Annex 4

Spatial comparisons of consumer prices, purchasing power parities and the International Comparison Program

1 Introduction

This annex deals with the problem of comparing price levels across different areas or regions within a country, as well as across countries. Even though international price comparisons are required to handle differences in currencies in different countries, the index number problems involved in price comparisons across countries mirror those encountered in comparisons over time. There is a large body of literature on cross-country comparisons of prices and real income undertaken under the auspices of the International Comparison Program (ICP). While not providing an exhaustive account of the related problems and relevant aggregation methods, this annex aims to achieve a degree of completeness in the coverage of the problem of consumer price comparisons in the manual by adding the spatial and international dimensions to the temporal comparisons dealt with in various chapters of the manual. The annex also attempts to identify possible avenues for a closer integration between spatial and temporal comparisons of consumer prices.

The main objectives of the annex are: (i) to provide a brief summary of the index number problems encountered in the process of international and inter-area price comparisons and to highlight the need for the development and use of specialized aggregation methods; (ii) to describe a few aggregation methods used in deriving purchasing power parities (PPPs) and spatial measures of price levels; (iii) to examine the relationship between the ICP and PPPs for cross-country comparisons with the CPI; and (iv) to explore the feasibility of integrating the ICP activities with the streamlined activities of national statistical offices for the compilation of the CPI.

The annex is also designed to provide an introduction for the statisticians in various national statistical offices who may currently be involved in consumer price index (CPI) compilation to the issues and methods involved in spatial comparisons of consumer prices. The annex outlines some of the principal differences in the approaches to spatial comparisons. Countries embarking on inter-area or regional consumer price comparisons, as well as those countries that may participate in the ICP in the near future, may find the contents of the annex useful.

2 Differences between temporal and spatial comparisons

There are several major qualitative differences in the nature of price comparisons involved in the standard CPI comparisons over time and price comparisons over space involving regions or countries. These differences highlight the need for specialized methods for aggregating price data in deriving summary measures of price levels, as well as specific types of data requirements associated with cross-country and inter-area comparisons.

The foremost difference is the absence of a natural ordering of price and quantity observations in the context of cross-country or inter-area comparisons. The CPI framework and methods are devised to measure changes over time. Therefore, the price observations appear in a chronological order. The presence of a natural ordering of price observations makes it possible to examine the feasibility and relative merits of the fixed and chain index numbers. For example, in the context of constructing price comparisons across countries within the OECD, or across states within the United States, it is impossible to arrive at an ordering which facilitates chained comparisons.

The multilateral nature of spatial comparisons is a distinguishing feature of price comparisons across regions and countries. When price levels of goods and services across different countries are compared, it is essential that such comparisons are undertaken for every pair of regions being considered. If the World Bank is interested in comparisons of real income in different countries, it is necessary for the Bank to be able to make comparisons between all pairs of countries involved. This multilateral nature of comparisons creates several problems. First, the number of comparisons (one for each pair) can be quite large, and presentation and use of such results may be quite unwieldy. For example, if a particular comparison exercise involves 20 countries, then it requires 190 ($20 \times 19/2$) separate binary comparisons involving distinct pairs of countries. Second, results from such a large tableau of binary comparisons require a degree of consistency. This requirement translates into the ''transitivity'' condition described below.

The uses and applications of inter-area price comparisons may differ significantly from general consumer price indices. The CPI is probably the most significant economic statistic produced in any country. It is not only used as a general measure of price changes over time, but often in assessing and calibrating monetary policy. Despite the conceptual similarities in price comparisons over time and across space, spatial price comparisons are useful in comparisons of standards of living and well-being in different regions of a country or across countries. Such comparisons are essential in assessing development and in ensuring more balanced growth in different regions. There is considerable demand for measures of CPI across different cities and different states and regions (rural versus urban) within countries. However, there are very few countries where inter-area price level comparisons are readily available. Kokoski et al. (1999) demonstrate the feasibility of deriving meaningful inter-area price comparisons within the United States using the price data collected across different states.

International comparisons of prices, in the form of PPPs from the ICP, are used by international organizations and individual researchers in assessing growth and productivity performance of countries, and also in

making meaningful comparisons of various national income aggregates (including government expenditure) across different countries. Currently, consensus is emerging among researchers and practitioners that price comparisons and PPPs are necessary in assessing the nature and extent of global poverty and its distribution across countries and regions of the world. Several recent research papers by Ward (2001), Prennushi (2001), Astin (2001) and Dwyer et al. (2001), presented at the joint World Bank–OECD Seminar on Purchasing Power Parities, 30 January to 2 February 2001, highlight a number of important applications of PPPs derived from international comparisons of prices under the ICP. Eurostat has recently embarked on a programme to extend the current ICP for its EU Member States to cover regional comparisons within different countries.

In recognition of the major analytical differences between the standard CPI comparisons over time and spatial and cross-country comparisons of consumer prices and PPPs, considerable research efforts have focused on the development of the data and methods necessary for spatial comparisons of prices. A brief summary of the results of this research is given below.

3 Data requirements for spatial comparisons

The basic data requirements for spatial comparisons are very similar to the data required for standard CPI calculation. The main components are the data on prices of a large range of products representative of the consumption baskets of households and information on weights associated with various product categories reflecting the importance attached to different products. Within the CPI, it is common practice to collect price quotations from different outlets scattered throughout the country. The selection of the outlets and areas from which prices are collected is based on complex multi-stage sampling designs. The expenditure weights are based on a classification of goods and services using a standard system such as the Classification of Individual Consumption according to Purpose (COICOP) or a similar national classification. The lowest level of product classification at which expenditure weights are available is used in identifying the elementary indices and higher-level indices at progressively higher levels of aggregation, leading ultimately to the total household expenditure level.

Spatial comparisons pose several problems in terms of identifying products that are to be priced from different areas, regions or countries involved in a comparison exercise. This problem is less severe when fairly similar or homogeneous areas are being compared. In cases where comparisons involve areas that are fairly heterogeneous, two problems arise. The first arises from major differences in the consumption baskets. For example, when comparisons are made between two states, say Minnesota and Florida, there may be major differences in the consumption baskets at the detailed level, even though the major expenditure categories may be identical. This problem is somewhat similar to the treatment of disappearing and new goods in the context of the CPI, but is more serious when cross-country comparisons are being attempted. The second problem arises from major differences in the quality of items. The quality differences may be measured through several product characteristics, with allowances and adjustments being made at the appropriate stage of index number calculation. Kokoski et al. (1999) demonstrate the feasibility of making inter-area price comparisons for heterogeneous goods.

