

# Methods of Aggregation above the Basic Heading Level: Linking the Regions

W. ERWIN DIEWERT

This chapter discusses the various methods for linking the five regions<sup>1</sup> of the International Comparison Program (ICP) and those countries in the Eurostat-OECD PPP Programme in such a way that the relative country volumes within each region are preserved.<sup>2</sup>

The 2005 ICP round of international comparisons was quite different from earlier ICP rounds for a number of reasons:

- Each of the six regions prepared its own list of representative (for its region) products. These products were priced by each country in the region over the reference year, whereas in previous ICP comparisons there was a single product list for all countries over all regions.
- Each region was allowed to use its preferred multilateral method to make comparisons within the region. Five of the six regions chose to use the Gini-Éltető-Köves-Szulc (GEKS) method, and the Africa region used the additive Iklé-Dikhanov-Balk (IDB) method.<sup>3</sup>
- Each region was also allowed to use slightly different methods for constructing basic heading purchasing power parities (BH PPPs) across countries in the region.<sup>4</sup>
- To link the regions, another separate product list was developed, and these products were priced by 18 *Ring countries* of which two or more were in each region.<sup>5</sup> This additional set of prices across regions allowed the 155 BH PPPs to be linked across regions.
- At the final stage of aggregation of the 2005 ICP (the subject of this chapter), the within-region aggregate volumes for each country in a region were linked across the six regions, giving rise to a complete set of 146 aggregate PPPs and relative volumes for each country.

The final stage in the process, which involved linking the regions, proved to be the most challenging. The various regions involved in the 2005 ICP wanted the final linking procedure to leave unchanged the relative volumes for the countries in their region—that is, they sought a so-called *fixity constraint* on the final step, which links the regional parities into a set of global parities. The

problem associated with linking the regions in a way that would preserve the within-region parities was not new to the 2005 ICP; Kravis, Heston, and Summers (1982, 117–23) and Heston (1986) discussed this problem many years ago.

Diewert (2004, 46–47) suggested a class of methods for linking the regions in the 2005 ICP round that would preserve relative volumes for country aggregates within a region, but would at the same time link the various regions so that volumes could be compared across all countries in the ICP comparison in a consistent manner—that is, his suggested interregional linking method would satisfy the fixity constraint. The basic idea behind the method is as follows. Once information on country expenditures for each basic heading (BH) category in each country in a region is determined, along with the 155 corresponding BH PPPs, then the BH expenditures in the local currencies can be divided by the corresponding PPPs to obtain country volumes in consistent comparable units by basic heading category. These volumes (or loosely speaking, quantities) can then be added across countries within the region in order to obtain *total regional volumes* or “quantities” by basic heading category. Then, with an appropriate choice of BH prices for each region, these BH regional “prices” and “quantities” can be linked using any multilateral method such as the GEKS method (see chapter 5).<sup>6</sup> Once the regional real expenditure shares have been determined by the chosen multilateral method, they can be combined with each region’s country shares of regional real expenditures to give each country’s share of world real expenditures. Note that this method will not affect the country shares of regional expenditures that are determined independently by each region. Note also that this method is very similar to the methods used by national income accountants to determine the annual real output of a nation by aggregating over quarterly data.

The problem with this class of methods for linking the regions is determining exactly how the *regional* basic heading “prices” should be chosen. These regional “price” or PPP vectors should be chosen so that the overall method leads to country (and regional) relative volumes that are independent of the choice of both the numeraire region and the numeraire countries within each region. Section 6.1 of this chapter explores several options for the regional BH prices. Option 1 does lead to relative volumes that are independent of the choice of numeraire regions and numeraire countries within each region, but the resulting method involves the use of country exchange rates, which is not a desirable feature. Option 2 does not involve the use of exchange rates, but the resulting method is dependent on the choice of the regional numeraire countries, which again is not a desirable feature.<sup>7</sup> Finally, option 3, which was suggested by Sergeev (2009b), does lead to relative volumes that are independent of the choice of numeraire regions and numeraire countries within each region, and this method for linking the regions does not involve the use of exchange rates.

All of the variants of the method just described for linking the regions are based on two important properties within each region: (1) volumes are added up across countries in the region to obtain total regional volumes by BH category of expenditure; and (2) a single set of BH reference PPPs or “prices” is applied to those regional volumes for each region.

As mentioned earlier, the first property is analogous to practices in the *System of National Accounts* in which annual quantities by expenditure category are obtained by adding up quarterly or monthly quantities. Therefore, this aspect of the linking method is not necessarily problematic. However, in the national accounting context the structure of prices tends to be very similar from quarter to quarter within a year. Thus in the national accounts, the annual price for a commodity is usually taken to be an annual unit value—that is, the total expenditure or revenue for the commodity is summed up over, say, quarters and divided by the total quantity used or produced during the year, and this unit value is used as the annual price of the commodity.<sup>8</sup> Using an annual unit value price in the context of aggregating over time is generally not a problem, because the price of a commodity usually does not vary too much within a year. However, when one is aggregating

quantities across countries within a region, the variations in commodity prices can be very large, and, generally speaking, no single vector of commodity prices will be representative of the structure of commodity prices for every country in the region. Thus the property of these methods in which a regional quantity vector is calculated by addition over country quantity vectors means that all of these methods are essentially *additive methods* (within each region) and hence subject to *substitution bias*.<sup>9</sup> For the 2011 ICP, then, other methods for linking the region should be considered, but while respecting within-region parities.

Sections 6.2, 6.3, and 6.4 of this chapter consider alternative linking methods. Section 6.2 describes the first alternative method for constructing interregional parities while respecting within-region parities—see Heston (1986, 3) and Dikhanov (2007). Heston suggested using the Geary-Khamis (GK) method, whereas Dikhanov suggested that the GEKS method be used in a single comparison over all countries in the comparison. It will generate shares of world real expenditures for each country. One can then simply add up the shares of world real expenditures of the countries in each region in order to obtain the regional shares. These between-region shares plus the within-region shares (determined independently by the regions) will generate an overall set of shares that respect the within-region parities.

Section 6.3 discusses a variant of the method used in section 6.2. Instead of relying on the GEKS method to generate individual country shares for each country in the comparison, one could use Robert J. Hill's minimum spanning tree (MST) or similarity or spatial linking method to form the shares of each country's output in the world aggregate.<sup>10</sup> As noted in chapter 5, the basic idea behind this method is to link countries through a series of bilateral comparisons in which the overall comparison rests on a chain of bilateral comparisons between countries that are most similar in their (relative) price structures.

Section 6.4 of this chapter discusses a “new” method for linking the regions while respecting the within-region parities.<sup>11</sup> In section 6.5, regional parities are computed for a small artificial data set using the methods suggested in sections 6.1–6.4. In section 6.6, another numerical example based on a subset of the 1985 ICP data is computed to again illustrate the fact that different methods can give quite different results.

The methods suggested in this chapter that could be used to form aggregate PPPs for the 146 countries in the comparisons (over the entire set of 155 basic heading product groups) while respecting the fixity constraint within regions could also be used to form PPPs for *subsets* of the 155 BH product groups—for example, the same techniques could be used to form a set of consumption PPPs for the 146 countries.

## 6.1 Variants of Diewert's Suggested Method for Linking the Regions

This section begins by defining what basic data are needed to link the ICP regions in a fashion that will respect the parities and relative volumes determined within each region. It is assumed that  $R$  regions are in the comparison, and region  $r$  is made up of  $C(r)$  countries for  $r = 1, 2, \dots, R$ . In the 2005 ICP,  $R$  was equal to 6, with varying numbers of countries in each region. Assume there are  $N$  basic heading commodity groups and each country in the comparison has collected expenditure data on these  $N$  commodity groups in its own currency. In the 2005 ICP,  $N$  was equal to 155. Let  $E_{rcn}$  be the expenditure (in the currency of country  $c$  in region  $r$ ) on commodity class  $n$  for the reference year for  $r = 1, \dots, R$ ;  $c = 1, \dots, C(r)$ ;  $n = 1, \dots, N$ . Assume that these country expenditure data have been collected.

Also assume that each region  $r$  has constructed a purchasing power parity for each BH commodity group  $n$  and each country  $c$  in region  $r$ ,  $\alpha_{rcn}$  say, for  $r = 1, \dots, R$ ;  $c = 1, \dots, C(r)$ , and  $n = 1, \dots, N$ . Assume that country  $c = 1$  is the numeraire country in each region, and so the PPP for this numeraire country is set equal to unity so that

$$(6.1) \quad \alpha_{r1n} = 1; \quad r = 1, \dots, R; \quad n = 1, \dots, N.$$

Thus for  $c \neq 1$ ,  $\alpha_{rcn}$  is the price in the currency of country  $c$  in region  $r$  of a bundle of the products in BH commodity group  $n$ , which costs one currency unit in the currency of the numeraire country (country 1) for region  $r$ .

The next assumption is that the central office has constructed a set of *interregional linking PPPs*,  $\beta_m$ , that link the PPPs of the numeraire country in each region with the numeraire country in the numeraire region, which is assumed to be region 1. Thus assume that

$$(6.2) \quad \beta_{1n} = 1; \quad n = 1, \dots, N.$$

For  $r \neq 1$ ,  $\beta_m$  is the price in the currency of country 1 in region  $r$  of a bundle of commodity  $n$ , which costs one currency unit in the currency of the numeraire country (country 1) in region 1, the numeraire region.

Now the two sets of PPPs can be multiplied to form a consistent set of *world basic heading PPPs*,  $\gamma_{rcn}$ , defined as

$$(6.3) \quad \gamma_{rcn} \equiv \beta_m \alpha_{rcn}; \quad r = 1, \dots, R; \quad c = 1, \dots, C(r); \quad n = 1, \dots, N.$$

Thus  $\gamma_{rcn}$  is the price in the currency of country  $c$  in region  $r$  of a bundle of BH commodity group  $n$ , which costs one currency unit in the currency of the numeraire country (country 1) in the numeraire region (region 1).

If the numeraire region is changed, what happens to the world BH PPPs,  $\gamma_{rcn}$ , defined by (6.3)? Suppose region 2 replaces region 1 as the numeraire region. In this case, the original linking BH PPPs,  $\beta_m$ , should be replaced by the new linking PPPs, defined as

$$(6.4) \quad \beta_m^* \equiv \frac{\beta_m}{\beta_{2n}}; \quad r = 1, \dots, R; \quad n = 1, \dots, N$$

and the original set of world BH PPPs,  $\gamma_{rcn}$ , defined by (6.3) should be replaced by the following new set of world BH PPPs:

$$(6.5) \quad \gamma_{rcn}^* \equiv \beta_m^* \alpha_{rcn} = \frac{\gamma_{rcn}}{\beta_{2n}}; \quad r = 1, \dots, R; \quad c = 1, \dots, C(r); \quad n = 1, \dots, N.$$

Thus the new set of world BH PPPs is equal to the old set of world BH PPPs except for a divisor that depends on the BH commodity  $n$  under consideration. Hence the net effect of switching the numeraire region is to leave the PPPs unchanged except that the new set of PPPs is measured in a *new system of units*. Instead of measuring commodity units in terms of a dollar's worth of purchases of BH commodity  $n$  in the numeraire country of region 1, commodity units are now measured in terms of a peso's worth of purchases of BH commodity  $n$  in the numeraire country of region 2.

