A note on the Methodology for Linking the Regions within the ICP

Prasada Rao

To be presented at the TAG Meeting

Global Office

2nd Technical Advisory Group Meeting
February 17-19, 2010
Washington DC
## Table of Contents

1. **Notation**  
   - 1

2. **Regional price comparisons**  
   - 1

3. **Linking regional comparisons at the basic heading level**  
   - 2
     - Sergeev’s critique  
     - An additional problem with multiple linking countries within regions  
     - 3

4. **Linking the Regions at levels above the Basic Heading Level**  
   - 4
     - Regional Linking Factors using Regional Expenditure Weights  
     - 5

5. **Conclusions**  
   - 5

References:  
   - 6
A note on the Methodology for Linking the Regions within the ICP

D.S. Prasada Rao
School of Economics
The University of Queensland, Brisbane

Subsequent to the publication of the ICP 2005 results and as a prelude to the upcoming 2011 round of the ICP, the issue of methodology for linking regions is being revisited. Sergeev has been an active contributor to these discussions leading to a short paper by Sergeev (2009) entitled, “the Evaluation of the Approaches Used for the Linking of the Regions in the ICP 2005” outlining the properties of the Diewert methodology and offering a critique identifying possible problems with the methodology. The main objective of this short note is to summarise the approaches discussed in Diewert (2008) which formed the basis for the linking of regional data within the 2005 round of the International comparison program and the exposition provided in a recent paper by Diewert (2010). The linking procedure used in ICP 2005 has the following steps: (i) PPPs of currencies in each region are computed setting PPP for the numeraire country currency equal to 1. These parities are available for each region at the basic heading level as well as at higher levels of aggregation including PPP at the GDP level; (ii) A set of ring countries are identified for the purpose of linking regional comparisons. In ICP 2005, 18 countries were selected for this purpose with each region represented by at least two countries; (iii) Price surveys are conducted in the ring countries to collect prices on the ring-country product list; and (iv) Price data from the ring countries were aggregated using the methodology described in Diewert (2008, 2010) to yield global comparisons. Various issues related to the regional linking methodology are discussed below.

1. Notation

The notation used here is essentially the one used in various papers by Diewert (2008, 2010). An attempt is made to translate notation used in Sergeev (2009) to be consistent with Diewert’s notation. Suppose that there are R regions in the comparison with C(r) countries in region r. We assume that there are N commodity groups, each commodity group represents a basic heading within the ICP. Let L(r) (r=1,…,R) be the number of link countries selected from region r. ∑ L(r) = L = total number of link countries. Therefore there are L price vectors representing price data from the ring countries that need to be aggregated in order to derive linking factors to link different regions.

2. Regional price comparisons

The first step in the compilation of PPPs is the computation of PPPs for all the C(r) countries in each of the regions. At the basic heading level, these PPPs are derived using the CPD method.

Let αrcn represent purchasing power parity of currency of country c in region r for the basic heading (or commodity group) n for r=1,2,…,R; c=1,…,C(r) and n =1,2,…,N.

---

1 Basic heading within the ICP represents the most disaggregated level for commodity groups for which expenditure share weights are available. 155 basic headings were used in ICP 2005.
2 Even though the recommended method is the CPRD method which is designed to make use information on representativity, only the CPD method was implemented in most of the regions. The Eurostat-OECD region used a variant of the EKS method.
We assume that country $c=1$ in each region is selected to be the numeraire country so that its parity is set to 1. That is, $\alpha_{rcn}=1$ for each $r$. These parities are estimated through the following CPD regression model:

$$\ln p_{ircn} = \alpha_{rcn} + \eta_{rn} + u_{icrn}$$

where $p_{ircn}$ is the price of $i$-th commodity in $n$-th basic heading observed in country $c$ in region $r$.

The basic heading parities, $\alpha_{rcn}$, are then aggregated using expenditure data available at the basic heading level using the EKS method. Additive methods like the Geary-Khamis and Ikle methods were also used in the 2005 round of the ICP.