Changes in quality are likely to be more gradual in the case of temporal comparisons, but can be a serious problem when comparisons across countries are attempted. The ICP follows the principle of identity in dealing with the problem of quality differences across countries. A comprehensive list of products with detailed product specifications is developed at the planning stages of any cross-country comparison exercise. These items are priced in different countries from various outlets distributed across the country, a procedure very similar to that used in the CPI. Development of the product listing is, however, a difficult step, with the degree of difficulty depending upon the size and heterogeneity of the group of countries involved. The use of a product listing, based on the identity principle, can have serious implications for the representativeness of the product list of the consumption baskets in different countries. There are several operational procedures used by international organizations in handling these problems relating to the compilation of price data. A more detailed account of the problems and recommended solutions can be found in *ICP Handbook* (United Nations, 1992) and the recent publication by OECD (1999) on its international comparison work.

Once the price data are compiled, the next stage in the CPI compilation is the aggregation of item-specific price changes to measure price movements for various categories of consumption expenditure. At this stage, it is necessary to have information on consumption patterns. This information is usually drawn from household expenditure surveys. These surveys are regularly conducted in most countries by the respective national statistical organizations. For purposes of making inter-area comparisons of consumer prices, the corresponding requirement is the availability of household expenditure survey data specific to each area included in the comparisons. In many cases, for reasons relating to sampling and statistical reliability, detailed expenditure pattern data may not be available for all the regions.

Spatial comparisons of consumer prices pose specific problems because of the non-overlapping nature of the consumption baskets, major differences in the quality of items priced in different regions and countries, and the non-availability of crucial data on region-specific expenditure patterns. These problems require the development of new analytical techniques that can handle major differences in quality. National statistical offices may need additional financial resources in order to provide reliable and meaningful price comparisons between different cities, areas and regions within countries, and to compile reliable data for the more difficult task of inter-country comparisons of prices and real consumption.

4 Aggregation methods for spatial comparisons

This section briefly describes the types of aggregation methods that are commonly used in cross-country comparisons of prices. Since most of these methods have been developed in the context of the ICP, and are equally valid for inter-area or regional comparisons, the discussion below uses countries as spatial entities. This section is further divided into three parts. The first deals with the notation and conceptual framework necessary to deal with multilateral spatial comparisons. The second describes the construction of elementary indices for aggregation of prices when no quantity or expenditure information is available. Finally, a small selection of index number methods used in spatial price comparisons are presented.

4.1 Notation and conceptual framework

Consider the case involving comparisons across M countries, and price and quantity data on N commodities. These commodities refer to goods and services that are priced in all the countries. If the commodities refer to items below the elementary level at which no quantity or expenditure share data are available, we make use of only the price data. At this stage, all the problems relating to non-overlapping commodity lists and existence of quality differences are set aside so that the main focus is just on the aggregation issues. Let $p^j = [p_1^j, \ldots p_N^j]$ and $q^j = [q_1^j, \ldots q_N^j]$ represent the price and quantity vectors from country j ($j = 1, 2, \ldots, M$). In the case of international comparisons, all the prices are expressed in the respective national currency units. As in the case of the CPI computation, the problem is one of decomposing the differences in the value aggregates

$$V^j = \sum_{i=1}^{N} p_i^j q_i^j \tag{A4.1}$$

into measures of price and real expenditure components.

Since there are M sets of price and quantity vectors and, therefore, $M(M-1)/2$ binary comparisons between all distinct pairs of countries, a simpler notation is used in this annex in the place of the notation generally used in the manual. Let I_{jk} denote the (consumer) price index number for country k with country j as the base. If j and k are, respectively, the United States and India, and if $I_{jk} = 22.50$, then the index is interpreted to mean that 22.50 Indian rupees have the same purchasing power as one US dollar for the goods and services covered in computing the index. Thus the index can also be interpreted as the PPP between the currencies of j and k. This interpretation is consistent with the meaning accorded to the CPI. Since currency denominations are involved here, a proper measure of relative price level differences can be obtained if the PPP is compared to the exchange rate prevailing at the time when comparisons are made.

Because of the multilateral nature of spatial comparisons, when M countries are involved, it is necessary to provide comparisons between all pairs of countries. Thus, it becomes necessary to compute each and every entry in the following matrix of binary comparisons:

$$I = \begin{bmatrix} I_{11} & I_{12} & I_{1k} & I_{1M} \\ I_{21} & I_{22} & I_{2k} & I_{2M} \\ I_{j1} & I_{j2} & I_{jk} & I_{jM} \\ I_{M1} & I_{M2} & I_{Mk} & I_{MM} \end{bmatrix} \tag{A4.2}$$

Several points concerning the matrix, I, are worth noting. First, the matrix can be large if the number of countries (or regions) involved is large. Second, the results recorded in the matrix need to be internally consistent. All the index number issues and various approaches discussed in the manual apply directly to each binary comparison involving two countries. Diewert (1986, 1999b) provides a summary of the microeconomic theoretical and test approaches to cross-country comparisons. Thus, it is possible to apply Fisher, Törnqvist, Walsh or other index number formulae described in the manual.

In order to ensure meaningful interpretation of the results from multilateral cross-country comparisons, the index number methods applied need to satisfy a number of basic requirements, only the most important of which are discussed below. Kravis et al. (1982), OECD (1999) and United Nations (1992) provide a complete list of these requirements.

Transitivity. An index number formula I_{jk} is said to satisfy the transitivity property if and only if for all choices of j, k and ℓ (j, k, $\ell = 1, 2, \ldots, M$), the index satisfies

$$I_{jk} = I_{jl} \times I_{lk} \tag{A4.3}$$

Equation (A4.3) requires that the application of a formula to make a direct comparison, I_{jk}, should result in the same numerical measure as an indirect comparison between j and k through a link country ℓ. Note that the transitivity property ensures internal consistency of index numbers in the matrix given in equation (A4.2). It guarantees that the PPP for two currencies, say A and B, is the same whether it is derived through a direct comparison of A and B or through an indirect comparison that compares A with C and C with B, which are then combined to provide an indirect PPP for A and B. This requirement arises mainly from the spatial nature of the comparisons where no natural ordering of the countries involved could be imposed without a value judgement. Most of the commonly used index number procedures do not satisfy this requirement. The following result is useful in constructing transitive index numbers.

An index number formula I_{jk} satisfies the transitivity property in (A4.3) if and only if there exist M positive real numbers $\lambda_1, \lambda_2, \ldots, \lambda_M$, such that

$$I_{jk} = \frac{\lambda_k}{\lambda_j} \tag{A4.4}$$

for all j and k.

