to detect changes in final product, and changes between the labour inputs and the service delivered; failure to undertake such monitoring ultimately leads to biased estimates of GDP.

One final concern regarding time based methods is their applicability in the presence of "subcontracting", where some portion of labour is supplied by another establishment[25]. This causes additional complications when attempting to use time-based methods as a compromise to model pricing. In such a case, the complication that arises is that the "hours worked" being reported may not align with turnover measures. For example, if a charge-out rate is being approximated by dividing revenue by wage rates, any subcontracted work will be excluded from the wages component but included (implicitly) in the "revenue" component. The use of any such derived approximation to charge-out rate is then problematic; the level of the price will clearly be higher than if subcontracting had not been employed, and it is unclear whether movements in resulting

---

[21] See subsequent discussions regarding actual changes in services due to IT.

[22] See **APPENDIX 2: Use of time based methods for SPPI and resultant volume measures** for an algebraic proof.

[23] I'd like to thank Peter Roemer from DeStatis for this excellent suggestion and resulting discussion, made at the OECD-Eurostat workshop on SPPI's, October 16-19 Luxembourg, 2006.

[24] Some care must be taken here to avoid circularity: for example, avoiding the case where SPPI's are used in the determination of productivity changes which are then used in the calculation of SPPI's.

[25] Note that pricing methods other than those based on working time may also need to be applied with caution in the presence of subcontracting; see the discussion in section 8.

prices are caused by compositional shifts in the amount of work subcontracted or whether they reflect a real price change for the service being approximated. Consequently, the use of time-based methods where subcontracting is employed is more problematic than in other circumstances, and may be a less preferred option because of these issues. In any event, when considering a wage-rate method, "revenues should be accrual and **net of any subcontract costs** to provide comparable prices for consecutive periods[26]". An excellent example of the steps needed to consider when employing time-based methods in the presence of subcontracting is provided in Gac *et al* (2008).

These techniques do not entirely eliminate the risk of bias arising from the use of time based methods (that is, the bias of equating billable hours with service provision). Indeed, survey instruments designed to detect changes in productivity themselves increase respondent burden and compliance costs, and become "yet another ball in the balancing act". However their proposed use attempts to mitigate the chief concerns associated with the use of time-based methods, and at the very least provides indicators as to when such biases are present.

7   Impact of changes in the production function

Voorburg Group discussions, particularly with respect to business services, frequently engage a theme of the impact of changing production functions (especially changing technology). The introduction of new technology is frequently addressed as a change in the service production function. Assuming that providers of services are efficient users of inputs (for example, labour and technology), the change in technology should show an increase in output for a fixed level of labour. Therefore any derived volume measure should show an increase. It is hence imperative that any SPPI (measuring price of the final service to constant quality) should show a decrease.

However the introduction of changes to the production function may also change the characteristics of the final product. Even discounting services where timing of service delivery is a paramount characteristic of utility, new technologies (and other changes) allow more complex or more comprehensive services to be delivered from the same or even lower levels of inputs. This results in efficiencies for the service provider and also results in an improved product for the purchaser.

The key challenges for national statistical offices are then determining when service production has changed, and additionally ascertaining whether the final product (service) is different in any way as a consequence of this changed production function. This latter requirement will necessarily mean working closely with industry experts (especially respondents) to assist in identifying changes in services delivered, and then assisting the quantification of this change in quality

For example, consider the introduction of technological change in the management consultancy industry. Such changes may include software or other analysis tools, and allow consultancy services to be undertaken to the same quality but with less labour. Such changes may also allow simultaneous, concurrent or integrated reorganisation of different aspects of a client's key business, which may not have been possible prior to the introduction of the new technology. In this case the service being delivered has changed.

Continuing this example, since the service being delivered is complex it is appropriate that price changes are measured via model pricing. The challenge for the national statistical office is then to ensure that technological changes are reported during the regular collection cycle, and that the model is updated to reflect both the change in the production function (how it is produced) and the change in output (what is being delivered). The practical downside is that meeting this challenge introduces additional respondent burden and additional cost, and in some instances requires considerable expertise with respect to the service products under consideration.

---

[26] OECD-Eurostat (2005), page 45

8   Bundled Services and Multi-modal solutions

The seemingly innocuous activity of transporting freight illustrates an issue that has occurred in discussion of all types of service industries: where are the boundaries for scope, and when are activities included, excluded or measured elsewhere? These questions arise from two distinct situations.

On the one hand, service providers are increasingly adopting a pricing mechanism known as *bundling*, where a combination of many service types (that is, distinct products) is included on one bill of service. The issue for national statistical office is that whilst the fees levied per bill are frequently readily available from respondents, the composition of services varies from customer to customer, and from period to period. Any change in average "price per bill" is then in part due to pure price change and also due in part to change in both customer mix and service mixes for individual customers. Examples have historically included services such as telecommunications, but the transport and storage industries are seeing "complete logistic solutions" which integrate storage, freight, freight forwarding and even courier services into one combined service.