Now suppose that the numeraire country in a region is changed. If the change of numeraire country is outside region 1 (say, in region 2 where there is a switch from country 1 to country 2 as the numeraire country), then nothing will happen to the world BH PPPs,  $\gamma_{rcn}$ . The old  $\gamma_{rcn} = \beta_{rn} \alpha_{rcn}$  will remain the same for all regions  $r$  except for  $r = 2$ , and for the second region  $\alpha_{2cn}$  will be replaced by  $\frac{\alpha_{2cn}}{\alpha_{22n}}$  for  $n = 1, \dots, N$  and  $c = 1, \dots, C(2)$ , and  $\beta_{2n}$  will be replaced by  $\beta_{2n} \alpha_{22n}$  for  $n = 1, \dots, N$ . Thus the new  $\gamma_{rcn}^*$  will still equal the old  $\gamma_{rcn}$  for all  $r$ , including  $r = 2$ .

If the change of numeraire country is within region 1 so that the new numeraire country in region 1 is country 2 in place of country 1, then the within-region parities for region 1 become  $\frac{\alpha_{1cn}}{\alpha_{12n}}$  for  $n = 1, \dots, N$ , and the new set of interregional linking BH PPPs becomes  $\frac{\beta_{2n}}{\alpha_{12n}}$  for  $n = 1, \dots, N$ . Thus in this case, the new set of world BH PPPs becomes

$$(6.6) \quad \gamma_{rcn}^* = \frac{\gamma_{rcn}}{\alpha_{12n}}; \quad r = 1, \dots, R; c = 1, \dots, C(r); n = 1, \dots, N.$$

Thus in this change of the numeraire country within the numeraire region, the new set of world BH PPPs is equal to the old set of world BH PPPs except for a divisor that depends on the BH commodity  $n$  under consideration. Hence as it is for the change in the numeraire region, the net effect of switching the numeraire country within the numeraire region is to leave the BH PPPs unchanged except that the new set of BH PPPs is measured using a *new system of units*.

Recall that a knowledge of the expenditures (in the local currencies) of country  $c$  in region  $r$  by BH commodity class  $n$ ,  $E_{rcn}$ , is assumed for  $r = 1, \dots, R; c = 1, \dots, C(r); n = 1, \dots, N$ . This expenditure information is used, along with the consistent set of world BH PPPs,  $\gamma_{rcn}$  as defined by (6.3), to define consistent (across countries) *volumes* or *imputed quantities*,  $Q_{rcn}$ , for each BH commodity group  $n$  and each country  $c$  in each region  $r$  as

$$(6.7) \quad Q_{rcn} \equiv \frac{E_{rcn}}{\gamma_{rcn}}; \quad r = 1, \dots, R; c = 1, \dots, C(r); n = 1, \dots, N.$$

These imputed quantities or volumes are measured in units that are comparable across countries and regions. Thus these “quantities” can be added across countries within a region. The resulting regional totals,  $Q_{rn}$ , are also comparable across regions—that is, they define *regional total volumes by commodity class* as

$$(6.8) \quad Q_{rcn} \equiv \sum_{c=1}^{C(r)} Q_{rcn}; \quad r = 1, \dots, R; n = 1, \dots, N.$$

Now form *regional volume* or *imputed quantity* vectors from  $Q_{rn}$ :

$$(6.9) \quad Q^r \equiv [Q_{r1}, \dots, Q_{rN}]; \quad r = 1, \dots, R.$$

Note that if the numeraire region or the numeraire countries within a region are changed, then, if the world BH PPPs are changed in a consistent manner, the regional volume vectors will be identical to the initial regional volume vectors defined by (6.7)–(6.9), except that the units of measurement for these vectors may have been changed by the change in numeraires. In other words, a change in the numeraire region or numeraire country within a region will lead to identical regional quantity vectors  $Q^r$  or to a new set of  $Q^{r*}$  that is equal to  $\hat{\delta} Q^r$  where  $\hat{\delta}$  is a diagonal matrix in which the diagonal elements reflect changes in the units of measurement of the  $N$  commodity groups.

Regional BH PPP or “price” vectors  $P^r$  are now needed to match up with the regional volume vectors  $Q^r$  defined by (6.9). At this point, there are at least three possible strategies, and each is discussed in turn in the sections that follow.

### Option 1: Conversion to Regional Currency Unit Values

The strategy using this option would be to first convert country expenditures in each BH commodity classification category into common regional expenditures (in a numeraire country’s currency) using market exchange rates for the reference year. Then, deflate these regional commodity expenditures by the corresponding regional quantities defined by (6.8) in order to form regional unit value prices, which will be used as the regional prices. As noted at the outset of this chapter, this is the same type of strategy used by national income accountants in forming annual price and quantity vectors from subannual information.

Suppose the *reference year exchange rate* for country  $c$  in region  $r$  is  $\varepsilon_{rc}$  for  $r = 1, \dots, R$  and  $c = 1, \dots, C(r)$ . Assume that country 1 in each region is the numeraire region as usual so that

$$(6.10) \quad \varepsilon_{r1} = 1; \quad r = 1, \dots, R.$$

Thus  $\varepsilon_{rc}$  for  $c \neq 1$  indicates how many units of the currency of the numeraire country (country 1) in region  $r$  is equal to one unit of the currency unit of country  $c$  in region  $r$ . These market exchange rates can be used to convert country expenditures (in the country’s currency) on commodity class  $n$  within region  $r$  into region  $r$  numeraire currency units  $V_{rcn} \equiv \varepsilon_{rc} E_{rcn}$ , and then these country expenditures on BH commodity class  $n$  in a common regional currency can be summed to regional totals  $V_m$  defined as

$$(6.11) \quad V_m \equiv \sum_{c=1}^{C(r)} \varepsilon_{rc} E_{rcn}; \quad r = 1, \dots, R; \quad n = 1, \dots, N.$$

Now *regional PPPs* or *unit value prices*  $P_m$  can be formed by taking the regional values defined by (6.11) and dividing them by the corresponding regional volume or quantity totals  $Q_m$  defined by (6.8) so that

$$(6.12) \quad P_m \equiv \frac{V_m}{Q_m}; \quad r = 1, \dots, R; \quad n = 1, \dots, N.$$

Finally, form *regional unit value price vectors* from the components  $P_m$  so that

$$(6.13) \quad P^r \equiv [P_{r1}, \dots, P_{rN}]; \quad r = 1, \dots, R.$$

Now consider what happens to the regional total expenditures on commodity class  $n$ ,  $V_m$ , if a numeraire country within a region is changed or there is a change in the numeraire region. If the numeraire region is changed, nothing happens to regional expenditures defined by (6.11) because it is not necessary to relate the regional exchange rates across regions. However, if the numeraire country within region  $r$  is changed, then it can be seen that all the regional values for this region will change by a scalar factor—that is, if in region  $r$  the numeraire country is changed from country 1 to country 2, then the new market exchange rates will be  $\frac{\varepsilon_{rc}}{\varepsilon_{r2}}$  for  $c = 1, \dots, C(r)$ , and hence the new regional totals for region  $r$  will be

$$(6.14) \quad V_m^* \equiv \sum_{c=1}^{C(r)} \left[ \frac{\varepsilon_{rc}}{\varepsilon_{r2}} \right] E_{rcn} = \frac{V_m}{\varepsilon_{r2}}, \quad n = 1, \dots, N.$$

Thus the new regional expenditure totals on the  $N$  commodity groups in region  $r$ ,  $V_{rcn}^*$ , will be equal to a scalar multiple  $\left(\frac{1}{\varepsilon_{r2}}\right)$  times the old regional expenditure totals in region  $r$ ,  $V_{rm}$ .

This is a careful exposition of the regional unit value method for forming regional price and quantity vectors that could be used to link the regions by relying on a multilateral index number method with the regional price and quantity vectors as the input vectors to the method.

Now consider linking the  $R$  regions using a *multilateral index number method* (see chapter 5). Diewert (1999) defined such methods in terms of share functions for the  $R$  regions—that is, he looked at a system of *shares of world output or real expenditures*,  $s_1(P^1, \dots, P^R; Q^1, \dots, Q^R), \dots, s_R(P^1, \dots, P^R; Q^1, \dots, Q^R)$ , where  $s_r(P^1, \dots, P^R; Q^1, \dots, Q^R)$  is region  $r$ 's share of world output and considered the axiomatic properties of such multilateral systems. The next question to be addressed: what properties does the multilateral method have to satisfy to ensure that the regional shares are independent of the choice of the numeraire region and the numeraire countries within the regions?

From the previous discussion, it is clear that the multilateral method must satisfy the following two properties:

*Property 1: Invariance to changes in the units of measurement.*

Thus let  $\delta_n > 0$  for  $n = 1, \dots, N$  and define  $\delta$  as the  $N \times N$  matrix with the elements  $\delta_n$  running down the main diagonal. Then this property requires that the multilateral share system satisfy the following equations:

$$(6.15) \quad s_r(\hat{\delta}P^1, \dots, \hat{\delta}P^R; \hat{\delta}^{-1}Q^1, \dots, \hat{\delta}^{-1}Q^R) = s_r(P^1, \dots, P^R; Q^1, \dots, Q^R); \quad r = 1, \dots, R.$$

*Property 2: Homogeneity of degree zero in the regional price vectors.*

Let  $\lambda_1 > 0, \dots, \lambda_R > 0$ . Then this property requires that the multilateral share system satisfy the following equations:

$$(6.16) \quad s_r(\lambda_1 P^1, \dots, \lambda_R P^R; Q^1, \dots, Q^R) = s_r(P^1, \dots, P^R; Q^1, \dots, Q^R); \quad r = 1, \dots, R.$$

This homogeneity property means that it is the relative regional prices that affect the interregional volume comparisons and not the absolute level of regional prices.

These two properties are also sufficient to imply that a multilateral method using the unit value regional price and quantity vectors defined earlier will generate regional shares that are independent of the choice of the numeraire region and the numeraire countries within the regions. The GEKS multilateral method satisfies these two properties (see Diewert 1999, 33).

A drawback of this method for linking the regions is that it brings market exchange rates into the picture. Because these rates are often far removed from their corresponding PPPs, it is desirable to avoid their use in constructing the interregional PPPs and relative volumes. The following two methods make use of the regional “quantity” vectors  $Q^r$  defined by (6.8), but the corresponding regional “price” vectors  $P^r$  do not use exchange rates.