The proof of this result is straightforward (Rao and Banerjee, 1984). The result is important since it shows that, when the transitivity property is satisfied, all that is necessary is to measure M real numbers $\lambda_1, \lambda_2, \ldots, \lambda_M$, and then all the necessary indices in (A4.2) can be calculated using these M numbers, thus reducing the dimension of the problem involved. Two important

points may be noted. First, the numbers λ_j in equation (A4.4) are not unique, since any scalar multiplication of a vector of λ_j can also lead to the same matrix of index numbers as that derived from the original λ_j. Therefore, these λ_j need to be determined (in any empirical exercise) up to a factor of proportionality. Second, these λ_j can be interpreted as the PPP of currencies involved. This particular result formed the basis of the work of statisticians such as Geary (1958) and Khamis (1970), who proposed aggregation methods designed to compute PPPs directly from the price and quantity data without invoking the index number literature.

Base invariance. An index number formula is said to be base invariant if a comparison between a given pair of countries (j, k) is invariant to the order in which the countries are listed. This implies that multilateral comparisons should be invariant to all possible permutations of the data set. For example, consider a set of transitive comparisons derived using a particular country (say the United States) as a star country. Under this scheme, price comparison between any pair of countries, say A and B, is effected through the United States which serves as a link country. Therefore,

$$PPP_{A,B} = PPP_{A,USA} \times PPP_{USA,B}$$

This scheme is inadmissible under the base invariance criterion since the choice of the star country clearly affects the PPP of currencies of countries A and B. Further, the United States is accorded a special status, in the form of a link country, in deriving transitive multilateral comparisons.

Characteristicity. This is a requirement outlined in Drechsler (1973). This property requires that any set of multilateral comparisons satisfying the transitivity property should retain the essential features of the binary comparisons constructed without the transitivity requirement. Since condition (A4.3) implies that a transitive comparison between a pair of countries j and k is necessarily influenced by the price and quantity data for all the other countries, the characteristicity property requires that distortions resulting from adherence to the transitivity property should be kept at a minimum. Balk (2001) shows that a complete adherence to the characteristicity principle in its extreme (complete preservation of all binary comparisons) would imply that price indices, and hence PPPs, cannot depend upon any quantity or expenditure share weights. This is an extreme result, which is to be avoided in all index number comparisons. The Elteto–Koves–Szulc (EKS) method for multilateral comparisons, discussed below, has its origins in the characteristicity property.

4.2 Index number methods for spatial comparisons

Spatial price comparisons in general, and international comparisons in particular, use index number methods for aggregating price and quantity data at two different levels. The first is the basic heading level. This is normally the lowest level of aggregation at which expenditure data and weights are available. These basic headings usually consist of a fairly homogeneous group of items that are priced in different outlets in the countries. The subsequent levels of aggregation lead to indices for broad expenditure categories, and finally to the whole consumption basket.

4.2.1 Aggregation below the basic heading level

Two commonly used index number methods are described below. These procedures explicitly allow for the possibility that price data may not be available for all items in the product list constructed for a given international comparison exercise. Such a situation is possible in the case of temporal comparisons, but is usually limited to a small number of commodities that are either disappearing or new goods.

The Elteto–Koves–Szulc (EKS) Method. A variant of the original method proposed in Elteto and Koves (1964) and Szulc (1964) is generally used in aggregating price data below the basic heading level. The EKS method involves two stages. In the first stage, binary comparisons are constructed using price relatives of those commodities for which prices are available in both countries. If n_{jk} is the number of commodities that are priced in both countries, then the current practice within ICP constructs a binary elementary index using the following formula:

$$I_{jk} = \prod_{i=1}^{n_{jk}} \left[\frac{p_i^k}{p_i^j} \right]^{1/n_{jk}} \tag{A4.5}$$

Obviously, these indices are not transitive, since each index is based on prices of a different set of commodities. The EKS procedure is then used in deriving a transitive set of indices. The resulting formula for the construction of elementary indices for spatial comparisons is given by

$$I_{jk}^{EKS} = \prod_{\ell=1}^{M} \left[I_{j\ell} I_{\ell k} \right]^{1/M} \tag{A4.6}$$

The elementary index number formula in (A4.5) is similar to the formula used in the construction of the CPI. The principal difference results from the fact that not all commodities are priced in all countries and that there is a need for transitivity at all stages of aggregation. The properties of these indices are discussed in Chapter 20 of the manual.

The OECD (1999) uses a slightly modified variant of the binary indices shown in equation (A4.5). A formula that mimics the standard Fisher index is used, but without the use of any expenditure shares since the aggregation is below the basic heading level. This procedure tries to account for the fact that not all commodities for which prices are collected are really characteristic or important in one or both of the countries. The procedure takes explicit account of those commodities which are starred, indicating that the item is important in a given country. The modified EKS method uses the same formula as above, but the binary index on the right-hand side is replaced by:

$$I_{js} = \left\{ \prod_{i \in M(s)} \left[\frac{p_i^s}{p_i^j}\right]^{\frac{1}{n(s)}} \prod_{i \in M(j)} \left[\frac{p_i^s}{p_i^j}\right]^{\frac{1}{n(j)}} \right\}^{\frac{1}{2}} \qquad \text{(A4.7)}$$

where $n(s)$ and $n(j)$ are, respectively, the number of starred items in countries s and j; $M(s)$ and $M(j)$ are, respectively, the sets of commodities that are starred (considered representative) in different countries.

Use of equations (A4.5) and (A4.6) for the construction of spatial CPI numbers at the basic heading level has its problems. The most important problem is that these formulae do not take into consideration whether or not the commodities priced in different countries are "representative" of the consumption in different countries within the basic heading. A related problem is whether or not the coverage of the commodities, priced with respect to the basic heading to which they belong, is adequate. These issues are currently being researched, and Rao (2001b) offers a modified approach that attaches weights proportional to coverage and representativeness.

While there has been much research on the properties of index number formulae for the construction of elementary indices within the CPI framework (Diewert (1995a), Dalén (1992) and Turvey (1996)), there has been very little research on the properties of elementary indices within the context of international comparisons.

The Country–Product–Dummy (CPD) Method. The CPD method was originally proposed by Summers (1973) as a tool to deal with missing price observations. The method is a simple statistical device that can be used in deriving the PPPs for a particular basic heading by simply regressing the logarithm of observed prices against a set of dummy variables, defined with respect to commodities and countries. Thus the procedure involves the model:

$$\ln p_i^j = \eta_1 D_1 + \eta_2 D_2 + \ldots + \eta_n D_n + \pi_1 D_1^* + \pi_2 D_2^* + \ldots + \pi_M D_M + u_i^j \qquad \text{(A4.8)}$$

where D_i $(I = 1,2,\ldots,n)$ and $D_j^* (j = 1, 2, \ldots, M)$ are, respectively, dummy variables for the N commodities in the basic heading and M countries involved in the comparisons.