When presented with a service where the pricing mechanism combines a number of discrete services, the appropriate mechanism to employ for development of SPPI's is component pricing. In the component pricing method[27], a technique similar to model pricing, a service product or a base model of the service is agreed upon with the surveyed establishment[28]. The key feature of model selection in this instance is that the resulting service is comprised of discrete elements for which actual prices can be determined. Continuing the examples above, this could be fixed numbers and types of telephone calls in the case of telecommunications, or specific combinations of storage and road freight for the combined "logistic solutions". Under such an approach the resulting service may itself be fictitious (in the sense that no specific customer happens to have such a combination of discrete elements). The statistical office then combines this information to compute a price index for the composite service.

However, the approach recommended above is susceptible to potential flaws. In the example of telecommunications, the aggregate of individually provided services provides a price of a telecommunications product. However the aggregate estimate in the "logistic solutions" example has services that are different types of products – that is, the bundle of services covers different types of service products. Following the "logistic solution" example, this would involve aggregation of measures for road freight, for storage and for courier services. The price for the bundle does not represent any one type of service product (and, in this particular example, the bundle indeed cuts across different industries). So where component cost can provide a solution to the issue of bundling for some services, for others a modified approach is required.

It is not immediately clear how to proceed in these cases, although one possible suggestion that warrants further attention ("disaggregated cost component") is presented below:

If we consider the elements of the component cost pricing method, we see that key features are:
- elements of an individual service type that can be individually priced each period using real transaction prices;
- weighting data that allows the priced elements to be combined into aggregates for the service; and
- (implicitly) higher level weighting data that allows results from individual establishments to be recombined to produce broader measures of price change for the service product.

---

[27] This section draws heavily from OECD-Eurostat (2005), 2.6

[28] In component pricing, prices of sub-components should be based on prices of truly transacted services. In this respect the method differs from model pricing where sub-component prices are largely estimated.

It may then be possible to utilise the sub-establishment weights inherent within the component cost model to combine service components along product lines within an establishment, and then rather than combining the products *within* the establishment (in this example, storage and road freight), combine individual components *across* establishments. So for example, the SPPI survey is conducted so that elements of the "logistic solutions" are priced individually as via component cost, but road freight measures are combined with road freight measures from other establishments, rather than within the "logistic solutions" establishment. This approach would continue to require identification of elements that can be repeatedly priced from period to period, and would further require derivation and maintenance of detailed sets of weighting data.

A variation on this approach might be to consider a full model price approach (where prices for the elements are estimated rather than observed directly), and then combine these prices of the elements not within the models for each establishment but, again, across establishments. This would then necessitate the mapping of each model element to an appropriate service type.

Further work is needed with regards to both the problem of bundling across service types, and possible solutions.

In addition to the bundling issue, a second situation arises when the delivery of one service itself requires the use of another service. Examples include accountancy services which provide services of a management consultancy nature, management consultancy services which in turn require the development of software (that is, software consultancy), and road freight services that use sea, rail or air transport as part of the service provision mechanism[29].

The determination of the appropriate treatment of such services can be resolved by reference to SNA93 (intermediate consumption):

> 6.152 When goods or services produced within the same establishment are fed back as inputs into the production within the same establishment, they are not recorded as part of the intermediate consumption or the output of that establishment. On the other hand, deliveries of goods and services between different establishments belonging to the same enterprise are recorded as outputs by the producing establishments and must, therefore, be recorded as intermediate inputs by the receiving establishments.

That is, in the system of national accounts, services provided outside an establishment but used by that establishment are measured as intermediate consumption of that establishment. For SPPI's the implication is that, for example, a road freight service that uses a ferry is still considered a road freight service, provided that the ferry service is not produced on own-account within the same establishment[30].

Much discussion on this topic has centred on "double counting". In the road freight-ferry example, the typical discussions might note "the counting of a price movement of road freight that is (in essence) driven by price movements in sea freight will result in aggregate measurement of transport services where price movements are double counted; hence the price index is upwardly biased and the resulting volume estimate (GDP) is downwardly biased".

---

[29] The distinction between the two types of examples is important. On the one hand a "logistics company" may offer storage and delivery as a combined service – this is the "bundled product" type example, where the services are distinct and are combined only as part of the *pricing mechanism*. The second example is where one service is consumed as part of *service delivery* – such as road freight requiring the use of a ferry.

[30] That is, if the statistical units are defined according to the discussion in Section 4.

This line of argument perhaps has some merit when considering inflationary analysis, and is a concern typical of all gross sector price indexes. One solution to this dilemma is to construct aggregate indexes on a net sector basis[31]; that is, continuing the road freight-ferry example, any aggregate measure of change in output prices such as "transport services" excludes those transactions that occur intra-sector (in this case, specifically *excluding* the ferry service as purchased by the road freight establishment, but including the road freight service).