## Option 2: Use of Regional Numeraires as Regional Price Weights

It is not necessary to bring in market exchange rates to convert regional expenditures into a common currency. Instead, one can simply use the PPPs for each numeraire country 1 in each region  $r$



(relative to the numeraire country in the numeraire region,  $\beta_m$ , as described earlier) as the price for commodity  $n$  in region  $r$ —that is, one can define the regional price for commodity  $n$  in region  $r$  as

$$(6.17) \quad P_m \equiv \beta_m; \quad r = 1, \dots, R; n = 1, \dots, N.$$

Then use (6.13) to form the *regional price* or *PPP vectors*  $P^r$  in the usual way. The regional *total volume* or *quantity vectors*,  $Q^r$ , are defined as before by (6.7)–(6.9), and these equations do not involve exchange rates. Finally, use these regional price and quantity vectors,  $P^r$  and  $Q^r$ , to calculate the share functions for the  $R$  regions using a favored multilateral method,  $s_1(P^1, \dots, P^R; Q^1, \dots, Q^R), \dots, s_R(P^1, \dots, P^R; Q^1, \dots, Q^R)$ , where  $s_r(P^1, \dots, P^R; Q^1, \dots, Q^R)$  is region  $r$ 's share of world output.

Unfortunately, as Sergeev (2009b) points out, this method is not invariant to the choice of the numeraire countries within the regions. Thus this method should not be used in the 2011 ICP.

### Option 3: Use of the Geometric Average of the Regional Numeraires as Regional Price Weights

Sergeev (2009b) suggested a way to avoid the lack of numeraire invariance in option 2: within each region take the *geometric mean of the country parities* over all countries in the region. Thus (6.17) is replaced by<sup>12</sup>

$$(6.18) \quad P_m \equiv \prod_{c=1}^{C(r)} \gamma_{rcn}^{\frac{1}{C(r)}}; \quad r = 1, \dots, R; n = 1, \dots, N.$$

Then use (6.13) to form the *regional price vectors*,  $P^r$ , in the usual way. Finally, use the regional “price” and “quantity” vectors defined by (6.9),  $P^r$  and  $Q^r$ , to calculate the share functions for the  $R$  regions using a favored multilateral method,  $s_1(P^1, \dots, P^R; Q^1, \dots, Q^R), \dots, s_R(P^1, \dots, P^R; Q^1, \dots, Q^R)$ , where  $s_r(P^1, \dots, P^R; Q^1, \dots, Q^R)$  is region  $r$ 's share of world output. If the multilateral method satisfies property 1 (invariance to changes in the units of measurement), then the resulting regional shares will be invariant to the choices of both numeraire countries and the numeraire region.

As noted earlier, option 2 can be ruled out as a method for linking the regions for the 2011 ICP because of its lack of country numeraire invariance within the regions. However, there is a good case for ruling out options 1 and 3 as well because all the methods in this section *impose a set of common prices to add up the volumes or quantities within a region*—that is, the methods impose a form of *additivity*. Additive multilateral methods are subject to *substitution bias* if three or more countries are in the comparison.<sup>13</sup>

Thus the two sections that follow discuss *nonadditive methods* for linking the regions that avoid the substitution bias inherent in the methods discussed in this section.

## 6.2 The Global Comparison GEKS Method

Recall the definitions in (6.7) that defined the volume or “quantity” of commodity  $n$ ,  $Q_{rcn}$ , that was finally demanded by purchasers in country  $c$  in region  $r$ . Define country  $c$  in the region  $r$  volume vector,  $Q^r$ , in the usual way, using the definitions in (6.7) to define the components  $Q_{rcn}$ :

$$(6.19) \quad Q^r \equiv [Q_{rc1}, \dots, Q_{rcN}]; \quad r = 1, \dots, R; c = 1, \dots, C(r).$$



The country “price” vectors  $P^{rc}$  that correspond to the country “quantity” vectors  $Q^{rc}$  defined by (6.19) are in turn defined using the global set of PPPs,  $\gamma_{rcn}$ , defined by (6.3). Recall that  $\gamma_{rcn}$  is the price in the currency of country  $c$  in region  $r$  of a bundle of commodity  $n$ , which costs one currency unit in the currency of the numeraire country (country 1) in the numeraire region (region 1). The basic heading price vector for country  $c$  in region  $r$ ,  $P^{rc}$ , is defined as

$$(6.20) \quad P^{rc} \equiv [\gamma_{rc1}, \dots, \gamma_{rcN}]; \quad r = 1, \dots, R; c = 1, \dots, C(r).$$

Thus there are country “price” and “quantity” vectors ( $P^{rc}$ ,  $Q^{rc}$ ) for all  $C(1) + C(2) + \dots + C(R)$  countries in the ICP.

At this stage, any multilateral method could be used to form price comparisons between each individual country participating in the ICP. One such multilateral method with good axiomatic and economic properties is the GEKS.<sup>14</sup> The algebra for this method works as follows. First, define the *Fisher quantity index* (Fisher 1922) for country  $c$  in region  $r$  relative to country  $d$  in region  $s$  as<sup>15</sup>

$$(6.21) \quad Q_F \left( \frac{rc}{sd} \right) \equiv \left[ \frac{P^{rc} \cdot Q^{rc} P^{sd} \cdot Q^{rc}}{P^{rc} \cdot Q^{sd} P^{sd} \cdot Q^{sd}} \right]^{\frac{1}{2}}; \quad r = 1, \dots, R; c = 1, \dots, C(r); \\ s = 1, \dots, R; d = 1, \dots, C(s)$$

where  $P^{rc} \cdot Q^{rc} \equiv \sum_{n=1}^N P_{rcn} Q_{rcn}$  denotes the inner product between the vectors  $P^{rc}$  and  $Q^{rc}$ . If the base country is fixed (i.e., fix region  $s$  and country  $d$  in region  $s$ ) and if  $r = 1, \dots, R$  and  $c = 1, \dots, C(r)$ , then the Fisher indexes defined by (6.21) can be interpreted as the volume of each country  $rc$  in the comparison relative to the base country  $sd$ , and then these relative volumes can be normalized into a set of shares of world product using country  $sd$  as the base country—that is, a set of country “star” shares are obtained for each  $rc$  with country  $sd$  as the “star” country.<sup>16</sup> The GEKS method is then used to take the geometric mean of all of these country parities over all possible “star” bases. Thus define these geometric mean relative parities as

$$(6.22) \quad Q(rc) \equiv \left[ \prod_{s=1}^R \prod_{d=1}^{C(s)} Q_F \left( \frac{rc}{sd} \right) \right]^{\frac{1}{[C(1) + \dots + C(R)]}}; \quad r = 1, \dots, R; c = 1, \dots, C(r).$$

Now normalize the relative parities given by (6.22) into country shares of world product. Thus define the sum of the parities defined by (6.22) as  $\sigma$ :

$$(6.23) \quad \sigma \equiv \sum_{r=1}^R \sum_{c=1}^{C(r)} Q(rc).$$

The GEKS share of world real final expenditure for country  $c$  in region  $r$  can now be defined as  $Q(rc)$  divided by  $\sigma$ :

$$(6.24) \quad s_{rc} \equiv \frac{Q(rc)}{\sigma}; \quad r = 1, \dots, R; c = 1, \dots, C(r).$$

Following Heston (1986) and Dikhanov (2007),<sup>17</sup> one can aggregate over the individual country shares of world product defined by (6.24) within each region to obtain the following *GEKS regional shares of world output*:

$$(6.25) \quad S_r \equiv \sum_{c=1}^{C(r)} s_{rc}; \quad r = 1, \dots, R.$$

The between-region shares  $S_1, S_2, \dots, S_R$  defined by (6.25) can be used in conjunction with the within-region shares in each region to obtain a system of world product shares for each country. The resulting parities will respect the regional parities that are independently determined by the regions.

Why should the GEKS regional shares defined by (6.25) be preferred over the various regional shares defined in the previous section? The GEKS shares are consistent with broader patterns of substitutability between commodities—that is, if the preferences of each country can be represented by certain homothetic preferences (that can approximate arbitrary homothetic preferences to the second order), then the GEKS country shares will give exactly the “right” relative volumes across countries.<sup>18</sup>

The following section describes a variant of the method used in this section.

### 6.3 Spatial Comparisons Based on Similar Price Structures

The GEKS multilateral method treats each country “star” parity as equally valid, and hence an averaging of the parities is appropriate under this hypothesis. However, are all bilateral comparisons of volume between two countries equally accurate? One could argue that the answer to this question is yes if the relative prices in countries A and B are very similar. Then the Paasche and Laspeyres quantity indexes will be very close, and therefore it is likely that the “true” volume comparison between these two countries (using the economic approach to index number theory) will be very close to the Fisher volume comparison. On the other hand, if the structure of relative prices in the two countries is very different, it is likely that the structure of relative quantities in the two countries will also be different. Then the Paasche and Laspeyres quantity indexes will likely differ considerably, and one can no longer be certain that the Fisher quantity index will be close to the “true” volume comparison. These considerations suggest that a more accurate set of world product shares could be constructed if an initial bilateral comparison were made between the two countries that have the most similar relative price structures. Then look for a third country that has the price structure most similar to those of the first two countries and link this third country to the comparisons of volume between the first two countries and so on. At the end of this procedure, Hill’s *minimum spanning tree* would be constructed: a path between all countries that minimizes the sum of the relative price similarity measures. A key aspect of this methodology is the choice of the measure of similarity (or dissimilarity) of the relative price structures of two countries. Various measures of the similarity or dissimilarity of relative price structures have been proposed by Aten and Heston (2009), Diewert (2009), Hill (2009), and Sergeev (2001, 2009a). Diewert (2009, 207) suggested the following *weighted log quadratic (WLQ) measure of relative price dissimilarity*,  $\Delta_{\text{WLQ}}(p^1, p^2, q^1, q^2)$  (the smaller the measure, the more similar is the structure of relative prices between the two countries):

$$(6.26) \quad \Delta_{\text{WLQ}}(p^1, p^2, q^1, q^2) \equiv \sum_{n=1}^N \left( \frac{1}{2} \right) (s_n^1 + s_n^2) \left[ \ln \left( \frac{p_n^2}{p_n^1 P_F(p^1, p^2, q^1, q^2)} \right) \right]^2$$

where  $P_F(p^1, p^2, q^1, q^2) \equiv \left[ \frac{(p^2 \cdot q^1 p^2 \cdot q^2)}{(p^1 \cdot q^1 p^1 \cdot q^2)} \right]^{\frac{1}{2}}$  is the Fisher ideal price index between countries 2 and 1, and  $s_n^c \equiv \frac{p_n^c \cdot q_n^c}{p^c \cdot q^c}$  is the country  $c$  expenditure share on basic heading category  $n$  for  $c = 1, 2$  and  $n = 1, \dots, N$ .