3. **Linking regional comparisons at the basic heading level**

The basic input into calculations at this stage consists mainly of the price data collected by the ring countries, $L(r)$, $r=1,2,...R$. The item list used for ring country price surveys is different from the region-specific item lists used in regional comparisons. Let $p_{ircn}^L$ represent the price of $i$-th commodity observed in link country $c$ in region $r$. The Diewert-approach suggests:

1. the conversion of the prices of link countries in region $r$ into the currency of numeraire country (country 1 in each region)

   $$p_{ircn}^L = p_{ircn}^L / \alpha_{rcn} \text{ for } c=1,2,...L(r) \quad (1)$$

   At this step, ring country prices in different regions are converted into the currency of the corresponding numeraire country in that region.

2. Use the CPD model to derive regional linking parities at the basic heading level

   $$\ln p_{ircn}^L = \beta_{rn} + \eta_{n} + u_{icrn} \quad (2)$$

   Here, $\beta_{rn}$ represents the PPP for the $n$-th basic heading and a given region $r$ with respect to the numeraire region, region 1.

3. The PPPs of currencies within each region can now be converted into the currency of the numeraire region. If region 1 is taken as the numeraire region, then the world PPPs at the basic heading level, $\gamma_{rcn}$, are given by

   $$\gamma_{rcn} = \beta_{rn} \alpha_{rcn} \quad (3)$$

At the end of these three steps, a tableau of PPPs for 155 basic headings are generated using the currency of the numeraire country (say country 1) in the numeraire region (say region 1). It is fairly straightforward to show that this procedure yields PPPs for the regions which are invariant to the choice of the numeraire region and the numeraire countries within the regions (see page 3 of Diewert, 2010). Further, this tableau provides the basic price data for determining the regional linking factors at levels above the basic heading level.
Sergeev’s critique

Though there is little criticize in the methodology described in equations (1) to (3), Sergeev (2009) raises an interesting issue that arises due to the presence of different number of ring/link countries from different regions, i.e., $L(r)$ varies with $r$. In ICP 2005, there were 5 countries from Africa and 4 countries from Asia compared to 4 from the OECD\textsuperscript{3}. The main point is that within the CPD framework, it can be shown that the international average prices are averages of prices from the linking countries and therefore regions with greater number of linking countries may be seen to exert a larger influence.\textsuperscript{4} However, the rationale for including more than one country of a region in the ring list is that for those regions that are large and those that exhibit a large variation in prices it is necessary to use more price data drawn from a diverse set of countries representing the region. There are two possible solutions to the question raised in Sergeev (2009).

(i) Express the uncertainty attached to prices from a region through a larger variance for the corresponding disturbance term in the CPD model and then apply generalised least squares method which accords less weight to those observations with larger variance. This eliminates the problem discussed by Sergeev. However, it is difficult to know the extent of variability in prices across countries within a region.

(ii) Alternatively, the suggestion made in Sergeev’s paper may be implemented. Instead of using country specific prices for each of the link countries in a region, simply take the geometric mean of the prices after converting them into the currency of the numeraire country. In this case, we simply use

$$p_{ln}^{L(r)} = \prod_{c=1}^{L(r)} \left[ p_{ln}^{L(r)}/p_{r1n} \right]^{1/L(r)}$$

In the next step, use these geometric means as inputs into the CPD model to generate regional linking factors described in equation 2. Use of the geometric mean ensures that each region is represented by a single vector of prices in the CPD regression model thus ensuring that no region exerts more than proportionate influence on the basic heading level parities that are used in deriving linking factors at higher levels of aggregation.

In the absence of information about variability in price structures within a region, the use of geometric means as suggested by Sergeev may be the simplest way of ensuring all the regions are treated symmetrically within the CPD calculations at the BH level within the linking process.