Once this regression equation is estimated, the PPP for currency of country k with country j as base can be obtained by

$$\text{PPP}^j = \exp(\hat{\pi}_j) \qquad \text{(A4.9)}$$

where $\hat{\pi}_j$ is the estimator of π_j in equation (A4.8). Then the desired index at the basic heading level is given by

$$I_{jk} = \frac{\text{PPP}_k}{\text{PPP}_j}. \qquad \text{(A4.10)}$$

The exponential of the difference in the estimates of π_j and π_k obtained from the regression equation provides the necessary index number.

The CPD model offers a number of generalizations that can explicitly account for a number of data-related problems. The model can be easily generalized to account for differences in quality measured through a set of product characteristics. The feasibility of this approach to inter-area comparisons of consumer prices within the United States was demonstrated in Kokoski et al. (1999). Rao and Timmer (2000) examined the feasibility of using a generalized CPD model to incorporate various measures of reliability in the context of aggregating unit value ratios to provide comparisons at the manufacturing branch level. Rao (2001b) discusses a few model specifications that are appropriate for aggregation below the basic heading level within the context of the ICP.

The EKS method, described in equations (A4.6)–(A4.8), is the procedure that is currently used by all the international organizations for purposes of aggregation below the basic heading level. The CPD and EKS methods yield identical basic heading parities when all the commodities are priced in all the countries. Ferrari and Riani (1998) and Ferrari et al. (1996) present a number of analytical results relating to these methods.

While the sampling issues relating to the selection and distribution of outlets and the frequency of price quotations are considered important for the construction of the elementary indices within the CPI, issues that are crucial for the construction of PPPs below the basic heading for inter-country comparisons are quite different. Issues of quality differences and non-availability of goods and services in all the comparison countries are far more important in spatial comparisons of consumer prices.

4.2.2 Aggregation above the basic heading level

This section presents a small selection of the range of aggregation methods used in the context of spatial comparisons. A more comprehensive analysis of the spatial aggregation methods developed over the past three decades is presented by Balk (2001).

This level of aggregation is similar to the stage where elementary indices are aggregated to derive the overall CPI. In the case of temporal comparisons involving two time periods, all the methods and approaches described in the manual are appropriate, and in most cases the national statistical offices use the Laspeyres, Fisher or some variants of these formulae for CPI construction. However, the multilateral nature of spatial comparisons necessitates slightly different approaches to their construction.

A number of index number methods for aggregation above the basic heading level have been developed over the past three decades but, in the interest of brevity, only the principal methods are discussed below. These are the Geary–Khamis and EKS methods for international comparisons, the principal aggregation methods used in various international comparison exercises by the ICP, the OECD, Eurostat and FAO.

Several approaches to the construction of multilateral index numbers that satisfy transitivity and base invariance properties are discussed below. Four distinct approaches emerged during ICP work carried out between 1970 and the early 2000s. The first and most straightforward is the EKS approach, which uses binary results as building-blocks for multilateral comparisons. The second is the Geary–Khamis approach, which

provides a methodology for computing PPPs of currencies and international average prices of commodities, using the price–quantity data at the basic heading level. The third is the stochastic approach based on the CPD method and its generalizations that can be used in econometrically estimating the PPPs using a regression framework. The fourth and the last approach discussed here is the linking approach to the construction of chained comparisons based on the concept of the minimum spanning tree. This is generating considerable interest and is explored further in the last section of this annex. These four approaches are by no means exhaustive, but they represent major strands of research and development in this area.

The EKS Method. The EKS system is a simple method of generating transitive multilateral index numbers from a system of binary index numbers, with the property that the resulting multilateral indices deviate the least (according to a specific criterion) from the binary indices. Since the seminal paper by Drechsler (1973), it has been well recognized that (transitive) multilateral systems necessarily deviate from their binary counterparts and therefore result in a loss of "characteristicity". The EKS system is designed to minimize such loss of characteristicity. The original EKS system uses the Fisher binary indices, but the work of Caves, Christensen and Diewert (1982b) and Rao and Banerjee (1984) recognizes that other binary indices could be used in conjunction with the EKS technique. For any pair of countries j and k, if F_{jk} represents the Fisher binary index, then

$$EKS_{jk} = \prod_{\ell=1}^{M} \left[F_{j\ell}.F_{\ell k}\right]^{1/M} \qquad \text{(A4.11)}$$

provides the EKS index.

There are several notable features of the EKS technique. First, it is based on the premise that direct binary comparisons, derived using any chosen formula, provide the best comparison between pairs of countries. Second, even though the EKS index in equation (A4.11) is defined using the Fisher index, this approach can be applied in conjunction with any other index number formula. For example, the Fisher index in equation (A4.11) may be replaced by another superlative index, such as the Törnqvist index. Caves, Christensen and Diewert (1982b) suggest the use of a Törnqvist-based EKS formula for spatial comparisons. Third, the EKS index in equation (A4.11) is the multilateral index that deviates the least from the matrix of non-transitive binary indices, when the deviations are measured using a logarithmic distance function. Finally, the EKS index can be interpreted as a simple geometric mean of all indirect comparisons between j and k through all possible link countries.

The simple unweighted nature of the EKS index has attracted attention in recent years. Since different binary comparisons have different levels of reliability, measured using various criteria, it is necessary to reflect these differences in defining weighted EKS index numbers. Rao and Timmer (2000), Rao et al. (2000) and Rao (2001b) provide illustrations of how weighted EKS indices can be generated in order to account for various data-related problems.

The Geary–Khamis (GK) Method. The GK method was originally proposed by Geary (1958) and subsequently developed by Khamis (1970, 1972 and 1984). The GK method has been the principal aggregation method in most of the ICP phases to date. Since 1996, the OECD has produced and published international comparisons based on both the EKS and GK methods.