Diewert (2009, 2008) also suggested the *weighted asymptotically quadratic (WAQ) measure of relative price dissimilarity*,  $\Delta_{\text{WAQ}}(p^1, p^2, q^1, q^2)$ :<sup>19</sup>

$$(6.27) \quad \Delta_{\text{WAQ}}(p^1, p^2, q^1, q^2) \equiv \sum_{n=1}^N \left( \frac{1}{2} \right) (s_n^1 + s_n^2) \left\{ \left[ \frac{p_n^2}{p_n^1 P_F(p^1, p^2, q^1, q^2)} - 1 \right]^2 + \left[ \frac{P_F(p^1, p^2, q^1, q^2) p_n^1}{p_n^2} - 1 \right]^2 \right\}.$$

If prices are proportional for the two countries so that  $p^2 = \lambda p^1$  for some positive scalar  $\lambda$ , then  $P_F(p^1, p^2, q^1, q^2) = \lambda$ , and the measures of relative price dissimilarity defined by (6.26) and (6.27) will equal its minimum of 0. Thus the smaller is  $\Delta_{\text{WLQ}}(p^1, p^2, q^1, q^2)$  or  $\Delta_{\text{WAQ}}(p^1, p^2, q^1, q^2)$ , the more similar is the structure of relative prices in the two countries. These two measures of price dissimilarity were the measures of relative price dissimilarity preferred by Diewert (2009), and they will be used in sections 6.5 and 6.6 in numerical examples to illustrate the MST method for making multilateral comparisons. The following measure of relative price dissimilarity is used in the numerical examples as well; it was also used by Kravis, Heston, and Summers (1982, 105) and Hill (1999a, 1999b, 2001, 2004). It is essentially a normalization of the relative spread between the Paasche and Laspeyres price indexes, and so it is known as the *Paasche-Laspeyres spread (PLS) relative price dissimilarity measure*:

$$(6.28) \quad \Delta_{\text{PLS}}(p^1, p^2, q^1, q^2) \equiv \max \left\{ \frac{P_L}{P_P}, \frac{P_P}{P_L} \right\} - 1$$

where  $P_L \equiv \frac{p^2 \cdot q^1}{p^1 \cdot q^1}$  and  $\frac{P_P}{P_L} \equiv \frac{p^2 \cdot q^2}{p^1 \cdot q^2}$ . As noted in chapter 5, a major problem with this measure of relative price dissimilarity is that it is possible for  $P_L$  to equal  $P_P$  and yet  $p^2$  could be very far from being proportional to  $p^1$ .

As noted, Hill's MST method of similarity or spatial linking using the three measures of relative price dissimilarity just described will be illustrated in the sections 6.5 and 6.6. using small numerical examples. Basically, instead of using the GEKS country shares  $s_{rc}$  defined by (6.24) in the previous section, the shares generated by the minimum spanning tree (relying on the three dissimilarity measures) are used to link all of the countries in the ICP. Once these country shares  $s_{rc}$  have been defined, again use equations (6.25) in order to form the regional shares  $S_r$ .

The next section describes yet another method proposed for linking the regions.

## 6.4 A Least Squares Method for Linking the Regions

Recently, another method for linking the regions while respecting regional parities was proposed by Robert J. Hill (2010). In order to explain how this method works, it is useful to define some new notation that is related to the notation used at the beginning of this chapter. As in section 6.1, it is assumed that there are  $R$  regions, and that region  $r$  is made up of  $C(r)$  countries for  $r = 1, 2, \dots, R$ . It is also assumed that each region  $r$  has constructed a purchasing power parity for a national accounts aggregate such as consumption or aggregate final demand. Within each region  $r$ , the country  $c$  aggregate PPP is denoted by  $\alpha_{rc} > 0$ , say, for  $r = 1, \dots, R$  and  $c = 1, \dots, C(r)$ . Assume that country  $c = 1$  is the numeraire country in each region, and so the PPP for this numeraire country

is set equal to unity so that  $\alpha_{r1} = 1$  for  $r = 1, \dots, R$ . As usual, each region would like these relative PPPs to be respected in a world comparison of the same national accounts aggregate.

In addition to the regional PPPs  $\alpha_{rc}$ , Hill (2010) assumes that a global or worldwide comparison has been undertaken for the same aggregate using some method such as GEKS. The world PPP for country  $c$  in region  $r$  is denoted by  $\gamma_{rc} > 0$  for  $r = 1, \dots, R$ , and  $c = 1, \dots, C(r)$ . It is assumed that country 1 in region 1 is the numeraire country in the global comparison so that  $\gamma_{11} = 1$ .

It would be ideal if the global parities (the  $\gamma_{rc}$ ) were completely consistent with the regional parities, the  $\alpha_{rc}$ . This will be the case if there exists a set of *interregional parities*,  $\beta_r > 0$  for  $r = 1, \dots, R$  such that the following equations hold:

$$(6.29) \quad \gamma_{rc} = \beta_r \alpha_{rc}; \quad r = 1, \dots, R; \quad c = 1, \dots, C(r).$$

In general, it would not be possible to find  $\beta_1, \dots, \beta_R$  so that equations (6.29) hold. Thus Hill follows the example of Gini (1924) and the other founders of the GEKS method and chooses  $\beta_r$  in order to solve the following *least squares minimization problem*:

$$(6.30) \quad \min_{\beta_1, \dots, \beta_R} \sum_{r=1}^R \sum_{c=1}^{C(r)} [\ln \gamma_{rc} - \ln \beta_r - \ln \alpha_{rc}]^2.$$

The first-order conditions for the minimization problem (6.30) lead to the following solution for the interregional PPPs:

$$(6.31) \quad \beta_r^* = \prod_{c=1}^{C(r)} \left[ \frac{\gamma_{rc}}{\alpha_{rc}} \right]^{\frac{1}{C(r)}}; \quad r = 1, \dots, R.$$

Thus the interregional parity for region  $r$ ,  $\beta_r^*$ , is equal to the geometric mean of all the ratios  $\frac{\gamma_{rc}}{\alpha_{rc}}$  (the global PPP for country  $c$  in region  $r$ ,  $\gamma_{rc}$ , to the corresponding regional PPP for country  $c$  in region  $r$ ,  $\alpha_{rc}$ ) over all countries  $c$  in the  $r$ -th region. Looking at the minimization problem, one can see that each ratio  $\frac{\gamma_{rc}}{\alpha_{rc}}$  can be regarded as an estimate for  $\beta_r$ , and the overall estimator is the geometric mean of these country-specific estimators. Note the similarity between this method for linking the regions and Diewert's variant of the Country Product Dummy method for linking basic heading PPPs across regions while respecting regional BH parities.

Once the regional parities  $\beta_r^*$  have been determined by equations (6.31), the global PPPs that respect the within-region relative parities can be given by  $\beta_r^* \alpha_{rc}$  for  $r = 1, \dots, R$  and  $c = 1, \dots, C(r)$ .

What are the advantages and disadvantages of this method for linking the regions as opposed to the Heston-Dikhanov method? The Hill method has two advantages:

- It is a democratic method: each country  $c$  in region  $r$  contributes equally to the formation of the regional parity  $\beta_r^*$ .
- It is possible to work out standard errors for the regional parities  $\beta_r^*$ , treating each ratio  $\frac{\gamma_{rc}}{\alpha_{rc}}$  for  $c = 1, \dots, C(r)$  as an equal contributor to the overall geometric average that is equal to  $\beta_r^*$ .

However, the democratic nature of the Hill method could be regarded as a disadvantage as well: why should a tiny country in region  $r$  have the same weight as a very large country in the region in determining the regional parity  $\beta_r^*$ ? The Heston-Dikhanov method does not suffer from this problem: a large country in a region will contribute a large country share to the overall region's share of world product, whereas a small country will contribute only a small country share to the regional share.

The two small numerical examples that follow in sections 6.5 and 6.6 demonstrate how the different methods suggested in this and the previous three sections perform in practice. The first example uses an artificial data set, and the second example uses a small subset of the 1985 ICP data.

## 6.5 A Numerical Example Based on an Artificial Data Set

In this example, there are only four countries and two commodities. There are also two regions: region 1 consists of the first two countries, and region 2 consists of the second two countries (see table 6.1). The basic data for the countries are the within-region  $r$  basic heading PPP for commodity  $n$  for country  $c$  in region  $r$ ,  $\alpha_{rcn}$ ; the expenditure on commodity class  $n$  for country  $c$  in region  $r$  in domestic currency,  $E_{rcn}$ ; and the market exchange rate for country  $c$  in region  $r$ ,  $\varepsilon_{rc}$  (which does not depend on commodity  $n$ ), for  $r = 1, 2$ ;  $c = 1, 2$ ; and  $n = 1, 2$ . The basic heading interregional PPPs for the numeraire countries in each region,  $\beta_{rn}$ , are the parities for commodity  $n$  for region  $r$  relative to region 1.

Eight world basic heading PPPs are obtained using equations (6.3),  $\gamma_{rcn} \equiv \beta_{rn} \alpha_{rcn}$ . Thus for commodity class 1, the following world BH PPPs are obtained:  $\gamma_{111} = 1$ ;  $\gamma_{121} = 20$ ;  $\gamma_{211} = 2$ ;  $\gamma_{221} = 20$ . For commodity class 2, the following world BH PPPs are obtained:  $\gamma_{112} = 1$ ;  $\gamma_{122} = 4$ ;  $\gamma_{212} = 4$ ;  $\gamma_{222} = 16$ .

Working through the algebra in section 6.1, the option 1 shares of world output for regions 1 and 2 turn out to equal 0.45134 for region 1 and 0.54866 for region 2. As expected, the same regional shares are obtained no matter which region is chosen as the numeraire region and no matter which country is chosen as the numeraire country within a region.

The option 2 regional shares turn out to depend on the choice of the regional numeraire countries as expected. Letting country 1 in each region be the numeraire country, 0.45676 and 0.54324 are obtained as the two regional shares of the world product. Letting country 2 in region 1 be the numeraire country and maintaining country 1 in region 2 as the numeraire country results in 0.46287 and 0.53713 as the new regional shares of the world product. Letting country 1 in region 1 be the numeraire country and letting country 2 in region 2 be the numeraire country results in 0.46041 and 0.53959 as the new regional shares of the world product. The bottom line is that the option 2 regional shares are not invariant to the choice of the numeraire countries in the regions.<sup>20</sup>

The option 3 regional shares<sup>21</sup> (Sergeev option) turn out to be  $S_1 = 0.46186$  and  $S_2 = 0.53814$ . Recall that the option 1 regional shares were 0.45134 for region 1 and 0.54866 for region 2. Thus the option 1 and 3 regional shares differ by about 2.3 percent.