An additional problem with multiple linking countries within regions

\textsuperscript{3} The full list of ring countries for the ICP 2005 round consists of: Brazil; Cameroon; Chile; Egypt; Estonia; Hong Kong; Japan; Jordan; Kenya; Malaysia; Oman; Philippines; Senegal; Slovenia; South Africa; Sri Lanka; United Kingdom; and Zambia.

\textsuperscript{4} Equations (1) to (3) in Sergeev (2009) can be derived from the normal equations associated with the ordinary least squares method.
Though Sergeev’s critique has not raised this problem, there is an additional issue that arises when we are working with multiple linking countries. We first note that the price data used for linking are different from the price data used for intra-regional comparisons. The ring list of goods and services could be qualitatively different from the lists used in the regional comparisons. This means, conceptually we can think of two sets of PPPs between the link countries within a region: one based on price data collected as a part of the regional comparison and another one based on prices collected for the ring list. For example, if Malaysia and Philippines are two ring countries from the Asia-Pacific region, it is possible to derive a PPP for these two currencies one based on price data from the regional comparisons and another based on price data collected for the purpose of linking regions. If these two sets of parities are close to each other, we may have confidence in the whole process of linking. However, if these parities are wildly different then there is an obvious disconnect between prices observed for regional lists and those collected for ring country comparisons. Therefore, this is an issue that is worth further consideration.

4. **Linking the Regions at levels above the Basic Heading Level**

Suppose we have the BH level parities for regions computed using price data collected for the ring countries. In Diewert’s notation (2010), suppose we have interregional linking PPPs, $\beta_{rn}$ ($r=1,2,...R; n=1,2,...N$). We need data on expenditure weights at the regional level so that the BH PPPs can be aggregated to yield PPPs at levels of aggregation above the BH level. Diewert (2010) summarises a few alternative approaches that can be used and discusses the relative merits of these approaches. Methods employed at this stage of aggregation are the subject matter of active debates.

- **Option 1**: Convert country-specific expenditures into the currency of numeraire country using the market exchange rates and compile comparisons based on the expenditure totals and the implied unit values. As this approach requires the use of market exchange rates, this is not a procedure actively pursued in Diewert (2010).

- **Option 2**: The second option is to use the regional PPPs to define imputed quantities as: $Q_{rcn} = E_{rcn}/\gamma_{rcn}$ where $\gamma_{rcn} = \beta_{rn}\alpha_{rcn}$. These quantities are expressed in units which are comparable across countries and regions. In this case, the prices and quantities for each region are respectively given by,

$$Q_{rn} \equiv \sum_{c=1}^{C(r)} Q_{rcn}$$

and

$$P_{rn} \equiv \beta_{rn} \text{ for } r=1,2,...R \text{ and } n=1,2,...N$$

(4)

These price and quantity data can be aggregated using the EKS and other aggregation methods. The main point made in Sergeev (2009) is that this procedure is not invariant to the choice of the numeraire country within the regions. The relativities across commodity groups tend to change when the regional numeraire country is changed. Sergeev demonstrates this result using a numerical example presented in Tables 1 and 2 of his paper. This problem associated with Option 2 is also acknowledged in the recent note by Diewert (2010).

- **Option 3**: Diewert (2010) describes a third option which uses the geometric mean of the of the country parities over all countries in the region as regional price weights. He suggests the use of total volumes for regions as defined in (4) but replace the regional prices by:

$$P_{rn} = \prod_{c=1}^{C(r)} [\gamma_{rcn}]^{1/C(r)} \text{ for } r=1,2,...R \text{ and } n=1,2,...N. \quad (5)$$
Diewert rules out this option as it is equivalent to imposing a set of common prices to add up quantities within a region. The approach described in equation (4) appears to be a bit different from the approach described on page 6 of Sergeev (2010)’s note. Diewert (2010) also rules out the use of additive methods like the Geary-Khamis method as all the additive methods are subject to substitution bias.