The GK method provides a way of calculating PPPs of currencies of different countries from the observed price and quantity data (applied at the basic heading level). The concept of PPP is applicable even when the currency unit is the same in several areas of a country. The GK method simultaneously determines international average prices of different countries. Let P_i denote the international average price of ith commodity. The GK method is defined through the following system of interrelated equations, defined for each country j and each commodity i,

$$P_i = \frac{\sum_{j=1}^{M} p_i^j q_i^j / PPP_j}{\sum_{j=1}^{M} q_i^j} \qquad \text{and} \qquad PPP_j = \frac{\sum_{i=1}^{N} p_i^j q_i^j}{\sum_{i=1}^{N} P_i q_i^j} \qquad \text{(A4.12)}$$

These simultaneous equations are then solved to yield numerical values of PPPs and Ps, after selecting one of the currencies as a numeraire. Once the PPPs are solved, the spatial price index numbers are simply defined as

$$I_{jk} = \frac{PPP_k}{PPP_j} \qquad \text{(A4.13)}$$

One of the main reasons for the continued use of the GK method is "additivity". Additivity requires that aggregates, such as real domestic product, derived by converting national aggregates using PPPs, should be equal to aggregates derived through valuation of quantities at international prices. Thus additivity requires

$$\sum_{i=1}^{n} p_i^j q_i^j / PPP_j = \sum_{i=1}^{n} P_i q_i^j \qquad \text{(A4.14)}$$

This requirement is satisfied automatically by the PPPs and Ps derived from the GK system defined in equation (A4.12). The GK system is also useful in analysing the structure of real GDP and shares of different components across different countries. This system provides a framework within which internationally comparable national accounts could be constructed. However, the GK system is not rooted in standard economic theory and fails several test properties (Diewert, 1986). There has been considerable debate among practitioners concerning the average prices resulting from the GK system. The system has the potential to reflect the price structure of the richer countries, and therefore has the tendency to overstate the real income of the poorer countries.

Weighted Country–Product–Dummy (CPD) Method. It is possible to generalize the CPD method discussed in the context of aggregation below the basic heading level. Rao (1995) has generalized the CPD method by incorporating quantity and value data directly into the CPD method described in equation (A4.8). The basic idea behind this generalization comes from the fact that the standard CPD regression model attempts to track the logarithm of the observed prices using an unweighted residual sum of squares. In the spirit of the standard index number approach, where price index numbers are required to track price changes of more important commodities more closely, a more appropriate procedure would be to find estimates of the parameters that are likely to track important commodities more closely. This is achieved by minimizing a weighted residual sum of squares, with each observation weighted according to the expenditure share of the commodity in a given country. Thus the generalized CPD method suggests that estimation of the equation

$$\ln p_{ij} = \pi_1 D_1 + \pi_2 D_2 + \ldots + \pi_M D_M + \eta_1 D_1^* + \eta_2 D_2^* + \ldots + \eta_n D_n^* + u_{ij} \qquad \text{(A4.15)}$$

be conducted after weighting each observation according to its value share. This is equivalent to the application of ordinary least squares to the following transformed equation obtained by premultiplying equation (A4.15) by $\sqrt{w_{ij}}$. The resulting equation is:

$$\sqrt{w_{ij}} \ln p_{ij} = \pi_1 \sqrt{w_{ij}} D_1 + \pi_2 \sqrt{w_{ij}} D_2 + \ldots + \pi_M \sqrt{w_{ij}} D_M + \eta_1 \sqrt{w_{ij}} D_1^* + \ldots + \eta_n \sqrt{w_{ij}} D_n^* + v_{ij} \qquad \text{(A4.16)}$$

where $w_{ij} = p_{ij}q_{ij} / \sum_{i=1}^{N} p_{ij}q_{ij}$ is the value share of the ith basic heading in the jth country.

Rao (1995) has shown that the international prices and PPPs resulting from the estimates of parameters in equation (A4.13) are identical to those derived using the Rao method for international comparisons, described in Rao (1990). Thus the weighted CPD method may be considered as a bridge between the GK approach to international comparisons and the standard stochastic approach to index numbers.

Kokoski et al. (1999) outline a procedure which makes use of the CPD method to adjust for differences in quality characteristics of various consumer items for the construction of inter-area CPIs using United States data. Estimates derived from the CPD method are subsequently used in the application of the multilateral Törnqvist index derived after imposing the transitivity condition. Hence the approach used in Kokoski et al. (1999) may be described as a mixed stochastic and index number approach to multilateral spatial comparisons of CPI numbers.

Spatial linking and chaining approach. In recent years, a new approach to inter-area and inter-country comparisons of prices has been given serious consideration. This approach advocates spatial chaining of binary comparisons where links are identified using a procedure based on a measure of distance or reliability of binary comparisons involved. This approach is in sharp contrast to the general approach to multilateral comparisons, where either all the binary comparisons are used, as in the case of the EKS method, or all the price and quantity data are simultaneously used, as in the case of the GK and CPD approaches.

Using the graphical theoretical concept of minimum spanning trees, Hill (1999c, 1999d) proposed a method of deriving a system of transitive multilateral comparisons from a matrix of binary comparisons. The Hill approach is based on the fact that direct binary comparisons may not always be the best.

For any pair of countries j and k, Hill suggests a measure of distance (indicating the reliability of the binary comparison) using the Laspeyres–Paasche spread defined as

$$D(j,k) = \left| \ln \left(\frac{L(j,k)}{P(j,k)} \right) \right| \qquad \text{(A4.17)}$$

where $L(j, k)$ and $P(j, k)$ are, respectively, binary Laspeyres and Paasche price index numbers. Note that the same distance function emerges if price index numbers are replaced by quantity index numbers. $D(j, k)$ is equal to zero if the price structures or quantity structures are identical in countries, j and k. Thus, this distance function serves as an indicator of similarity of price and quantity structures in these countries.

Using a matrix of distances calculated for all pairs of countries, Hill (1999c, 1999d) suggests that a minimum spanning tree (MST) be extracted and used in constructing chained links between all pairs of countries. The MST has the property that a chained comparison between any pair of countries has the least distance, and therefore can be considered as the most reliable. It also has the property that the sum of the distances between all the links, in the MST, is the least when compared to all possible tree-configurations. For purposes of illustrating the concepts involved, Figure A4.1.1 shows the MST for Europe constructed using ICP data for Europe for the 1985 benchmark year.

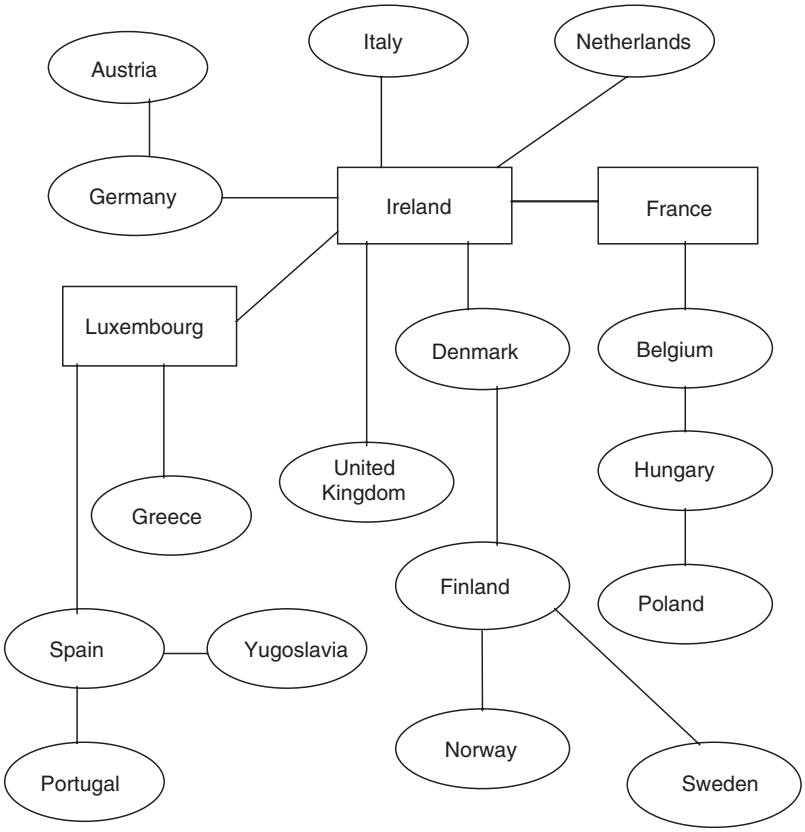
It is evident from the MST presented here that a comparison between Germany and Portugal is through a chain involving Ireland, Luxembourg and Spain. This chained comparison is deemed to be better than a direct comparison between these two countries. There are a number of issues yet to be resolved regarding the use of the MST. Nonetheless, an MST provides a formalization of a somewhat intuitive notion of linking dissimilar countries using a chain of similar countries.

Once the MST is identified, a transitive comparison between a given pair of countries in a particular exercise is constructed using binary indices calculated using a chosen formula, such as the Fisher or Törnqvist index, and the links indicated in the MST. Thus, if a comparison between Sweden and Denmark is needed, then the MST approach suggests the following index for this comparison:

$$I^{MST}_{\text{Sweden, Denmark}} = F_{\text{Sweden, Finland}} \times F_{\text{Finland, Denmark}}$$

where F denotes the Fisher index.

Figure A4.1 A minimum spanning tree: Europe



Since the MST provides a unique chain of links between any two countries, comparisons are uniquely defined. The spanning trees are, however, sensitive to the countries included, and the types of measures used in assessing the degree of reliability or comparability of any two countries. Aten et al. (2001) examine the sensitivity of the spanning trees and the resulting comparisons based on a range of measures including some similarity indices. Rao et al. (2000) applied the spanning tree approach to the construction of multilateral, agricultural input, output and productivity indices, using United States state-level agricultural production data.

5 Integration of the CPI and inter-area and international comparisons

The best available inter-area and international price comparisons for consumer goods and services use data from the national statistical offices compiling CPIs. In recent critical reviews of the ICP by Ryten (1998), and of the OECD–Eurostat PPPs by Castles (1997), the accuracy and reliability of international comparisons have been assessed as being less than adequate. Castles notes the difficulties encountered in comparing like goods from country to country, and Ryten argues for the need to secure greater support from national statistical offices in compiling price data for the ICP. Both reviews recommended examination of the feasibility of integrating ICP work with that of the normal CPI work undertaken by the national statistical offices. Since this annex deals with spatial and international comparisons, it is appropriate now to examine and identify possible steps involved in a more integrated approach to the CPI and ICP activities.

This section first provides an assessment of potential benefits that could flow from the integration of CPI and ICP activities with statistical systems in general, and price statistics in particular, at both global and national levels. It is useful to note here that the ICP provides cross-country comparisons of prices of goods and services that enter private consumption, government consumption and investment. Thus ICP encompasses all components of gross domestic product (GDP). In contrast, the CPI focuses mainly on movements in prices of consumer goods and services. The nature and scope of the CPI and ICP activities and the limits on the extent of an integrated approach are discussed briefly. The final subsection identifies a number of useful initiatives that could provide a framework for a practical, more integrated approach to these important activities involving inter-temporal, inter-area and international comparisons of consumer prices and the ICP.

5.1 Benefits from CPI and ICP integration

Globalization and the resulting expansion in international trade and financial flows have led to an ever-increasing demand for internationally comparable

statistics that can be used in assessing the economic performance of nations. The ICP plays a major role in meeting this demand by providing internationally comparable national income aggregates, such as private and government consumption and capital formation. The ICP has also provided valuable information on relative international prices of goods and services, at a reasonably disaggregated level, which is used by researchers around the world. The most popular by-products of the international comparisons are the Penn World Tables and the World Development Indicators, which are considered invaluable sources of data for research on global inequality, poverty and econometric analysis of productivity growth, and the study of catch-up and convergence among nations.

The potential benefits of the ICP are somewhat diminished by the long lags in making international comparisons available to potential users. The coverage of the ICP is not extensive and the number of countries covered varies across regions. The limited coverage of ICP in some regions reflects the resource needs associated with the compilation of price data specifically for the purpose of the ICP. It is in this area where significant benefits can be derived if the ICP activities can be integrated with the CPI work of the national statistical offices.

At the global level, the potential benefits are many and varied. A few are listed below:

- increased country coverage, leading to a better framework for extrapolations;
- improved quality of the estimates resulting from the use of extensive price data collected for CPI purposes, rather than basing comparisons on products with specifications that may not be representative of the consumption baskets of the countries involved;
- benefits from research on methods for quality adjustment. Such methods are necessary to make adjustments for differences in product quality across countries;
- the development of regional PPPs, which are likely to make ICP results more consistent with domestic price movements and more acceptable to national governments;
- construction of internationally comparable national accounts, in a common currency unit, complementing the existing national accounts in national currency units. Such accounts will be a useful addition to the international statistical data bases that will enable global-level research on country and regional economic performance, and long-term catch-up, and convergence among countries;
- reliable estimates of PPPs along with domestic rates of inflation, providing a complete matrix of temporal–spatial price differences that can be used to better understand the factors influencing national price levels and exchange rate movements.

Several benefits may also be derived by the national statistical offices from an integrated approach to the compilation of CPI and PPPs. The nature of benefits derived will, however, vary depending on the stage of development of the countries involved. For more developed countries with well-established statistical agencies and programmes, the benefits derived are through the synergies arising out of a joint approach to spatial and temporal comparisons. These are as follows:

- Recent work on CPI and ICP manuals addresses the important issue of quality change over time and quality differences across countries. Efforts are being channelled into finding suitable statistical methods that can be used in the measurement of price-level changes leading to reductions in potential biases.
- The treatment of new and disappearing products within the CPI is also an important problem associated with increased globalization, the expansion of free trade and the removal of tariff barriers. Thus national statistical offices, especially in developed countries, are often confronted with the problem of accounting for new goods, and goods subject to rapid quality changes.
- The regular compilation of the national accounts statistics and the measurement of price change through the CPI, and extending to inter-area and international comparisons, can provide information needed to assess the levels of, and movements in, real income.

In the case of developing countries with inadequate statistical infrastructure, benefits from an integrated approach could be significant and aid statistical capacity building in such countries. These benefits include:

- strengthening statistical infrastructure and institution building. Efforts to implement the CPI manual and its recommendations, along with participation in the ICP activity, are likely to identify deficiencies in a country's infrastructure and the lack of institutional capacity. In some countries, it may be necessary to strengthen management and planning functions, recruit new staff, and provide training to conduct household expenditure and other general price surveys;
- strengthening of data collection, processing systems and dissemination. Designing and conducting sample surveys to international standards is a lengthy and demanding exercise. It may be necessary to develop a systematic plan to improve survey designs to coordinate surveys of economic and business entities, and to conduct periodic censuses of economic activity. Computerization of data-processing activities may also be required;
- improvement in the measurement of income inequality and poverty. Improved price and expenditure data can be used to improve national and international estimates of poverty;
- improved regional comparisons. PPPs compiled for regions and for neighbouring countries can provide useful insights into the dynamics of regional development, and help identify the regions with special needs and required assistance. Comparisons with neighbouring countries, in terms of relative price differences and real income changes, can provide powerful incentives to pursue policies for growth and low inflation.

There are many potential applications of the PPPs from the ICP. Several papers (by Astin (2001), Ward (2001) and Prennushi (2001), in particular) at the recent OECD–World Bank (2001) seminar dealt with some important applications of PPPs at Eurostat and the OECD, and at the global level in the assessment of poverty and inequality.

5.2 Salient features of CPI and ICP integration

In order to identify strategies for closer integration of temporal CPI compilation with wider inter-area and cross-country comparisons of consumer prices and PPPs at the level of GDP, it is necessary to examine the main features and the context in which such an integration will occur. The important issues are the scope and coverage of these two endeavours from the perspective of national statistical offices, and a general framework of price comparisons within which both these activities are placed. This section deals with these two aspects.

The scope and coverage of the CPI and ICP are vastly different. The CPI is a measure of changes, over time, in prices of goods and services that belong to the consumption baskets of households in a given country. In contrast, the ICP provides a measure of price-level differences across countries, covering all components of the expenditure side of the national accounts. The main components of GDP used in ICP comparisons are household consumption, government consumption, capital formation and net exports. In line with the *SNA 1993*, the ICP merges the portion of government expenditure that provides goods and services to households with private consumption to form household consumption. Thus, the scope and coverage of goods and services in the ICP is much wider than that of the CPI. The household consumption concept used is, however, almost identical to that used in constructing the CPI. Any integration of CPI and ICP work will necessarily be confined to the household consumption aggregate of the national accounts.

Within the ICP, price quotations from different countries are obtained for a large number of goods and services with very well-defined product specifications. This approach, described as the ''tight specification'' approach, is used in the ICP. The tight specification approach requires a sufficient description of the product so that it is uniquely defined in the ''law of one price'' sense, and so that it can be recognized in a range of localities and time periods wherever and whenever it is available. The product listing is determined on the basis of the group of countries included in an international comparison exercise. While this approach provides a solution to the problem of quality variations across countries, the commodities priced for ICP purposes may not be representative of the consumption baskets in respective countries. Thus, commodities priced may not be representative of the items consumed in the countries, which usually make up the bundle of goods and services for the construction of the CPI.

The degree of success of integrating ICP activity with the CPI compilation depends upon the extent to which these two activities can draw on a common pool of data and information available at the national level. The intersection of data sets for the CPI and the ICP is represented in figure A4.2.

The following marked areas are of particular significance:

(1) Common set of goods and services between the CPI and ICP lists. These price data can be used directly in an integrated approach to these two activities.

(2) and (3) Subset of ICP goods and services for which prices can be derived after making quality adjustments to products listed in the CPI basket. These are the goods and services that are not identical, but close enough for quality adjustments to be made based on the characteristics of the goods and services.

(4) Set of goods and services in the CPI basket that has no direct component in the ICP basket.

(5) ICP basket of goods and services under headings of the expenditure side of the national accounts that has no direct relevance or correspondence with CPI.

Figure A4.2 shows that if the integration between the ICP and the CPI is to be successful, the ICP comparisons should necessarily be restricted to country groupings where the basket of goods and services representing household consumption within the ICP has a significant overlap with the country-specific CPI baskets. Such an overlap can be achieved only when country groupings within the ICP comparisons exhibit similarities in their CPI consumption baskets. This has implications for the ICP and its regionalization programme.

In examining the CPI and ICP activities, it is also necessary to consider a range of temporal and spatial price comparisons of interest to national statistical offices in the course of providing a comprehensive set of economic statistics for policy-makers and other analysts. It is possible to consider these activities in a sequence indicating the progression involved in these price comparison activities (figure A4.3).

The schematic diagram uses the standard CPI activity of estimating annual or quarterly price changes for the nation as a whole. In most countries, national CPI figures are supplemented by area-specific CPI estimates for either capital cities or regions within the country. A natural progression, where data permit, is to undertake spatial comparisons of prices. At present, very few