However, if the national accounts are produced on a supply-use framework[32], it is vital that national accountants are able to get measures of price change for the delivery of both road freight services and sea freight services, since the sea freight service (in this example, the ferry) is an intermediate input into road freight. That is, when considering detailed product level indexes, double counting is not a concern and instead the SPPI should measure all outputs, including those transactions used as intermediate consumption elsewhere. The caveat that applies here is to differentiate as to when the service provided is intermediate consumption, and when it is produced on own-account.

9   Summary

A range of SPPI themes recurring at Voorburg Group discussions has been discussed. The key points raised were:

a) SPPI's for the purpose of measuring service sector GDP should be produced on a product basis and not an industry basis; use of industry SPPI's in place of products SPPI's is only appropriate under strict assumptions regarding levels of secondary production, horizontal and vertical diversification, production functions, market operations and price discrimination.

b) Scope of SPPI's should be all output of businesses, both to domestic and non-resident purchasers; the definition of an export of service should be determined by the residency of the purchaser and not necessarily by where the service takes place.

c) The use of time based methods for SPPI's generally results in a biased measure of price change and consequently a biased measure of GDP. Explicitly, the resulting measure of GDP is equivalent to a measure of "hours worked", and fails to capture changes due to labour productivity.

d) This bias can be in some part mitigated through use of realised rates of revenue, collection of data in "bands of expertise" where the activities of such bands are closely monitored, and through use of mechanisms explicitly designed to measure changes in productivity.

e) Care also needs to be taken when applying time based methods in the presence of subcontracting (and outsourcing).

f) When new technology changes the production function for a service, an assessment needs to be made as to whether the actual service has changed.

g) The issue of bundling of services (a pricing mechanism) can be resolved through the component cost approach. However, this technique is limited when the "bundle" cuts across service types. Further work is required in this area.

Services consumed in the delivery of other services (intermediate consumption) need to be priced in their respective product level price indexes, with care taken to remove those instances where such services are produced on own-account.

---

[31] This is the approach adopted by the Australian Bureau of Statistics (ABS) for the manufacturing sector PPI's

[32] See discussion in Section 4

References

Berger M (2007), *Addressing cross-cutting issues arising in the development of SPPI's for use in the measurement of Service Sector GDP,* Australian Bureau of Statistics (ABS), 22nd Voorburg Group meeting

Gac D, Nieminen S, Norrman V, Varjonen S (2008) *Average daily charge-out rates in IT in the context of offshore outsourcing*, to be presented at 23nd Voorburg Group meeting

OECD-Eurostat (2005), *Methodological Guide For Developing Producer Price Indices For Services*

UN et al (1993) [Inter-Secretariat Working Group on National Accounts: Eurostat, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank], *System of National Accounts 1993*

UN et al (2002) [United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations Conference on Trade and Development, World Trade Organization], *Manual on Statistics of International Trade in Services*

APPENDIX 1:   Key changes introduced in 2008 revision

- Changed method of citation throughout and added references at the end

- Product versus industry:

    o Emphasised that different destinations means different products and cited SNA

    o Included discussion regarding infrastructure – industry based SPPI's arise as a consequence of industry turnover activities

    o Added "constant quality" requirement for using industry PPI's as proxy for product SPPI's

- Exports:

    o Indicated that description of international trade was a simplification, indicated where simplifications had been made, and added appropriate references

    o Added quality dimension to the discussion regarding exports

- Time based methods:

    o Changed an example from architectural services to legal service

    o Suggested possible indicators that respondent burden is impacting measurement

    o Added exception to "time based measures give bias" discussion (CPC 8512, Personnel Services)

    o Emphasised key issue of consistency between prices and labour inputs

    o Emphasised "time based measures can result in GDP bias" message

    o Clarified discussion on subcontracting to use the same terminology as in OECD-Eurostat (2005)

    o Cited Gac et al (2008) as good example of how to approach subcontracting

- Change in production function

    o Changed discussion from focus on IT changes to all types of changes to production function (with IT as an example)

    o Included cost implications arising from the need to detect and assess changes in production functions

APPENDIX 2: Use of time based methods for SPPI and resultant volume measures

Consider a typical time based method where charge-out rates are calculated as average revenue or turnover per hour of labour. If a price index is constructed from these data, changes in the price index are simply measuring changes in hourly rates. If these price changes in turn are used in deflation of national accounts value aggregates, the resulting volume (and volume change) measures are only measuring changes in hours worked. These volume measures do not consider other inputs, margins, or changes in productivity (labour or otherwise).

$$P = \frac{Turnover}{Hours}$$

$$V = Turnover$$

$$V = P \times Q$$

$$Q = \frac{V}{P}$$

$$= \frac{Turnover}{\left(\dfrac{Turnover}{Hours}\right)}$$

$$= Hours$$

$$\Delta Q = \Delta Hours$$