**Option 4:** Diewert (2010) puts forward the use of a global EKS method which derives global comparisons satisfying fixity in two steps. In Step 1 a global EKS, without respecting fixity, is derived using the price tableau based on equation 3 whereby the price of commodity or BH n for country c in region r, is given by $\gamma_{rcn}$. Using these prices and the observed expenditures in respective quantities, one can derive implicit quantities. In the next step, use GEKS method on this data for the whole world to determine the world product and the shares of the world product for country c in region r. Aggregating these shares over countries within a region, one can obtain regional shares of the world output. In the second step, these regional shares are then distributed to countries within the region using the shares implicit in the regional comparisons. This approach appears to be consistent with the use of the GEKS methodology for comparisons within each region and it maintains fixity. Sergeev (2010)’s note appears to be favourable to the use of this approach.

**Option 5:** Diewert (2010) also discusses spatial comparisons derived using chained links established on the basis of the similarity of price structures. Diewert’s suggestion makes use of the spanning tree methodology is developed in various papers by Hill including his recent paper, Hill (2009). This topic deserves a separate paper in its own right and no attempt is made here to discuss this approach.

**Regional Linking Factors using Regional Expenditure Weights**

Sergeev (2009) also describes a simple approach for the derivation of regional linking factors. On the price side, BH parities for different regions using the ring country price data as described in equation (2) leading to price of n-th BH in region r given by $\beta_{rn}$ derived using CPD estimation in equation (2) are used. His proposal suggests the use of country-specific expenditure share data based on expenditures in national currencies. Let $e_{cn}$ represent expenditure on BH n in country c expressed in the currency of country C. These data are available from the national accounts. Using these, define the expenditure shares as

$$w_{cn} = \frac{e_{cn}}{\sum_{j=1}^{N} e_{cj}} \text{ for } n=1,2,…N \text{ and } c = 1,2,…C(R)$$  \hspace{1cm} (6)

Expenditure shares of the countries within a region are then averaged using either a plutocratic approach or a democratic approach. Under plutocratic approach, the average expenditure share patterns are likely to be dominated by the patterns observed for large countries. Sergeev (2009 or 2010) argues for the use of democratic weights where simple arithmetic average of expenditure patterns of the countries in region r is used as the expenditure pattern for region r. The average expenditure shares and the BH level parities compiled using ring country price data may then be combined using a standard EKS or weighted CPD type of aggregation method.

**Conclusions**

1. There is a general consensus regarding the compilation of PPPs at the basic heading level. The original methodology proposed in Diwiert (2004 and 2008) appears to be a robust procedure for aggregation at the BH level.
2. In the presence of more than one ring country within each region, which is usually the case, a simple method of avoiding unequal weights to price data from different regions is to compute a geometric mean of the prices for different ring countries within each region and use them as an inputs into the computation of BH parities between regions.

3. The question of how we make optimum use of the presence of a core list of products that will be priced by all the countries need to be discussed at further. Annex 2 of Sergeev (2009) briefly looks at this problem and provides expressions for the normal equations associated with the estimation of parameters of the CPD model in the presence of a core list.

4. Based on the material presented in Diewert (2010) and Sergeev (2009), it is quite clear that linking above the basic heading level (macro-linking) needs to be revisited. It is necessary to have a robust discussion about the relative merits of what I consider are the three alternative approaches: The Global Comparison EKS Method; Chaining methods using the Minimum Spanning Tree approach and alternative measures of similarity of price structures; and the use of democratic expenditure weights derived using expenditure shares computed using country-specific data in national currency units.

5. A question we will need to address at some stage is whether the reported differences in regional shares derived using different approaches are statistically significant. If we start computing some measures of variability in the form of standard errors for the PPPs and the regional shares, we may find that a number of differences reported in tables 5 and 14 in Diewert (2010) may not be significantly different statistically. However this aspect needs further work.

References:


