Journal of Econometrics 191 (2016) 407–413

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

A least squares approach to imposing within-region fixity in the International Comparisons Program

Robert J. Hill

Department of Economics, University of Graz, Austria

ARTICLE INFO

Article history:
Available online 17 December 2015

JEL classification:
C43
E31
O47

Keywords:
ICP
Within-region fixity
Multilateral price index
GEKS

ABSTRACT

The International Comparisons Program (ICP) compares the purchasing power of currencies and real income across countries. ICP is broken up into six regions. Global results are then obtained by linking these regions together at both basic heading level and the aggregate level in a way that satisfies within-region fixity (i.e., the relative parities of a pair of countries in the same region are the same in the global comparison as in the within-region comparison). Standard multilateral methods violate this within-region fixity requirement and hence cannot be used to construct the global results. A method is proposed here that resolves this problem by altering the price and quantity indexes by the least-squares amount necessary to ensure that within-region fixity is satisfied. This method is then compared – both in terms of its underlying structure and empirically – with other methods for imposing within-region fixity.

1. Introduction

The International Comparisons Program (ICP) is a huge undertaking coordinated by the World Bank in collaboration with the Organization for Economic Cooperation and Development (OECD), Eurostat, International Monetary Fund (IMF) and United Nations (UN). It compares the purchasing power of currencies and real income of almost all countries in the world (see Deaton and Heston, 2010). These purchasing power parities can be used to compute internationally comparable national income aggregates, including gross domestic product (GDP) and per capita income. The results inform the debate over trends in global inequality, and the relative performance of individual countries and regions.

ICP data are used in the construction of the Penn World Table (probably the most widely used data set in the economics profession), the World Development Indicators published by the World Bank, the World Economic Outlook published by the IMF, and the Human Development Index (HDI) published by the United Nations Development Program (UNDP). The World Bank uses the ICP results to measure regional and global poverty, providing estimates of the number of people living on less than US$1 or $2 per day, and by the World Health Organization (WHO) and United Nations Educational, Scientific and Cultural Organization (UNESCO) to compare expenditures on health and education respectively across countries (see Rao, 2013). ICP results are also used by the IMF when computing its special drawing rights (SDRs)—which determine budget contributions and voting power (see Silver, 2010).

The most recent ICP comparisons were made in 2005 and 2011. The 2011 results were released in late April 2014. Both ICP 2005 and 2011 are broken up into six separate regional comparisons. The regions are Asia Pacific, South America, OECD–Eurostat, Confederation of Independent States (CIS), Africa, and West Asia.[1]

ICP 2011 divides GDP into 155 basic headings. A basic heading is the lowest level of aggregation at which expenditure weights are available (typically obtained from the national accounts). A basic heading consists of a group of similar products defined within a general product classification. One of the food basic headings for example is ‘rice’. The products in this heading vary to some extent from region to region. Some examples of possible ‘rice’ products include the following: long grain rice, parboiled long grain rice, non-parboiled long grain rice, jasmine rice, basmati rice, white rice, medium grain brown rice, and short-grained rice. To ensure the prices are comparable, the physical characteristics (e.g., the weight of the bag of rice) and the economic characteristics (e.g., whether it is a brand) are specified. Each country collects multiple price quotes from different locations on each product in a heading. These are then combined to produce an average price for each product in each country.

[1] The Caribbean constitutes a subregion of South America that is linked in at the end in ICP 2011. Similarly, the Pacific islands are linked to Australia and New Zealand. There are also a small number of singleton countries that are treated separately.

E-mail address: robert.hill@uni-graz.at.

http://dx.doi.org/10.1016/j.jeconom.2015.12.011

Within-region basic heading price indexes in both ICP 2005 and 2011 are calculated from the country-product average prices using the country-product-dummy method (see Summers, 1973). To obtain a global set of results it is then necessary to link the results across regions. An important complication that arises in this region-linking process is that the global results are required to satisfy within-region fixity at all levels of aggregation from basic heading up to GDP. In other words, the relative parities of a pair of countries in the same region must be the same in the global comparison as in the within-region comparison. There are two reasons for imposing within-region fixity. The first is essentially political. The European Union (EU) uses the official EU results to calculate budget contributions and the disbursement of grants and aid, and hence only wants one set of within-EU parities in the public domain. More generally, in other regions there is also concern that the availability of multiple within-region parities could generate confusion. The second reason is that the within-region comparisons are probably more reliable than the global comparisons. This partly reflects the inherent regional structure of ICP with each region having its own product list for each basic heading. However, it is also the case that countries within a region tend to have more similar levels of economic development, thus making it easier to compare them.

In ICP 2005 and 2011, the basic heading price indexes are linked across regions using the region-product-dummy method (see Diewert, 2008). Once the regions have been linked at basic heading level, in principle the aggregate level results can be computed using a standard multilateral method (such as GEKS – named after Gini, 1931; Eltetö and Köves, 1964; Szulc, 1964, – or Geary–Khamis).[2] However, this would lead to a violation of within-region fixity.

The question then is how to link the regional results at the aggregate level in a way that maintains within-region fixity? In this paper a new and flexible method – referred to as the Least Squares Fixity (LSF) method – is proposed for doing this (see Section 2). The LSF method is optimal in the sense that it alters the multilateral price indexes by the least-squares amount necessary to ensure that within-region fixity is satisfied.

The within-region comparisons at the aggregate level in ICP 2011 were made using the GEKS method.[3] The underlying rational of the LSF method is similar to that of GEKS, which alters Fisher price indexes by the least-squares amount necessary to ensure transitivity. Also, like GEKS, the LSF method effectively takes geometric means of all possible chained comparisons between a pair of reference countries (see Section 3), and treats prices and quantities symmetrically (see Section 4). In all these senses, therefore, LSF can be viewed as an extended version of GEKS that satisfies within-region fixity.

In ICP 2011, within-region fixity at the aggregate level is imposed using the Country Approach with Redistribution (CAR) method proposed by Kravis et al. (1982) (see also Heston, 1986, 2010). I describe the CAR method in Section 5. In Section 6 I compare the underlying algebraic structures of the LSF and CAR methods within a more general taxonomy. In Section 7 I then compare the LSF and CAR methods empirically using ICP 2005 data. While for most regions the choice between LSF and CAR has little impact, the same is not true for the CIS region where the resulting per capita incomes, relative to the US, differ by on average 8.6% depending on which method is used. My overall conclusions are summarized in Section 8.

[2] See Hill (1997) for a survey of standard multilateral methods.

[3] In ICP 2005, all regions used GEKS except Africa, which used the IDB method (see Iklé, 1972; Dikhanov, 1997; Balk, 1996).

2. The least squares fixity (LSF) method

The imposition of within-region fixity requires the development of methods designed specifically for this task. However, before doing this it is illuminating first to consider a least squares derivation of the GEKS price index.

The GEKS method alters intransitive bilateral price indexes $P_{j,k}^{Bilat}$ (e.g., Fisher or Törnqvist) – where j and k denote countries – by the logarithmic least squares amount necessary to obtain transitivity.[4] The GEKS problem can be written in regression format as follows:

$$\ln P_{j,k} = \ln P_k - \ln P_j + \epsilon_{j,k}, \tag{1}$$

where $\epsilon_{j,k}$ is a white noise error term. In matrix notation, we have that

$$p^{Bilat} = Dp + \epsilon,$$

where for the case of four countries p^{Bilat}, D and p are defined as follows:

$$p^{Bilat} = \begin{pmatrix} P_{1,2}^{Bilat} \\ P_{1,3}^{Bilat} \\ P_{1,4}^{Bilat} \\ P_{2,3}^{Bilat} \\ P_{2,4}^{Bilat} \\ P_{3,4}^{Bilat} \end{pmatrix}, \qquad D = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

$$p = \begin{pmatrix} \ln(P_1) \\ \ln(P_2) \\ \ln(P_3) \\ \ln(P_4) \end{pmatrix}.$$

Solving the ordinary least squares problem: $\text{Min}_p (p^{Bilat} - Dp)^T (p^{Bilat} - Dp)$, yields the solution: $\hat{p} = (D^T D)^{-1} D^T p^{Bilat}$, which reduces to:

$$\text{GEKS:} \quad \frac{\hat{P}_k}{\hat{P}_j} = \prod_{i=1}^{N} \left(\frac{P_{i,k}^{Bilat}}{P_{i,j}^{Bilat}} \right)^{1/N}. \tag{2}$$

Returning now to the problem of imposing within-region fixity, at this point it is useful to introduce some notation. The regions in the comparison are denoted here by $A, B, C, \ldots$ etc., or by $K = 1, \ldots, \mathcal{K}$ depending on the context. There are N_A countries in region A, N_B countries in region B, etc. P_{Aa}^{region} denotes a within-region A price index for country Aa, with country $A1$ as the base (i.e., $P_{A1}^{region} = 1$). Similarly, P_{Bb}^{region} denotes a within-region B price index for country Bb with country $B1$ as the base (i.e., $P_{B1}^{region} = 1$). $P_{Aa}^{unfixed}$ and $P_{Bb}^{unfixed}$ denote price indexes for countries Aa and Bb, obtained from a global comparison using a standard multilateral method with country $A1$ as the base country (i.e., $P_{A1}^{unfixed} = 1$). The unfixed price indexes do not satisfy within-region fixity.

The within-region (except Africa in 2005) and unfixed price indexes in ICP 2005 and 2011 are all calculated using GEKS. The difference between P_{Aa}^{region} and $P_{Aa}^{unfixed}$ is that, in the case of P_{Aa}^{region}, the GEKS formula is only applied to the countries in region A, while the unfixed price indexes $P_{Aa}^{unfixed}$ are obtained by applying the GEKS formula globally.

The requirement of within-region fixity implies that the global price index P_{Kk}^{global} for country k in region K must take the following form:

$$P_{Kk}^{global} = \lambda_K P_{Kk}^{region}, \quad \text{for } k = 1, \ldots, K \text{ and } K = 1, \ldots, \mathcal{K}. \tag{3}$$

[4] While GEKS itself was first proposed by Gini (1931), this least squares derivation was developed by Eltetö and Köves (1964), and Szulc (1964).

The question therefore is how to compute the regional scalars λ_K for each region?[5]

Least-squares optimal regional scalars $\hat{\lambda}_K$ can be derived by applying ordinary least squares to the following regression equation:

$$\ln\left(\frac{P_{Kk}^{unfixed}}{P_{Kk}^{region}}\right) = \ln\lambda_K + \epsilon_{Kk},$$

$$\text{for } k = 1, \ldots, K \text{ and } K = 1, \ldots, \mathcal{K}, \tag{4}$$

where $\mathcal{K}$ denotes the number of regions. Again ϵ_{Kk} is a white noise error term.

In matrix notation (4) can be rewritten as follows:

$$y^p = D\mu^p + \epsilon,$$

where for example focusing on the case of three regions A, B and C (each containing two countries) we have that:

$$y^p = \begin{pmatrix} y_{A1}^p \\ y_{A2}^p \\ y_{B1}^p \\ y_{B2}^p \\ y_{C1}^p \\ y_{C2}^p \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mu^p = \begin{pmatrix} \ln(\lambda_A) \\ \ln(\lambda_B) \\ \ln(\lambda_C) \end{pmatrix}.$$

Each element y_{Kk}^p of the y^p vector is calculated as follows:

$$y_{Kk}^p = \ln\left(\frac{P_{Kk}^{region}}{P_{Kk}^{unfixed}}\right).$$

Solving $\text{Min}_{\mu^p}(y^p - D\mu^p)^T(y^p - D\mu^p)$, yields the ordinary least squares solution: $\hat{\mu}^p = (D^TD)^{-1}D^Ty^p$. The resulting regional scalars for regions $K = 1, \ldots, \mathcal{K}$ take the following form:

$$\hat{\lambda}_K = \left[\prod_{k=1}^{N_K}\left(\frac{P_{Kk}^{unfixed}}{P_{Kk}^{region}}\right)^{1/N_K}\right]. \tag{5}$$

The global price index for country j in region K is hence:

$$P_{Kj}^{global} = P_{Kj}^{region}\left[\prod_{k=1}^{N_K}\left(\frac{P_{Kk}^{unfixed}}{P_{Kk}^{region}}\right)^{1/N_K}\right]. \tag{6}$$

The global price indexes derived above deviate from the unfixed price indexes by the least squares amount necessary to satisfy within-region fixity.[6] They are referred to henceforth as Least Squares Fixity (LSF) global price indexes.[7]

3. An alternative perspective on the LSF method

It follows from (6) that the LSF global price index of country Bk relative to country Aj can be written as follows:

$$\frac{P_{Bk}^{global}}{P_{Aj}^{global}} = \left(\frac{P_{Bk}^{region}}{P_{Aj}^{region}}\right)\left[\frac{\prod_{b=1}^{N_B}\left(\frac{P_{Bb}^{unfixed}}{P_{Bb}^{region}}\right)^{1/N_B}}{\prod_{a=1}^{N_A}\left(\frac{P_{Aa}^{unfixed}}{P_{Aa}^{region}}\right)^{1/N_A}}\right]$$

$$= \prod_{a=1}^{N_A}\prod_{b=1}^{N_B}\left(\frac{P_{Aa}^{region}}{P_{Aj}^{region}} \times \frac{P_{Bb}^{unfixed}}{P_{Aa}^{unfixed}} \times \frac{P_{Bk}^{region}}{P_{Bb}^{region}}\right)^{1/(N_A\times N_B)}. \tag{7}$$

Each term $(P_{Aa}^{region}/P_{Aj}^{region}) \times (P_{Bb}^{unfixed}/P_{Aa}^{unfixed}) \times (P_{Bk}^{region}/P_{Bb}^{region})$ in (7) can be thought of as a chained price index that compares countries Aj and Bk via countries Aa and Bb (i.e., the chaining path is $Aj \to Aa \to Bb \to Bk$). The overall global price index $P_{Bk}^{global}/P_{Aj}^{global}$ is the geometric average of the chained price indexes obtained by using all possible chain paths from Aj to Bk. For example, suppose there are three countries in region A, denoted by $A1$, $A2$, and $A3$, and two countries in region B, denoted by $B1$ and $B2$. There are then a total of six possible paths from say $A1$ to $B1$. These and their corresponding chained price indexes are listed below:

$$A1 \to B1: \quad \left(\frac{P_{B1}^{unfixed}}{P_{A1}^{unfixed}}\right)$$

$$A1 \to B2 \to B1: \quad \left(\frac{P_{B2}^{unfixed}}{P_{A1}^{unfixed}} \times \frac{P_{B1}^{region}}{P_{B2}^{region}}\right)$$

$$A1 \to A2 \to B1: \quad \left(\frac{P_{A2}^{region}}{P_{A1}^{region}} \times \frac{P_{B1}^{unfixed}}{P_{A2}^{unfixed}}\right)$$

$$A1 \to A2 \to B2 \to B1: \quad \left(\frac{P_{A2}^{region}}{P_{A1}^{region}} \times \frac{P_{B2}^{unfixed}}{P_{A2}^{unfixed}} \times \frac{P_{B1}^{region}}{P_{B2}^{region}}\right)$$

$$A1 \to A3 \to B1: \quad \left(\frac{P_{A3}^{region}}{P_{A1}^{region}} \times \frac{P_{B1}^{unfixed}}{P_{A3}^{unfixed}}\right)$$

$$A1 \to A3 \to B2 \to B1: \quad \left(\frac{P_{A3}^{region}}{P_{A1}^{region}} \times \frac{P_{B2}^{unfixed}}{P_{A3}^{unfixed}} \times \frac{P_{B1}^{region}}{P_{B2}^{region}}\right).$$

The global price index between $A1$ and $B1$ in this case is obtained by taking the geometric mean of these six chained price indexes.