Figure A4.2 Price data for CPI and ICP activities

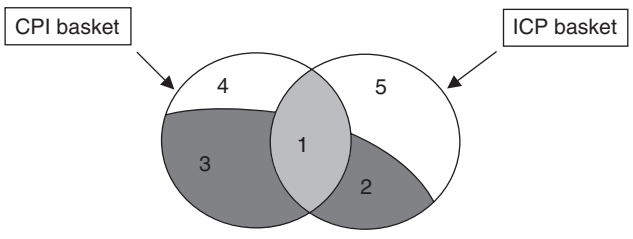
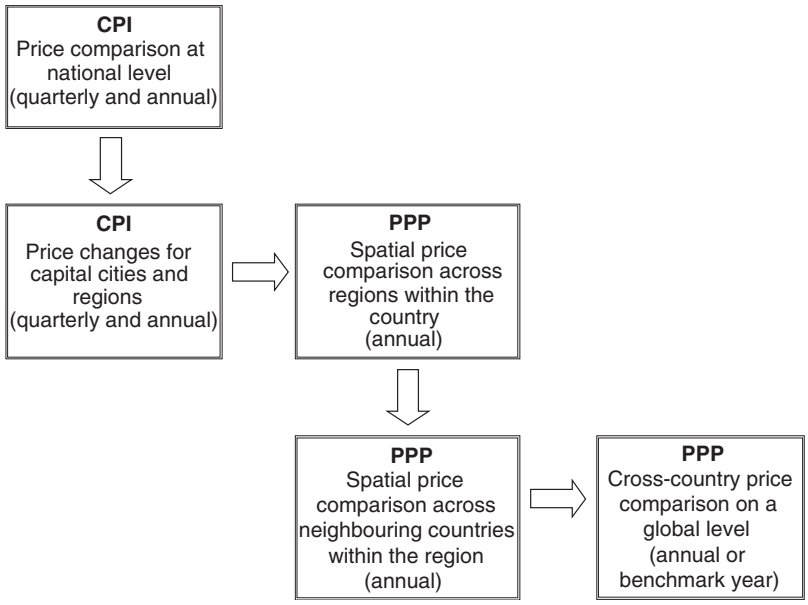


Figure A4.3 A sequence of price comparisons



countries appear to produce such indices on a regular basis. A fairly significant jump from this level is to undertake price comparisons, on a bilateral or multilateral basis, with geographically contiguous countries or countries within a political or economic grouping. The last element in this chain is the participation of the national statistical offices in a global price comparison exercise such as the ICP. At present, national statistical offices are mostly involved at the two extreme ends of this spectrum. The general level of involvement of national statistical offices in the ICP is likely to be more enthusiastic, however, when they gain experience from their participation in inter-area comparisons within the country and intra-regional comparisons involving countries in geographical proximity.

5.3 Two core strategies for CPI and ICP integration

Based on the brief discussion of the CPI and ICP activities undertaken by participating national statistical offices, it is possible to identify two major strategies that will result in a level of integration between these two activities that can benefit both programmes and the systems of economic statistics in the participating countries. Both of these strategies emerge from the need to maximize the flow of data from the CPI and ICP, and at the same time provide a framework for improving temporal and inter-area consumer price comparisons within a country.

Use of characteristics approach. This approach was proposed in Zieschang et al. (2001) and alluded to in Rao (2001a). The characteristics approach begins with a market study by a national statistical office analyst to determine a set of price-determining characteristics. These product characteristics, such as size, features, nature of the sale transaction, type of outlet, and so on, are determined according to the available information about the impact of the detailed characteristic on price at a point in time or over a specified reference period, such as a year. Under this approach, product prices are collected and at the same time the product characteristics are also recorded.

In the standard CPI context, the product listing remains fairly constant, except in cases where an old variety or product is replaced by a new one. When spatial comparisons are undertaken, however, the overlap in products may be limited. In such cases, the characteristics approach becomes useful. This approach is in direct contrast with pricing very specific products in all the countries, or in areas within a country, thus limiting the overlap and the usefulness of the resulting CPIs.

The characteristics approach requires price and characteristics data for a sufficiently wide variety of detailed commodities or specifications in the item group to estimate a regression model of price on characteristics. Such models are known as hedonic regression models, where the logarithm of price is regressed on various characteristics. For spatial comparisons, this approach is very similar to the country–product–dummy (CPD) method discussed above in this annex. Kokoski et al. (1999) describe a method of producing "exact", characteristics-adjusted, economic index number comparisons between areas that allow for hedonic equation parameters to differ from one area to another. This method can be classified as a variant of the weighted CPD method along with a multilateral Törnqvist index number.

The main advantages of the characteristics approach is that it is not critically based on the overlap in commodity bundles for different areas or countries, but on the sample size and number of commodities for which prices and quality characteristics are collected. The sample size needs to be large enough to enable efficient estimation of the parameters involved.

In order to facilitate cooperation with the ICP, in addition to having a well-structured database for the CPI that is extended to cover characteristics, the interests of national statistical offices are ultimately best served by implementing product and characteristic classification schemes consistent with internationally agreed standards. In order to make this approach operational, it is necessary that such classification schemes are established by making use of some of the existing classifications, for example the Central Product Classification (CPC) or the Classification of Individual Consumption according to Purpose (COICOP). These product classifications need to be extended to include a core set of standard characteristics for each category within the classification.

Notwithstanding the benefits of the characteristics approach, implementing a scheme of product characteristics classification requires all the countries or parties involved to agree to a specific standard and to allocate the necessary resources for such an endeavour. Until comprehensive data sets with price and characteristics of products become available, it may be necessary to explore other possible approaches that can be used in conjunction with the present ''tight specification'' approach used in ICP exercises.

Linking approach to international comparisons. If maximizing the overlap with the CPI is one of the principal objectives of the ICP, then multilateral cross-country comparisons need to be built up from bilateral comparisons, where pairs of countries are identified on the basis of the maximum overlap in their national CPI baskets. Once such pairs of countries are identified, then multilateral comparisons can be built using chains constructed on the basis of links. This approach is somewhat similar to the MST approach proposed by Hill. While the basic criterion in Hill's approach uses variability in price relatives, measured using the Laspeyres–Paasche spread, the approach suggested here requires measures of overlap of price data as the principal criterion.

The linking approach needs a multi-stage framework. In the first stage, it is necessary to identify groups of countries to form regions or clusters. The principal criteria that should be used are, first, the extent of overlap that can be achieved between pairs of countries within the group and some measures of similarity in the expenditure patterns. In the second stage, an MST approach should be used to identify the exact links within a regional cluster of countries. Once multilateral regional PPPs are constructed, and PPPs for the GDP and its major components are derived, then the next stage will involve linking various regional comparisons to derive a set of global comparisons and PPPs.

Application of the chaining procedure represents a major shift from the present approach to ICP work. Currently, comparisons within the ICP are based essentially on a top-down approach, where a commodity listing based on the ''tight specification'' approach is determined in the first instance, and then price data from different countries are collected. Where the ICP work is regionalized, regions are essentially determined on geographical considerations and not on any data-based considerations. If cross-country comparisons are to be based on nationally available CPI data, it is necessary to use a bottom-up approach, where all the operational procedures, including the determination of clusters and links, need to be built with the data available from the national sources provided by the national statistical offices. The application of spatial linking procedures will minimize the need for quality adjustments of the type described under the characteristics approach to price data collection.

The integration of spatial and temporal comparisons has the potential to provide a consistent set of temporal, inter-area and inter-country comparisons, and at the same time improve the quality of the underlying comparisons. It is an exciting prospect, but several challenges need to be met before a truly integrated approach to spatial and temporal consumer price comparisons can be achieved.