Countries from third regions, such as C, drop out if they are included in the chain path. This is because the unfixed price indexes are transitive. For example, as shown below the path $Aj \to Cl \to Bk$ reduces to $Aj \to Bk$:

$$\frac{P_{Aa}^{region}}{P_{Aj}^{region}} \times \frac{P_{Cl}^{unfixed}}{P_{Aa}^{unfixed}} \times \frac{P_{Bb}^{unfixed}}{P_{Cl}^{unfixed}} \times \frac{P_{Bk}^{region}}{P_{Bb}^{region}} = \frac{P_{Aa}^{region}}{P_{Aj}^{region}} \times \frac{P_{Bb}^{unfixed}}{P_{Aa}^{unfixed}} \times \frac{P_{Bk}^{region}}{P_{Bb}^{region}}.$$

Similarly, additional countries from either regions A or B when included drop out of the chain path since the within-region price indexes are also transitive. For example, as shown below, the path $Aj \to Al \to Aa \to Bb \to Bk$ reduces to $Aj \to Aa \to Bb \to Bk$:

$$\frac{P_{Al}^{region}}{P_{Aj}^{region}} \times \frac{P_{Aa}^{region}}{P_{Al}^{region}} \times \frac{P_{Bb}^{unfixed}}{P_{Aa}^{unfixed}} \times \frac{P_{Bk}^{region}}{P_{Bb}^{region}} = \frac{P_{Aa}^{region}}{P_{Aj}^{region}} \times \frac{P_{Bb}^{unfixed}}{P_{Aa}^{unfixed}} \times \frac{P_{Bk}^{region}}{P_{Bb}^{region}}.$$

An analogy can again be drawn here with the GEKS formula in (2), which transitivizes bilateral price indexes, $P_{j,k}$, in a similar way. As long as the bilateral indexes satisfy the country reversal test (i.e., $P_{k,j} = 1/P_{j,k}$) as Fisher for example does, this formula can be rewritten as follows:

$$\text{GEKS:} \quad \frac{P_k}{P_j} = \prod_{i=1}^{N}\left(P_{j,i}^F \times P_{i,k}\right)^{1/N}.$$

Written in this way, it can be seen that a GEKS comparison between countries j and k takes the geometric mean of all possible chaining paths between j and k, $P_{j,i} \times P_{i,k}$, through third countries i. In this sense, the LSF method in (7) can again be viewed as a natural extension of GEKS.

The representation of the LSF regional scalars of the least-squares method as a geometric mean of all possible paths between pairs of countries drawn one from each region is also useful for demonstrating that the LSF method gives equal weight to all regions. Suppose now we have two countries in region A (denoted by A1 and A2), and four countries in region B, denoted by B1, B2,

[5] Each level of aggregation will have its own set of λ_K regional scalars.

[6] As is stands, none of the global price indexes in (6) are normalized to 1. However, they can easily be rescaled to achieve whatever normalization is desired (e.g., $P_{A1}^{global} = 1$).

[7] A variant on the LSF method has been used by the OECD since 1990 (see Sergeev, 2005) to impose within-region fixity on the results of the Eurostat subregion within the OECD.

B3, and B4. Using A1 and B1 as the regional bases, this means we take a geometric mean of the price indexes obtained by chaining along the following eight paths from A1 to B1:

(i) $A1 \to B1$,
(ii) $A1 \to A2 \to B1$,
(iii) $A1 \to B2 \to B1$,
(iv) $A1 \to A2 \to B2 \to B1$,
(v) $A1 \to B3 \to B1$,
(vi) $A1 \to A2 \to B3 \to B1$,
(vii) $A1 \to B4 \to B1$,
(viii) $A1 \to A2 \to B4 \to B1$.

The geometric mean of these eight paths contains more within-region price indexes from region B than from region A. More specifically, paths (ii), (iv), (vi) and (viii) contain a within-region price index from region A, while paths (iii)–(viii) contain a within-region price index from region B. In other words, within-region A price indexes feature in only four of the eight chain paths, while within-region B price indexes feature in six of the eight chain paths. This does not imply though that region B is exerting a 'greater influence' on the overall results. What matters here is the link countries between regions. The within-region price indexes should be viewed as simply rebasing a particular chain path's comparison back into units of the numeraire of that region, rather than as exerting weight in the overall comparison. A1 is the link country in region A in paths (i). (iii), (v) and (vi). A2 is the link country for the other four paths. Similarly, B1 is the link country for region B in paths (i) and (ii), B2 is the link country in paths (iii) and (iv), etc. This means that each country in region A is the link country in four of the paths, while each country in region B is the link country in only two paths. Hence each country in region A has twice the weight in the comparison as compared with each country in region B. However, there are twice as many countries in region B. Overall the weight allocation for the two regions therefore is the same.

4. The least squares fixity (LSF) method applied to quantity indexes

The requirement of within-region fixity applied to quantity indexes implies that the global quantity index Q_{Kk}^{global} for country k in region K must take the following form:

$$Q_{Kk}^{global} = \phi_K Q_{Kk}^{region}, \quad \text{for } k = 1, \ldots, K, \text{ and } K = 1, \ldots, \mathcal{K}, \tag{8}$$

where ϕ_K denotes a quantity index regional scalar.

Applying the LSF method to quantity indexes, the objective now is to construct global quantity indexes that alter the unfixed quantity indexes by the minimum amount necessary in a least squares sense to satisfy within-region fixity in all regions. In regression format this can be written as follows:

$$\ln\left(\frac{Q_{Kk}^{unfixed}}{Q_{Kk}^{region}}\right) = \ln\theta_K + \epsilon_{Kk}, \quad \text{for } k = 1, \ldots, K \text{ and } K = 1, \ldots, \mathcal{K}. \tag{9}$$

Again ϵ_{Kk} is a white noise error term.

The solution (which is analogous to the price index case) is:

$$\hat{\phi}_K = \left[\prod_{k=1}^{N_K}\left(\frac{Q_{Kk}^{unfixed}}{Q_{Kk}^{region}}\right)^{1/N_K}\right]$$

and hence the LSF quantity index of country j in region K takes the following form:

$$Q_{Kj}^{global} = Q_{Kj}^{region}\left[\prod_{k=1}^{N_K}\left(\frac{Q_{Kk}^{unfixed}}{Q_{Kk}^{region}}\right)^{1/N_K}\right]. \tag{10}$$

The link between LSF price and quantity indexes is revealed by combining (6) and (10) to yield the following expression:

$$\frac{P_{Bk}^{global} \times Q_{Bk}^{global}}{P_{Aj}^{global} \times Q_{Aj}^{global}} = \left(\frac{P_{Bk}^{region} \times Q_{Bk}^{region}}{P_{Aj}^{region} \times Q_{Aj}^{region}}\right) \times \left[\prod_{b=1}^{N_B}\left(\frac{P_{Bb}^{unfixed} \times Q_{Bb}^{unfixed}}{P_{Bb}^{region} \times Q_{Bb}^{region}}\right)^{1/N_B}\right] \Big/ \left[\prod_{a=1}^{N_A}\left(\frac{P_{Aa}^{unfixed} \times Q_{Aa}^{unfixed}}{P_{Aa}^{region} \times Q_{Aa}^{region}}\right)^{1/N_A}\right]. \tag{11}$$

Assuming the methods used to compute the within-region and unfixed price and quantity indexes satisfy the product test (i.e., price index times quantity index equals value ratio), then it follows that the three components of (11) can be rewritten as follows:

$$\frac{P_{Bk}^{region} \times Q_{Bk}^{region}}{P_{Aj}^{region} \times Q_{Aj}^{region}} = \left(\frac{\sum_{i=1}^{I} p_{Bk,i}q_{Bk,i}}{\sum_{i=1}^{I} p_{B1,i}q_{B1,i}}\right) \Big/ \left(\frac{\sum_{i=1}^{I} p_{Aj,i}q_{Aj,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right), \tag{12}$$

$$\begin{aligned}\prod_{b=1}^{N_B}\left(\frac{P_{Bb}^{unfixed} \times Q_{Bb}^{unfixed}}{P_{Bb}^{region} \times Q_{Bb}^{region}}\right)^{1/N_B} &= \prod_{b=1}^{N_B}\left[\left(\frac{\sum_{i=1}^{I} p_{Bb,i}q_{Bb,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right) \Big/ \left(\frac{\sum_{i=1}^{I} p_{Bb,i}q_{Bb,i}}{\sum_{i=1}^{I} p_{B1,i}q_{B1,i}}\right)\right]^{1/N_B} \\ &= \left(\frac{\sum_{i=1}^{I} p_{B1,i}q_{B1,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right),\end{aligned} \tag{13}$$

$$\prod_{a=1}^{N_A}\left(\frac{P_{Aa}^{unfixed} \times Q_{Aa}^{unfixed}}{P_{Aa}^{region} \times Q_{Aa}^{region}}\right)^{1/N_A} = \prod_{a=1}^{N_A}\left[\left(\frac{\sum_{i=1}^{I} p_{Aa,i}q_{Aa,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right) \Big/ \left(\frac{\sum_{i=1}^{I} p_{Aa,i}q_{Aa,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right)\right]^{1/N_A} = 1, \tag{14}$$

where $i = 1, \ldots, I$ indexes the basic headings, $A1$ and $B1$ are the base countries in regions A and B, and $A1$ is the base country in the unfixed comparison.

Substituting (12)–(14) into (11), the product of the global price and quantity index reduces to:

$$\begin{aligned}\frac{P_{Bk}^{global} \times Q_{Bk}^{global}}{P_{Aj}^{global} \times Q_{Aj}^{global}} &= \left[\left(\frac{\sum_{i=1}^{I} p_{Bk,i}q_{Bk,i}}{\sum_{i=1}^{I} p_{B1,i}q_{B1,i}}\right) \Big/ \left(\frac{\sum_{i=1}^{I} p_{Aj,i}q_{Aj,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right)\right] \times \left(\frac{\sum_{i=1}^{I} p_{B1,i}q_{B1,i}}{\sum_{i=1}^{I} p_{A1,i}q_{A1,i}}\right) \\ &= \frac{\sum_{i=1}^{I} p_{Bk,i}q_{Bk,i}}{\sum_{i=1}^{I} p_{Aj,i}q_{Aj,i}}.\end{aligned}$$

In other words, the global linking method as defined in (6) and (10) satisfies the factor reversal test (i.e., the product test plus the price and quantity indexes have the same functional form).[8]

On rearrangement this becomes

$$\frac{P_{Bk}^{global}}{P_{Aj}^{global}} = \frac{\sum_{i=1}^{I} p_{Bk,i}q_{Bk,i}}{\sum_{i=1}^{I} p_{Aj,i}q_{Aj,i}} \Bigg/ \frac{Q_{Bk}^{global}}{Q_{Aj}^{global}} = \frac{\tilde{P}_{Bk}^{global}}{\tilde{P}_{Aj}^{global}}.$$

The direct global price indexes $P_{Bk}^{global}/P_{Aj}^{global}$ are identical to the implicit global price indexes $\tilde{P}_{Bk}^{global}/\tilde{P}_{Aj}^{global}$ derived from the direct global quantity indexes via the factor-reversal equation. The direct and implicit global quantity indexes are likewise identical. This means that it makes no difference to the LSF method whether the starting point is the price or quantity indexes.

5. The CAR method for imposing within-region fixity

Kravis et al. (1982) and Heston (1986, 2010) suggest an alternative method for imposing within-region fixity on the global aggregate results, which they refer to as the 'country approach with redistribution' (CAR) method. It is typically expressed in terms of volume shares as follows:

$$s_{Aj}^{global} = s_{A}^{unfixed} \times s_{Aj}^{region}, \tag{15}$$

where the volume shares are essentially rescaled quantity indexes.

$$s_{Aj}^{global} = \frac{Q_{Aj}^{global}}{\sum_{n=1}^{N} Q_{n}^{global}},$$

is the share of country Aj in total world income in the global comparison.

$$s_{A}^{unfixed} = \frac{\sum_{a=1}^{N_A} Q_{Aa}^{unfixed}}{\sum_{n=1}^{N} Q_{n}^{unfixed}},$$

is the share of region A in total world income obtained from the unfixed global comparison (where within-region fixity is not satisfied).

$$s_{Aj}^{region} = \frac{Q_{Aj}^{region}}{\sum_{a=1}^{N_A} Q_{Aa}^{region}},$$

is the share of country Aj in the total income of region A obtained from the within-region A comparison. Converting (15) into quantity indexes yields the following expression:

$$\frac{Q_{Aj}^{global}}{\sum_{n=1}^{N} Q_{n}^{global}} = \left(\frac{\sum_{a=1}^{N_A} Q_{Aa}^{unfixed}}{\sum_{n=1}^{N} Q_{n}^{unfixed}}\right)\left(\frac{Q_{Aj}^{region}}{\sum_{a=1}^{N_A} Q_{Aa}^{region}}\right).$$

An equivalent expression for another country Bk is as follows:

$$\frac{Q_{Bk}^{global}}{\sum_{n=1}^{N} Q_{n}^{global}} = \left(\frac{\sum_{b=1}^{N_B} Q_{Bb}^{unfixed}}{\sum_{n=1}^{N} Q_{n}^{unfixed}}\right)\left(\frac{Q_{Bk}^{region}}{\sum_{b=1}^{N_B} Q_{Bb}^{region}}\right).$$

Now dividing the latter by the former, and rearranging yields the following:

$$\frac{Q_{Bk}^{global}}{Q_{Aj}^{global}} = \left(\frac{Q_{Bk}^{region}}{Q_{Aj}^{region}}\right)\left(\frac{\sum_{b=1}^{N_B} Q_{Bb}^{unfixed}}{\sum_{b=1}^{N_B} Q_{Bb}^{region}}\right)\Bigg/\left(\frac{\sum_{a=1}^{N_A} Q_{Aa}^{unfixed}}{\sum_{a=1}^{N_A} Q_{Aa}^{region}}\right). \tag{16}$$

Hence the CAR quantity index regional scalars ϕ take the following form:

$$\phi_K = \frac{\sum_{k=1}^{N_K} Q_{Kk}^{unfixed}}{\sum_{k=1}^{N_K} Q_{Kk}^{region}}. \tag{17}$$

Hence from (8) we have that:

$$Q_{Bk}^{global} = Q_{Bk}^{region}\left(\frac{\sum_{k=1}^{N_K} Q_{Kk}^{unfixed}}{\sum_{k=1}^{N_K} Q_{Kk}^{region}}\right). \tag{18}$$

The key difference between the LSF and CAR formulas in (10) and (18) is that LSF takes geometric means of the unfixed and within-region quantity indexes while CAR takes arithmetic means. As a consequence LSF satisfies the factor reversal test while CAR does not.

6. A taxonomy of regional scalars

One way of deriving the regional scalars λ_K in (3) is as some average of the difference between the within-region and unfixed price indexes of each region. The problem can be framed in terms of mean value functions $M(\cdot)$. Two general formulas for calculating λ_K are the following:

$$\text{(i)}\quad \lambda_K = \frac{M(P_K^{unfixed})}{M(P_K^{region})}, \qquad \text{(ii)}\quad \lambda_K = M\left(\frac{P_K^{unfixed}}{P_K^{region}}\right).$$

For example, $M(\cdot)$ could be a mean of order r defined as follows:

$$M(x_K) = \left[\frac{1}{K}\sum_{k=1}^{K}(x_{Kk})^r\right]^{1/r} \qquad \text{for } r \neq 0,$$

$$M(x_K) = \left[\prod_{k=1}^{K}(x_{Kk})\right]^{1/K} \qquad \text{for } r = 0.$$

If instead our starting point is the quantity indexes, then as was noted in Section 4 within-region fixity implies that:

$$Q_{Kk}^{global} = \phi_K Q_{Kk}^{region},$$

where ϕ_K is a regional scalar defined on the quantity indexes. The general formulas above can equally well be applied to quantity indexes:

$$\text{(iii)}\quad \phi_K = \frac{M(Q_K^{unfixed})}{M(Q_K^{region})}, \qquad \text{(iv)}\quad \phi_K = M\left(\frac{Q_K^{unfixed}}{Q_K^{region}}\right).$$

[8] In deriving this result it was assumed that the total expenditure (i.e., $\sum_{i=1}^{I} p_{Kk,i}q_{Kk,i}$) is the same for country Kk in the within-region and unfixed comparisons. My understanding is that this is the case in ICP 2005 and 2011. If this condition is violated then the factor reversal test will only hold approximately.

Now from the product test, implicit global and within-region price indexes are derived from their quantity index counterparts as follows:

$$\tilde{P}_{Kk}^{global} = \left(\frac{\sum_{i=1}^{I} p_{Kk,i} q_{Kk,i}}{\sum_{i=1}^{I} p_{A1,i} q_{A1,i}} \right) \Big/ Q_{Kk}^{global},$$
$$\tilde{P}_{Kk}^{region} = \left(\frac{\sum_{i=1}^{I} p_{Kk,i} q_{Kk,i}}{\sum_{i=1}^{I} p_{K1,i} q_{K1,i}} \right) \Big/ Q_{Kk}^{region}. \tag{19}$$

Substituting for Q_{Kk}^{global} and Q_{Kk}^{region} from (19) into (8), the following expression is obtained:

$$\tilde{P}_{Kk}^{global} = \left(\frac{\sum_{i=1}^{I} p_{K1,i} q_{K1,i}}{\sum_{i=1}^{I} p_{A1,i} q_{A1,i}} \right) \left(\frac{\tilde{P}_{Kk}^{region}}{\phi_K} \right), \tag{20}$$

which implies that implicit regional scalars for price indexes $\tilde{\lambda}_K$ can be derived from the regional scalars for quantity indexes ϕ_K as follows:

$$\tilde{\lambda}_K = \left(\frac{\sum_{i=1}^{I} p_{K1,i} q_{K1,i}}{\sum_{i=1}^{I} p_{A1,i} q_{A1,i}} \right) \left(\frac{1}{\phi_K} \right).$$

When r is set to zero in the mean of order r function (i.e., when $M(x_K)$ is a geometric mean), then all four approaches reduce to the LSF method. When $r = 1$ (i.e., an arithmetic mean), we obtain the CAR method from approach (iii).

The framework above demonstrates that LSF and CAR are by no means the only methods available for constructing regional scalars from the unfixed price or quantity indexes. However, these two methods each have their own particular advantages. LSF alters the unfixed price and quantity indexes by the minimum least-squares amount necessary to attain within-region fixity and treats prices and quantities symmetrically, while CAR achieves fixity while retaining the unfixed regional volume shares.

7. An illustrative example using ICP 2005 data

Based on ICP 2005 data for the 146 participating countries, the resulting per capita incomes are relatively insensitive to the choice between GEKS, LSF and CAR. The within-region GEKS per capita incomes are provided in the Online Appendix (see Appendix A) in Table A-1. The base country in each region is the first country listed. For example, the per capita income of Hong Kong in the Asia-Pacific region is normalized to 1.

The LSF regional scaling factors λ_K for the price indexes are computed using the formula in (5). The corresponding CAR regional scaling factors are obtained by inserting (17) into (20). These scaling factors along with their quantity index counterparts ϕ_K are shown in Table 1.

It can be seen from Table 1 that the choice between LSF and CAR has its biggest impact on the CIS region. The LSF and CAR scalars for CIS differ by 10% (with the Asia-Pacific as the base).

The overall global LSF and CAR results are obtained by multiplying the within-region per capita income indexes in Table A-1 by their regional scalars from Table 2. These results are shown in the Online Appendix (see Appendix A) in Table A-2. The base country in Table A-2 is the USA and hence all per capita incomes are expressed in US dollars. Also shown in Table A-2 are the unfixed GEKS per capita incomes.

Table 1
Price and quantity regional scalars—2005.

Regional price scalars λ_K	LSF	CAR
Asia Pacific	1.0000	1.0000
South America	4.0919	4.0313
OECD-Eurostat	5.9534	5.8590
CIS	0.5119	0.4637
Africa	0.1000	0.0989
West Asia	22.1578	21.8292
Regional quantity scalars ϕ_K	**LSF**	**CAR**
Asia Pacific	1.0000	1.0000
South America	0.2830	0.2788
OECD-Eurostat	0.8457	0.8323
CIS	0.3809	0.3450
Africa	0.0774	0.0765
West Asia	0.7420	0.7309

Table 2
Percentage changes in per capita incomes (relative to USA).

	GEKS-LSF	GEKS-CAR	LSF-CAR
Asia Pacific	4.43	3.57	1.61
South America	2.95	2.84	0.11
OECD–Eurostat	3.25	3.25	0.00
CIS	7.35	12.29	8.64
Africa	3.41	3.04	0.50
Western Asia	4.96	4.92	0.10

The average percentage difference in the per capita incomes in region K relative to those in region A generated by the unfixed-GEKS, LSF and CAR methods in Table A-2 can be measured as follows:

$$Z_{XY}^{A1,K} = \frac{100}{K} \sum_{k=1}^{K} \left[\frac{\mathrm{Max}(X_{A1,Kk}, Y_{A1,Kk})}{\mathrm{Min}(X_{A1,Kk}, Y_{A1,Kk})} - 1 \right],$$

where $X_{A1,Kk}$ denotes the per capita income of country k in region K measured in units of the currency of country $A1$ calculated using method X, while $Y_{A1,Kk}$ measures the corresponding income measured using method Y. Methods X and Y here refer to unfixed GEKS, LSF and CAR. Here we use OECD–Eurostat as the base region, and the US as country $A1$. This means that $X_{A1,Kk}$ and $Y_{A1,Kk}$ are both measured in US dollars. The $Z_{XY}^{A1,K}$ coefficients derived from comparisons between each pair of methods for each region are shown in Table 2.

Consider for example the Z coefficient for Asia Pacific when unfixed GEKS is compared with LSF. This implies that on average the per capita income of Asia-Pacific countries (measured in US dollars) changes by 4.43% when unfixed GEKS is replaced by LSF as the method for linking the regions together. By construction, $Z_{XY}^{A1,K}$ equals zero for the OECD–Eurostat region when methods LSF and CAR are compared. This is because both methods impose within-region fixity. Two main themes emerge from Table 2. First, the LSF and CAR methods generate results that are closer to each other than they are to the unfixed-GEKS results. Second, while the results for most regions are relatively insensitive to the choice between LSF and CAR, the same is not true for the CIS region where, measured relative to the US, this choice affects the per capita incomes of countries by on average 8.64%.

8. Conclusion

The Least Squares Fixity (LSF) method proposed here for linking regions in an ICP context is very flexible. It can be combined with any multilateral method (e.g., GEKS or Geary–Khamis), and applied to either price or quantity indexes. It is a natural generalization of the GEKS method, and is an optimal solution in a least-squares sense to the problem of imposing within-region fixity on global results. CAR (the method actually used to impose within-region fixity in ICP 2011), by contrast has the attractive property that it maintains the volume shares of each region obtained from the unfixed comparison.

The LSF method warrants attention as a potential alternative to CAR in future rounds of ICP. The choice between these methods, however, has only a limited effect on the results for most regions, based on ICP 2005 data. The exception is the CIS region, where the results change by on average 8.6% relative to the US. It remains to be seen whether a similarly large effect is observed when ICP 2011 data are used.

More generally, the taxonomy developed here should help illuminate the underlying structure of methods for imposing within-region fixity. Given the regionalized structure of ICP and the requirement of within-region fixity, it is important that the implications of this requirement are better understood.

Acknowledgments

I thank the World Bank for providing the data, and members of the ICP Technical Advisory Group for their comments. The views expressed here are my own and do not necessarily represent those of the World Bank.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jeconom.2015.12.011.

References

Balk, B.M., 1996. A comparison of ten methods for multilateral international price and volume comparisons. J. Official Stat. 12, 199–222.
Deaton, A., Heston, A., 2010. Understanding PPPs and PPP-based national accounts. Amer. Econ. J.: Macroecon. 2 (4), 1–35.
Diewert, W.E., 2008. New Methodology for Linking the Regions, Discussion Paper 08-07. Department of Economics, University of British Columbia, Vancouver, Canada.
Dikhanov, Y., 1997. Sensitivity of PPP-Based Income Estimates to Choice of Aggregation Procedures. Development Data Group, International Economics Department, The World Bank, Washington D.C, January.
Eltetö, O., Köves, P., 1964. On a problem of index number computation relating to international comparison. Stat. Sz. 42, 507–518.
Gini, C., 1931. On the circular test of index numbers. Int. Rev. Stat. 9 (2), 3–25.
Heston, A., 1986. World Comparisons of Purchasing Power and Real Product for 1980. United Nations and Eurostat, New York: United Nations.
Heston, A., 2010. Productivity Adjustments as a Special Case of a General Regional Linking Problem, Paper presented to the 4th Technical Advisory Group Meeting of the International Comparisons Program 2011 Round. Washington, DC, October 20–22.
Hill, R.J., 1997. A taxonomy of multilateral methods for making international comparisons of prices and quantities. Rev. Income Wealth 43 (1), 49–69.
Iklé, D.M., 1972. A new approach to the index number problem. Quart. J. Econ. 86 (2), 188–211.
Kravis, I.B., Heston, A., Summers, R., 1982. World Product and Income: International Comparisons of Real Gross Product. Published for the World Bank by Johns Hopkins University Press, Baltimore.
Rao, D.S.P., 2013. The Framework of the International Comparison Program, Chapter 1 in the ICP 2011 Handbook. World Bank, Washington DC.
Sergeev, S., 2005. Calculation of the Results of the Eurostat GDP Comparison with the Use of Fixity. Paper prepared for ICP Technical Advisory Group.
Silver, M., 2010. IMF Applications of Purchasing Power Parity Estimates, IMF Working Paper WP/10/253, International Monetary Fund. Washington DC.
Summers, R., 1973. International price comparisons based upon incomplete data. Rev. Income Wealth 19 (1), 1–16.
Szulc, B.J., 1964. Indices for multiregional comparisons. Prz. Stat. (Stat. Rev.) 3, 239–254.