TABLE 6.1 Numerical Example: Four Countries in Two Regions

$n$	Region 1						Region 2						Regional basic heading parities	
	Country 1			Country 2			Country 1			Country 2			$\beta_{1n}$	$\beta_{2n}$
	$\alpha_{11n}$	$E_{11n}$	$\varepsilon_{11}$	$\alpha_{12n}$	$E_{12n}$	$\varepsilon_{12}$	$\alpha_{21n}$	$E_{21n}$	$\varepsilon_{21}$	$\alpha_{22n}$	$E_{22n}$	$\varepsilon_{22}$		
1	1	10	1	20	2	5	1	20	3	10	20	18	1	2
2	1	10	1	4	8	5	1	20	3	4	160	18	1	4

**TABLE 6.2** Star and GEKS Country Shares of World Product

	Region 1		Region 2	
	Country 1	Country 2	Country 1	Country 2
Region 1, country 1 star	0.43523	0.02879	0.30775	0.22823
Region 1, country 2 star	0.47078	0.03114	0.32258	0.17550
Region 2, country 1 star	0.42439	0.02897	0.30009	0.24655
Region 2, country 2 star	0.44331	0.04125	0.28296	0.23248
GEKS	0.44436	0.03226	0.30387	0.21951

Note: GEKS = Gini-Éltető-Köves-Szulc.

The regional shares for the methods explained in section 6.3 will now be calculated. The individual country shares of world output using each of the four countries as the “star” in the bilateral Fisher index number comparisons are listed in the first four rows of table 6.2. The table reveals some relatively large differences between these world shares, particularly for the small country 2 in region 1. The GEKS country shares of world product are listed in the last line of table 6.2.

The GEKS volume shares in the last row of table 6.2 are defined by equations (6.24)—that is, the entries for the last row of the table are the GEKS shares  $s_{11}$ ,  $s_{12}$ ,  $s_{21}$ , and  $s_{22}$ , respectively. Using these country shares, one can define the *regional GEKS shares*  $S_r$  by (6.25) so that, by using the entries in the last line of table 6.2,  $S_1$  and  $S_2$  are defined as

$$(6.32) \quad S_1 = s_{11} + s_{12} = 0.47662; \quad S_2 = s_{21} + s_{22} = 0.52338.$$

Thus the GEKS share of world output for region 1 is 0.47662, which is higher than the corresponding region 1 shares for the option 1 method (0.45134) and for the option 3 method (0.46479).

Finally, the MST or similarity or spatial linking method regional shares are calculated using the three dissimilarity measures defined by (6.26) (WLQ), (6.27) (WAQ), and (6.28) (PLS).

The first measure of relative price dissimilarity is the weighted log quadratic measure of relative price dissimilarity between countries 1 and 2,  $\Delta_{\text{WLQ}}(p^1, p^2, q^1, q^2)$ , as defined by (6.26). The  $4 \times 4$  matrix of relative price dissimilarity measures appears in table 6.3. For convenience in labeling the countries, set country 1 equal to country 1 in region 1, country 2 equal to country

**TABLE 6.3** Weighted Log Quadratic Relative Price Dissimilarities between Countries  $i$  and  $j$ 

	Country 1	Country 2	Country 3	Country 4
Country 1	0.00000	0.59465	0.12011	0.01057
Country 2	0.59465	0.00000	1.22741	0.25253
Country 3	0.12011	1.22741	0.00000	0.17933
Country 4	0.01057	0.25253	0.17933	0.00000

2 in region 1, country 3 equal to country 1 in region 2, and country 4 equal to country 2 in region 2.

Table 6.3 reveals that countries 1 and 4 have the most similar structures of relative prices, with a dissimilarity measure equal to 0.01057. The next pair of countries with the most similar structures of relative prices is 1 and 3; they have a dissimilarity measure equal to 0.12011. Thus countries 3 and 4 can be linked to country 1 using the bilateral Fisher quantity index between 4 and 1 and between 3 and 1. The next pair of countries most similar in structure is 3 and 4, with a dissimilarity measure equal to 0.17933. But these two countries are already linked, and so the next lowest measure of relative price dissimilarity is considered. The next most similar pair of countries is 2 and 4, with a dissimilarity measure equal to 0.25253. The bilateral Fisher quantity index is used to link country 2 to country 4, and so now all countries in the “world” have been linked using bilateral links. The volumes of all four countries relative to country 1 turn out to be 1.00000, 0.09305, 0.70711, and 0.52440. This result leads to the following vector of country shares of world output: 0.43019, 0.04003, 0.30419, and 0.22559. Adding up the shares of the countries in each region produces the following *spatially or similarity linked regional shares of world output*:

$$(6.33) \quad S_1 = s_{11} + s_{12} = 0.47022; S_2 = s_{21} + s_{22} = 0.52978.$$

Thus under the spatial linking method, 0.47022 is region 1’s estimated share of world output, compared with the Heston-Dikhanov-GEKS estimate of 0.47662 and the Sergeev option 3 estimate of 0.46479.

The second measure of relative price dissimilarity is the weighted asymptotically quadratic measure of relative price dissimilarity,  $\Delta_{\text{WAQ}}(p^1, p^2, q^1, q^2)$ , defined by (6.27). The corresponding  $4 \times 4$  matrix of relative price dissimilarity measures appears in table 6.4. The labeling of countries is the same as in table 6.3.

Table 6.4 shows that countries 1 and 4 have the most similar structures of relative prices, with a dissimilarity measure equal to 0.02135. The next pair of countries with the most similar structures of relative prices is 1 and 3 with a dissimilarity measure equal to 0.25736. The next pair of countries most similar in structure is 3 and 4, with a dissimilarity measure equal to 0.41607. But these two countries are already linked, so the next lowest measure of relative price dissimilarity is for countries 2 and 4, with a dissimilarity measure equal to 0.91013. The bilateral Fisher quantity index is used to link country 2 to country 4, and so now all countries in the “world” have been linked using bilateral links. One ends up with exactly the same minimum spanning tree produced using the weighted log quadratic measure of dissimilarity, and

**TABLE 6.4** Weighted Asymptotically Quadratic Relative Price Dissimilarities between Countries  $i$  and  $j$

	Country 1	Country 2	Country 3	Country 4
Country 1	0.00000	1.74216	0.25736	0.02135
Country 2	1.74216	0.00000	4.94541	0.91013
Country 3	0.25736	4.94541	0.00000	0.41607
Country 4	0.02135	0.91013	0.41607	0.00000



**TABLE 6.5** Paasche-Laspeyres Spread Relative Price Dissimilarities between Countries  $i$  and  $j$ 

	Country 1	Country 2	Country 3	Country 4
Country 1	0.00000	1.52000	0.12500	0.10000
Country 2	1.52000	0.00000	3.51000	0.13333
Country 3	0.12500	3.51000	0.00000	0.63333
Country 4	0.10000	0.13333	0.63333	0.00000

so all further analysis of this case is the same as that in the previous case. It is encouraging that Diewert's two most preferred measures of relative price dissimilarity give rise to exactly the same set of links.

The third measure of relative price dissimilarity is the Paasche-Laspeyres spread,  $\Delta_{PLS}(p^1, p^2, q^1, q^2)$ , defined by (6.28). The  $4 \times 4$  matrix of relative price dissimilarity measures generated by this measure appears in table 6.5. The countries are labeled in the same fashion as in table 6.3.

Working through the information in table 6.5, one can see that the minimum spanning tree using the Paasche-Laspeyres spread measure of dissimilarity is exactly the same as the MST using the previous two dissimilarity measures. Thus for this example all three dissimilarity measures generate the same set of bilateral linkages and hence the same estimates of country and region relative volumes.

A summary of the region 1 share of world output using the various methods appears in table 6.6.

The differences between the various methods are fairly substantial: a 4.2 percent difference in the share of region 1 for the highest share method (spatial linking with any of the three dissimilarity measures) compared with the lowest share method (regional unit values method).

The following section considers another numerical example based on actual ICP 1985 data that leads to even bigger numerical differences between the various methods.

**TABLE 6.6** Share of World Output for Region 1 Using Various Methods

	Region 1 share of world output
Option 1 (regional unit values method)	0.45134
Option 2: base countries: 1 in region 1; 1 in region 2	0.45676
Option 2: base countries: 2 in region 1; 1 in region 2	0.46287
Option 2: base countries: 1 in region 1; 2 in region 2	0.46041
Option 3: geometric mean average prices in each region	0.46186
Heston-Dikhanov-GEKS	0.47662
Spatial linking (weighted log quadratic dissimilarity)	0.47022
Spatial linking (weighted asymptotic quadratic dissimilarity)	0.47022
Spatial linking (Paasche-Laspeyres spread dissimilarity)	0.47022

Note: GEKS = Gini-Éltető-Köves-Szulc.

## 6.6 A Numerical Example Based on 1985 ICP Data

Yuri Dikhanov (1994 and 1997) generated some highly aggregated data (across basic heading groups) from the 1985 ICP on five consumption components for eight countries/economies:

- 1 = Hong Kong SAR, China
- 2 = Bangladesh
- 3 = India
- 4 = Indonesia
- 5 = Brazil
- 6 = Japan
- 7 = Canada
- 8 = United States.

Hong Kong SAR, China; Japan; Canada; and the United States can be considered to be “rich,” while Bangladesh, India, Indonesia, and Brazil can be considered to be “less rich.”

The five consumption categories are as follows:

- 1 = durables
- 2 = food, alcohol, and tobacco
- 3 = other nondurables, excluding food, alcohol, tobacco, and energy
- 4 = energy
- 5 = services.

The expenditure data (converted to U.S. dollars) and the volume or “quantity” data for the eight countries are listed in tables 6.7 and 6.8, respectively.

If the entries in table 6.7 (expenditures converted to U.S. dollars at market exchange rates) are divided by the entries in table 6.8 (quantities in comparable units), the basic heading prices (converted into U.S. dollars at market exchange rates) for each commodity class for each country are obtained. These prices are listed in table 6.9.

Thus the U.S. price level for each commodity group is set equal to 1, and the other prices are the average domestic prices for the commodity group converted into U.S. dollars at the 2005

**TABLE 6.7** Expenditures for Eight Countries/Economy and Five Consumption Categories

*US\$ millions*

Category	HK	BGD	IND	INDO	BRA	JPN	CAN	US
1	4,320	1,963	23,207	8,234	52,722	307,547	94,121	967,374
2	10,562	24,835	176,782	83,882	105,527	448,995	82,056	778,665
3	14,951	5,100	60,748	15,158	60,798	272,875	69,461	992,761
4	2,619	3,094	42,126	17,573	39,933	125,835	43,342	524,288
5	62,124	11,627	166,826	61,248	273,669	1,736,977	379,629	5,559,458

Source: Yuri Dikhanov, tabulations of 1985 ICP results.

Note: 1 = durables; 2 = food, alcohol, and tobacco; 3 = other nondurables, excluding food, alcohol, tobacco, and energy; 4 = energy; 5 = services. HK = Hong Kong SAR, China; BGD = Bangladesh; IND = India; INDO = Indonesia; BRA = Brazil; JPN = Japan; CAN = Canada; US = United States.

**TABLE 6.8** Quantities (Volumes) in Comparable Units for Eight Countries/Economy and Five Consumption Categories

Category	HK	BGD	IND	INDO	BRA	JPN	CAN	US
1	15,523	2,312	30,189	9,781	46,146	280,001	81,021	967,374
2	9,164	47,509	356,756	138,273	163,868	251,846	63,689	778,665
3	317,564	10,588	180,964	29,879	65,274	200,614	58,261	992,761
4	1,095	3,033	38,377	22,084	23,963	59,439	35,714	524,288
5	81,148	47,611	786,182	223,588	541,236	1,695,136	417,210	5,559,458

Note: See table 6.7 for consumption categories and country/economy abbreviations.

market exchange rates. For durables, India has the lowest price level at 0.77, and Canada has the highest at 1.16. For food, India has the lowest prices at 0.50, and Japan has the highest at 1.78. For other nondurables, India has the lowest price level at 0.33, and Japan has the highest level at 1.36. For energy, Indonesia has the lowest price level at 0.79, and Japan has the highest at 2.12. Finally, for services India has the lowest price level at 0.21, and Japan has the highest at 1.02. Thus the amount of price level variation across countries ranges from 38 percent for durables to 500 percent for services.

These data are used to compute relative consumption volumes for the eight countries using various multilateral methods. Instead of normalizing the relative volumes into shares of “world” consumption, the consumption of each country relative to the consumption of the United States is calculated. This is simply an alternative normalization of the country relative volumes.

The star method of constructing relative volumes for the eight countries is explained briefly in section 6.2. Basically, one country is chosen as the “star” country, and the Fisher quantity index of all other countries is calculated relative to the star country. Thus the volumes of all eight countries are given relative to that of the star country. Table 6.10 lists these star relative volumes, but they have been normalized so that the volume of country 8 (the United States) has been set equal to unity, thereby giving some indication of the variability in the data.

The Fisher star parities for the seven countries relative to the United States have the following relative volume ranges: Hong Kong SAR, China: 0.01257–0.01355 (7.7 percent variation); Bangladesh, 0.01277–0.01437 (12.5 percent); India, 0.14351–0.16439 (14.5 percent);

**TABLE 6.9** Prices or PPPs of Consumption Components for Eight Countries/Economy and Five Consumption Categories

*U.S. dollars*

Category	HK	BGD	IND	INDO	BRA	JPN	CAN	US
1	0.92250	0.84905	0.76872	0.84184	1.14250	1.09838	1.16169	1.0
2	1.15255	0.52274	0.49553	0.60664	0.64398	1.78282	1.28839	1.0
3	0.85123	0.48168	0.33569	0.50731	0.93143	1.36020	1.19224	1.0
4	2.39178	1.02011	1.09769	0.79573	1.66644	2.11704	1.21359	1.0
5	0.76556	0.24421	0.21220	0.27393	0.50564	1.02468	0.90992	1.0

Note: See table 6.7 for consumption categories and country/economy abbreviations.

**TABLE 6.10** Fisher Star Volumes Relative to That of United States

Category	HK	BGD	IND	INDO	BRA	JPN	CAN	US
Star 1	0.01346	0.01367	0.16021	0.05158	0.09192	0.27530	0.07444	1.00000
Star 2	0.01257	0.01277	0.14351	0.04660	0.08984	0.24629	0.07129	1.00000
Star 3	0.01275	0.01350	0.15178	0.04984	0.09040	0.25596	0.07328	1.00000
Star 4	0.01277	0.01341	0.14902	0.04894	0.09141	0.25496	0.07262	1.00000
Star 5	0.01323	0.01284	0.15169	0.04837	0.09035	0.26357	0.07372	1.00000
Star 6	0.01355	0.01437	0.16439	0.05322	0.09504	0.27724	0.07464	1.00000
Star 7	0.01343	0.01331	0.15387	0.05007	0.09105	0.27596	0.07429	1.00000
Star 8	0.01346	0.01277	0.15178	0.04894	0.09035	0.27724	0.07429	1.00000

Note: See table 6.7 for country/economy abbreviations.

Indonesia, 0.04660–0.05322 (14.2 percent); Brazil, 0.08984–0.09504 (5.8 percent); Japan, 0.24629–0.27724 (12.6 percent); and Canada, 0.07129–0.07464 (4.7 percent). Thus the variation in relative volumes is quite large, depending on which country is used as the base country in a comparison based on the use of Fisher star parities.

The GEKS, GK, and IDB methods<sup>22</sup> for comparing relative volumes were explained in previous sections and in chapter 5. Consumption volumes (relative to the United States) for the eight countries were computed using these methods, and they are listed in table 6.14, which appears later in this chapter.

Hill's MST spatial linking method was used as well. The relative volumes were computed using the three dissimilarity measures in equations (6.26)–(6.28). The first measure of relative price dissimilarity is the weighted log quadratic measure of relative price dissimilarity as defined by (6.26). The 8 × 8 matrix of relative price dissimilarity measures appears in table 6.11.

Table 6.11 reveals that the eight countries plus one economy fall into two groups that have similar price structures: the rich countries—Hong Kong SAR, China; Japan; Canada; and the United States (countries 1, 6, 7, and 8)—form one group, and the less rich countries—Bangladesh, India, Indonesia, and Brazil (2, 3, 4, and 5)—form the other group. The linking between the two

**TABLE 6.11** Weighted Log Quadratic Relative Price Dissimilarities between Eight Countries/Economy

	HK	BGD	IND	INDO	BRA	JPN	CAN	US
HK	0.00000	0.10056	0.11017	0.09067	0.07011	0.01381	0.03660	0.06143
BGD	0.10056	0.00000	0.01188	0.01165	0.05632	0.10506	0.13237	0.23223
IND	0.11017	0.01188	0.00000	0.03133	0.08980	0.13429	0.18955	0.29841
INDO	0.09067	0.01165	0.03133	0.00000	0.07084	0.07726	0.09610	0.19600
BRA	0.07011	0.05632	0.08980	0.07084	0.00000	0.09146	0.08770	0.14328
JPN	0.01381	0.10506	0.13429	0.07726	0.09146	0.00000	0.01904	0.05322
CAN	0.03660	0.13237	0.18955	0.09610	0.08770	0.01904	0.00000	0.02020
US	0.06143	0.23223	0.29841	0.19600	0.14328	0.05322	0.02020	0.00000

Note: See table 6.7 for country/economy abbreviations.

**TABLE 6.12** Weighted Asymptotic Quadratic Relative Price Dissimilarities between Eight Countries/Economy

	HK	BGD	IND	INDO	BRA	JPN	CAN	US
HK	0.00000	0.24097	0.27260	0.20885	0.15436	0.02914	0.10005	0.18062
BGD	0.24097	0.00000	0.02444	0.02455	0.11694	0.28243	0.32119	0.58729
IND	0.27260	0.02444	0.00000	0.07050	0.19169	0.36590	0.50994	0.85586
INDO	0.20885	0.02455	0.07050	0.00000	0.15038	0.20144	0.21328	0.45346
BRA	0.15436	0.11694	0.19169	0.15038	0.00000	0.21447	0.20876	0.36462
JPN	0.02914	0.28243	0.36590	0.20144	0.21447	0.00000	0.04023	0.11820
CAN	0.10005	0.32119	0.50994	0.21328	0.20876	0.04023	0.00000	0.04106
US	0.18062	0.58729	0.85586	0.45346	0.36462	0.11820	0.04106	0.00000

Note: See table 6.7 for country/economy abbreviations.

groups took place via Hong Kong SAR, China and Brazil.<sup>23</sup> The details of the spatial linking process are as follows. Country 7, Canada, is linked to country 8, the United States (the WLQ dissimilarity measure,  $\Delta_{WLQ}$ , equals 0.0202), and also to country 6, Japan ( $\Delta = 0.019$ ). Then country 6, Japan, is linked to country 1, Hong Kong SAR, China ( $\Delta = 0.0138$ ), which completes the linking of the rich countries. Country 2 acts as a star country for the poorer countries: country 2, Bangladesh, is linked to country 4, Indonesia ( $\Delta = 0.0116$ ), to country 3, India ( $\Delta = 0.0118$ ), and to country 5, Brazil ( $\Delta = 0.056$ ). Finally, the two groups of countries are linked via countries 1, Hong Kong SAR, China, and 5, Brazil ( $\Delta = 0.070$ ). The resulting MST volumes relative to the United States are listed in table 6.14.

The second measure of relative price dissimilarity is the weighted asymptotic quadratic measure of relative price dissimilarity, as defined by (6.27). The  $8 \times 8$  matrix of relative price dissimilarity measures appears in table 6.12.

The WAQ dissimilarity measures listed in table 6.12 are roughly two to three times the size of the WLQ measures listed in table 6.11. The lowest measure of dissimilarity is between Bangladesh and India ( $\Delta_{WAQ} = 0.02444$ ) and between Bangladesh and Indonesia ( $\Delta = 0.02444$ ). Then there is a shift to the rich countries, where the next lowest measure of dissimilarity is between Hong Kong SAR, China and Japan ( $\Delta = 0.02914$ ). The next lowest measure is between Japan and Canada ( $\Delta = 0.04023$ ) and then between Canada and the United States ( $\Delta = 0.04106$ ). Thus the rich countries are linked: Hong Kong SAR, China to Japan, then Japan to Canada, and then Canada to the United States. The next lowest measure of dissimilarity is between India and Indonesia, but Bangladesh has already been linked to both India and Indonesia, and so one must move to the next lowest measure of dissimilarity, which is between Bangladesh and Brazil ( $\Delta = 0.11694$ ). Thus all of the poor countries are now linked: Bangladesh is a poor country star, directly linked to India, Indonesia, and Brazil. Now to link the rich and poor countries, and the lowest dissimilarity measure between these two groups is again between Hong Kong SAR, China and Brazil. Thus the MST generated by the weighted asymptotic quadratic measure of relative price dissimilarity is exactly the same as the tree generated by the weighted log quadratic measure. Thus the MST (WLQ) relative volume parities will be exactly the same as the MST (WAQ) parities (see table 6.14).

The third measure of relative price dissimilarity is the Paasche-Laspeyres spread measure of relative price dissimilarity defined by (6.28). The  $8 \times 8$  matrix of relative price dissimilarity mea-

**TABLE 6.13** Paasche-Laspeyres Spread Relative Price Dissimilarities between Eight Countries/Economy

	HK	BGD	IND	INDO	BRA	JPN	CAN	US
HK	0.00000	0.06486	0.06860	0.08716	0.12845	0.01907	0.08539	0.09904
BGD	0.06486	0.00000	0.02484	0.02354	0.09876	0.09420	0.06389	0.02905
IND	0.06860	0.02484	0.00000	0.04596	0.00007	0.10421	0.13058	0.08134
INDO	0.08716	0.02354	0.04596	0.00000	0.05308	0.06806	0.02150	0.03897
BRA	0.12845	0.09876	0.00007	0.05308	0.00000	0.09387	0.11206	0.11177
JPN	0.01907	0.09420	0.10421	0.06806	0.09387	0.00000	0.01615	0.03390
CAN	0.08539	0.06389	0.13058	0.02150	0.11206	0.01615	0.00000	0.00133
US	0.09904	0.02905	0.08134	0.03897	0.11177	0.03390	0.00133	0.00000

Note: See table 6.7 for country/economy abbreviations.

asures appears in table 6.13.

The lowest measure of dissimilarity is between India and Brazil ( $\Delta_{PLS} = 0.00007$ ) and then between Canada and the United States ( $\Delta = 0.00133$ ). The next lowest measure of dissimilarity is between Japan and Canada ( $\Delta = 0.01615$ ) and then between Japan and Hong Kong SAR, China ( $\Delta = 0.01907$ ). Thus the rich countries are linked: Hong Kong SAR, China to Japan, then Japan to Canada, and then Canada to the United States, which is exactly the same set of linkages generated by the WLQ and WAQ measures of dissimilarity. The next lowest measure of dissimilarity is between Indonesia and Canada ( $\Delta = 0.02150$ ), and so the rich and poor countries are now linked by Indonesia and Canada! Recall that in the previous two spanning trees, the rich and poor countries were linked by Hong Kong SAR, China and Brazil. The next lowest measure of dissimilarity is between Bangladesh and Indonesia ( $\Delta = 0.02354$ ) and then between Bangladesh and India ( $\Delta = 0.02484$ ). Thus now all of the poor countries are linked: Indonesia to Bangladesh, Bangladesh to India, and India to Brazil. As mentioned earlier, rich and poor countries are linked via Indonesia and Canada. Thus the MST generated by the Paasche-Laspeyres spread measure of relative price dissimilarity is quite different from the trees generated by the WLQ and WAQ measures. The MST (PLS) relative volume parities are reported in table 6.14. It lists the country consumption volumes relative to those of the United States

**TABLE 6.14** Country/Economy Consumption Volumes Relative to the Those of the United States Using Six Multilateral Methods

Method	HK	BGD	INDIA	INDO	BRA	JPN	CAN	US
GEKS	0.01315	0.01332	0.15317	0.04966	0.09128	0.26556	0.07357	1.0
MST (WLQ)	0.01349	0.01310	0.14720	0.04779	0.09214	0.27596	0.07429	1.0
MST (WAQ)	0.01349	0.01310	0.14720	0.04779	0.09214	0.27596	0.07429	1.0
MST (PLS)	0.01349	0.01372	0.15420	0.05007	0.09184	0.27596	0.07429	1.0
GK	0.01386	0.01357	0.16258	0.05057	0.09613	0.27814	0.07431	1.0
IDB	0.01346	0.01392	0.16187	0.05143	0.09441	0.27076	0.07417	1.0

Note: GEKS = Gini-Éltető-Köves-Szulc; MST = minimum spanning tree; WLQ = weighted log quadratic; WAQ = weighted asymptotically quadratic; PLS = Paasche-Laspeyres spread; GK = Geary-Khamis; IDB = Iklé-Dikhanov-Balk. See table 6.7 for country/economy abbreviations.

for the six multilateral methods.

In table 6.14, the volume parities for the seven countries relative to the United States have the following ranges: Hong Kong SAR, China, 0.01315–0.01386 (5.4 percent variation); Bangladesh, 0.01310–0.01392 (6.3 percent); India, 0.14720–0.16258 (10.4 percent); Indonesia, 0.04779–0.05143 (7.6 percent); Brazil, 0.09128–0.09613 (5.3 percent); Japan, 0.26556–0.27814 (4.7 percent); and Canada 0.07357 to 0.07429 (1.1 percent). Thus the variation in relative volumes is quite large, depending on which multilateral method is used.

The relative consumption volumes generated by the four methods based on the use of a bilateral superlative index (the GEKS and the MST or similarity linking methods) are fairly close to each other, and the relative consumption volumes generated by the two additive methods (GK and IDB) are also fairly close to each other. However, the additive methods tend to overstate the consumption levels of the poorer countries (Bangladesh, India, Indonesia, and Brazil) relative to that of the United States.<sup>24</sup>

Based on table 6.14, it is difficult to choose between GK and IDB if an additive method is required: both methods tend to overstate the volumes of poor countries relative to those of rich countries, but the degree of overstatement seems to vary between poor countries.

Turning to methods based on the economic approach to multilateral comparisons, the MST method based on the Paasche-Laspeyres spread is not recommended because this measure of dissimilarity does not adequately distinguish dissimilar price vectors. In the empirical example just given, the WLQ and WAQ measures of dissimilarity gave rise to the same set of comparisons, and so for this example these two variants of the MST method cannot be distinguished from one another. The differences between the GEKS volume estimates and the MST (WLQ) estimates are smaller than the differences between the GEKS estimates and the two additive methods, but there are some significant differences.<sup>25</sup>

What are the advantages and disadvantages of using the GEKS method versus the MST (WLQ) or MST (WAQ) method? The GEKS method has the advantage of using all possible bilateral comparisons between each pair of countries in the comparison, and thus it is more robust to data problems in any one country. On the other hand, the MST method is very dependent on each set of bilateral comparisons in the final tree of comparisons, and so poor-quality data for any single country could adversely affect the overall quality of the comparison. But if the quality of the data is roughly the same across countries, the MST method is the spatial counterpart to the use of the chain principle in annual intertemporal comparisons—that is, using the MST method, the countries that have the most similar structures of relative prices are compared, and bilateral comparisons are generally regarded as being more accurate if the structure of relative prices is similar. Thus in the empirical example just presented, the United States and Canada (which have very similar structures of relative prices) are linked directly via the Fisher index between these two countries using the MST method, whereas under GEKS, links involving all other countries enter the comparison. Thus if data quality were uniformly high across countries, the MST method would seem to be preferred over the GEKS method.<sup>26</sup>

Now consider the problems associated with forming regional shares of “world” consumption, where the “world” is simply the eight countries listed earlier. Suppose the first four countries form an “Asian” region, region 1, and the remaining four countries form a “rest of the world” region, region 2. Obviously, the Heston-Dikhanov method for forming regional shares can be applied to the data in table 6.14. The consumption volumes (relative to that of the United States) listed there for the GEKS method, the three MST methods, the GK method, and the IDB method can be converted into shares of “world” consumption, and then the first four country shares can be summed to form the region 1 shares,  $S_1$  and  $S_2$ . The resulting region 1 shares for the six methods



**TABLE 6.15** Regional Consumption Quantities or Volumes  $Q_m$  by Consumption Category

	Consumption category				
	1	2	3	4	5
Region 1	57,805	551,702	238,995	64,589	1,138,529
Region 2	1,374,542	1,258,068	1,316,910	643,404	8,213,040

Note: See table 6.7 for consumption categories.

are listed in table 6.18.

It is also useful to calculate the region 1 shares of “world” consumption for the option 1 (Diewert) and option 3 (Sergeev) methods. Equations (6.7) and (6.8) are first used to calculate the country and regional quantity vectors; the country expenditures  $E_{rcn}$  are listed in table 6.7, and the country basic heading PPPs (which can be set equal to  $\gamma_{rcn}$ ) are listed in table 6.9. The resulting matrix of regional consumption volumes or “quantities” defined by (6.8),  $Q_m$ , are listed in table 6.15.

The regional unit value commodity prices  $n$  for each region  $r$ ,  $P_m$ , defined by (6.12) are listed in table 6.16.

As could be expected, the region 2 unit value commodity prices are all relatively close to the U.S. prices (which are all equal to unity) because three out of four of the region 2 countries are “rich” and hence have price structures similar to the U.S. structure. The region 1 unit value prices are all lower than the corresponding region 2 prices, and for commodity group 5, services, the region 1 unit value price is considerably lower. Now the Fisher quantity index for region 2 relative to region 1 can be calculated using the regional price and quantity data listed in tables 6.15 and 6.16. The resulting index is equal to 6.26177. Thus the relative regional consumption volumes are 1 and 6.17739, and when these volumes are converted into shares, the region 1 share of world consumption is 0.13771 and the region 2 share is 0.86229.

Finally, the option 3 additive method suggested by Sergeev (2009) is evaluated. To implement this method, it is necessary to compute the geometric mean of the regional basic heading prices listed in table 6.8. Recall equations (6.18). These regional geometric mean prices  $P_m^*$  are listed in table 6.17.<sup>27</sup>

The pattern of regional commodity prices is fairly similar in tables 6.16 and 6.17. To complete the analysis for this case, the Fisher quantity index for region 2 relative to region 1 is calculated using the regional quantity and price data listed in tables 6.15 and 6.17, respectively, and the resulting index is equal to 6.17739, which is very similar to the corresponding option 1

**TABLE 6.16** Regional Unit Value Consumption Prices or PPPs  $P_m$  by Consumption Category

	Consumption category				
	1	2	3	4	5
Region 1	0.82560	0.53663	0.40150	1.01274	0.26510
Region 2	1.03435	1.12493	1.05998	1.13987	0.96794

Note: See table 6.7 for consumption categories.

**TABLE 6.17** Regional Geometric Mean Prices or PPPs  $P_{rn}^*$  by Consumption Category

	Consumption category				
	1	2	3	4	5
Region 1	0.84377	0.65236	0.51405	1.20824	0.32287
Region 2	1.09882	1.10282	1.10861	1.43846	0.82863

Note: See table 6.7 for consumption categories.

Fisher index, which was equal to 6.26177. Thus the relative regional consumption volumes for the Sergeev method are 1 and 6.17739, and when these volumes are converted into shares, then the region 1 share of world consumption is 0.13933 and the region 2 share is 0.86067.

The resulting region 1 shares of “world” consumption generated by the various methods discussed earlier are summarized in table 6.18.

Although the variations in the region 1 shares that the various methods generate are not huge, they certainly are not negligible. The percentage differences between the various estimated shares and the preferred MST (WLQ) and MST (WAQ) shares are listed parenthetically in table 6.18. The preferred MST measure of the region 1 “world” share of real consumption is 3.7–7.6 percent below the region 1 shares generated by the other methods for linking the regions. More experimentation using data from the 2005 round of the ICP should be carried out before a definitive decision can be made on which method should be used to link the regions in the 2011 ICP.

## 6.7 Conclusion

Some tentative conclusions can be drawn from the analysis and examples presented in this chapter:

- The option 2 method should be ruled out for the 2011 ICP because, as Sergeev has pointed out, it is not invariant to the choice of the numeraire countries in the regions.
- The option 1 method should also be ruled out for the 2011 ICP because it depends on market exchange rates, which are not reliable and hence their use should be avoided

**TABLE 6.18** Share of World Output for Region 1 Using Various Methods

	Region 1 share of consumption
Option 1 (regional unit values method)	0.13771 (3.4 %)
Option 3: geometric mean average prices in each region	0.13933 (4.6 %)
GEKS	0.13815 (3.7 %)
GK	0.14243 (7.0 %)
IDB	0.14326 (7.6 %)
MST (PLS)	0.13831 (3.9%)
MST (WLQ) and MST (WAQ)	0.13316

Note: See table 6.14 for methodology abbreviations.

if possible.

- The option 3 method is not fully consistent with the economic approach to index number theory because the method is inherently an additive one (when constructing regional volume aggregates), and additive methods are subject to substitution bias.
- The use of the GK or IDB method should perhaps be ruled out for “headline” estimates for country and interregional parities for major aggregates because of their inherent substitution biases. These methods could be used, however, to provide users with analytical tables when they demand an additive method.<sup>28</sup>
- GEKS remains a viable method for constructing regional shares in a consistent manner.
- Various forms of spatial linking should also be considered for the 2011 ICP, but the use of this method should await more experimental results using the 2005 ICP database.

#### NOTES

1. The five geographic ICP regions in 2005 were Africa, Asia-Pacific, Commonwealth of Independent States (CIS), South America, and Western Asia. The Eurostat–Organisation for Economic Co-operation and Development (OECD) members constituted a sixth region.
2. The author is indebted to Yuri Dikhanov, D. S. Prasada Rao, Sergey Sergeev, and Frederic A. Vogel for their helpful comments.
3. These multilateral methods are described in chapter 5.
4. The Country Product Dummy (CPD) method was used by the Africa, Asia-Pacific, and Western Asia regions; the Country Product Representative Dummy (CPRD) method was used by the South America region; and the Gini-Éltető-Köves-Szulc (GEKS\*) method was used by the Eurostat-OECD and CIS regions. See chapter 4 for a review of these methods. For details on how the commodity lists were chosen and how the basic heading PPPs were linked across regions, see chapters 4 and 5. For a summary of the new methodologies used in the 2005 ICP, see Diewert (2010).
5. The 18 Ring countries were the Arab Republic of Egypt; Brazil; Cameroon; Chile; Estonia; Hong Kong SAR, China; Japan; Jordan; Kenya; Malaysia; Oman; the Philippines; Senegal; Slovenia; South Africa; Sri Lanka; United Kingdom; and Zambia.
6. This method flows from the work of Gini (1924, 1931), Éltető and Köves (1964), and Szulc (1964).
7. Sergeev (2009b) first noticed this problem with this method, which was used to link the regions in the 2005 ICP.
8. If the country experiences high inflation within the year, finding an appropriate annual price is more complicated—see Hill (1996).
9. See the discussion in chapter 5 on additive methods and substitution bias.
10. See Hill (1999a, 1999b, 2001, 2004). The work by Fisher (1922, 271–274) was a precursor to the work by Hill.
11. The method was recently suggested by Robert J. Hill (2010), but it is possible that Eurostat and OECD have been using variants of this method for some time.
12. This method of aggregation within a region is related to the geometric average price multilateral method originally suggested by Walsh (1901, 381,398). It was noted by Gini (1924, 106) and implemented by Gerardi (1982, 387). These authors used reference world prices that were the geometric mean over all countries in the world. They were applied to all countries, so that the resulting volume estimates were additive over all countries and all regions. In section 6.5

- of this chapter, the Sergeev method parities are computed using a subset of the 1985 ICP data. The regional parities and price parities  $P_m$  are equal to a regional constant times the parities defined by (6.18). This will retain the fixity of the final regional parities.
13. For an explanation of the problem, see Marris (1984, 52), Diewert (1999, 48–50), or chapter 5.
  14. See Gini (1924, 1931), Éltető and Köves (1964), Szulc (1964), and Diewert (1999, 31–37) for the properties of this method.
  15. Instead of using the Fisher ideal quantity index as the basic building block for this method, any other superlative quantity index could be used in this multilateral method. However, the Fisher index is generally preferred because of its strong axiomatic and economic properties; see Diewert (1976, 1992).
  16. Kravis (1984, 10) introduced this “star” terminology.
  17. Heston (1986, 3) suggested essentially the same methodology except that he suggested the world comparison be made using the GK multilateral method.
  18. On the other hand, the additive methods discussed in the previous section are consistent with homothetic preferences that can provide only a first-order approximation to arbitrary homothetic preferences. For further explanation of this point, see Diewert (1999, 31), who introduced the concept of a superlative multilateral system. The GEKS system is a superlative method, whereas the additive methods are not. Balk (2009, 82) recently provided an overview of various multilateral methods and endorsed the GEKS-Fisher method as a center stage method, particularly from the economic approach to international comparisons.
  19. For a discussion of how to deal with the problems with (6.26) and (6.27) that arise if any of the PPPs  $p_n^r$  are zero, see chapter 5.
  20. See table 6.6 for a summary of the differences due to the choice of different numeraires.
  21. The Sergeev regional price and quantity vectors  $P^r$  and  $Q^r$  for  $r = 1, 2$  turn out to be  $P^1 = [4.472, 2.000]$ ,  $P^2 = [6.325, 8.000]$ ,  $Q^1 = [10.1, 12.0]$ , and  $Q^2 = [11.0, 15.0]$ .
  22. See Geary (1958), Khamis (1972), and Iklé (1972).
  23. Another possible bilateral link between the two regions would be via Indonesia and Japan. They have a dissimilarity measure equal to 0.07726, which is a bit higher than the Hong Kong SAR, China and Brazil dissimilarity measure, which is equal to 0.07011.
  24. The GK volumes relative to the GEKS volumes (with U.S. volumes normalized to equal 1) were all higher for the seven non-U.S. countries by the following percentages: 5.4 percent, Hong Kong SAR, China; 1.9 percent, Bangladesh; 6.1 percent, India; 1.8 percent, Indonesia; 5.3 percent, Brazil; 4.7 percent, Japan; and 1.0 percent, Canada. The IDB volumes relative to the GEKS volumes (with U.S. volumes normalized to equal 1) were also all higher for the seven non-U.S. countries by the following percentages: 2.3 percent, Hong Kong SAR, China; 4.5 percent, Bangladesh; 5.7 percent, India; 3.6 percent, Indonesia; 3.4 percent, Brazil; 2.0 percent, Japan; and 0.8 percent, Canada.
  25. The MST (WLQ) volumes relative to the GEKS volumes (with U.S. volumes normalized to equal 1) differed by the following percentages: 2.6 percent, Hong Kong SAR, China; -1.7 percent, Bangladesh; -3.9 percent, India; -3.8 percent, Indonesia; 0.9 percent, Brazil; 3.9 percent, Japan; and 1.0 percent, Canada.
  26. However, data quality is not uniformly high across countries, and so this argument for the use of the MST method is not decisive.
  27. For comparison purposes, the vector of international prices generated by the GK method (with the price of commodity group 1 normalized to equal 1) is [1.00000, 1.16099, 0.99626, 1.22554, 0.86834] and the vector of international prices generated by the IDB method is [1.00000, 1.10438, 0.92265, 1.53623, 0.66525]. The U.S. prices are all equal to 1. Thus

when the price of durables is set equal to unity for all three price vectors, the GK and IDB price levels for food and for energy are above the corresponding U.S. price levels, while the GK and IDB price levels for other nondurables and services are below the U.S. levels. As expected, the GK prices are closer to the structure of U.S. prices, whereas the IDB prices are a more “democratic” average of U.S. and poorer country prices.

28. Based on some empirical evidence developed by Deaton and Heston (2010) using the entire set of 2005 ICP data, if it is desired to have the additive method parities approximate the GEKS parities, then IDB appears to be better than GK (the axiomatic properties of IDB may be more attractive to users than the axiomatic properties of GK). See chapter 5 for a discussion of the axiomatic properties of the GK and IDB methods.

## REFERENCES

- Aten, B., and A. Heston. 2009. “Chaining Methods for International Real Product and Purchasing Power Comparisons: Issues and Alternatives.” In *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, ed. D. S. Prasada Rao, 245–73. Cheltenham, U.K.: Edward Elgar.
- Balk, B. M. 2009. “Aggregation Methods in International Comparisons: An Evaluation.” In *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, ed. D. S. Prasada Rao, 59–85. Cheltenham, U.K.: Edward Elgar.
- Deaton, A., and A. Heston. 2010. “Understanding PPPs and PPP-Based National Accounts.” *American Economic Journal: Macroeconomics* 2 (4): 1–35.
- Diewert, W. E. 1976. “Exact and Superlative Index Numbers.” *Journal of Econometrics* 4: 114–45.
- . 1999. “Axiomatic and Economic Approaches to International Comparisons.” In *International and Interarea Comparisons of Income, Output and Prices*, Studies in Income and Wealth, Vol. 61, ed. A. Heston and R. E. Lipsey, 13–87. Chicago: University of Chicago Press.
- . 2004. “On the Stochastic Approach to Linking the Regions in the ICP.” Discussion Paper 04-16, Department of Economics, University of British Columbia, Vancouver.
- . 2009. “Similarity Indexes and Criteria for Spatial Linking.” In *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, ed. D. S. Prasada Rao, 183–216. Cheltenham, U.K.: Edward Elgar.
- . 2010. “New Methodological Developments for the International Comparison Program.” *Review of Income and Wealth*, series 56, special issue 1 (June): S11–S31.
- Dikhanov, Y. 1994. “Sensitivity of PPP-Based Income Estimates to Choice of Aggregation Procedures.” Paper presented at 23rd General Conference of the International Association for Research in Income and Wealth, St. Andrews, Canada. World Bank, Washington, DC, June 10.
- . 1997. “Sensitivity of PPP-Based Income Estimates to Choice of Aggregation Procedures.” Development Data Group, International Economics Department, World Bank, Washington, DC, January.
- . 2007. “Two Stage Global Linking with Fixity: Method 1 (EKS).” Unpublished paper, World Bank, Washington, DC.
- Éltető, O., and P. Köves. 1964. “On a Problem of Index Number Computation Relating to International Comparison.” *Statisztikai Szemle* 42: 507–18. <http://siteresources.worldbank.org/ICPINT/Resources/icppapertotal.pdf>.
- Fisher, I. 1922. *The Making of Index Numbers*. Boston: Houghton Mifflin.
- Geary, R. G. 1958. “A Note on Comparisons of Exchange Rates and Purchasing Power between Countries.” *Journal of the Royal Statistical Society Series A* 121: 97–99.

- Gerardi, D. 1982. "Selected Problems of Intercountry Comparisons on the Basis of the Experience of the EEC." *Review of Income and Wealth* 28: 381–405.
- Gini, C. 1924. "Quelques considérations au sujet de la construction des nombres indices des prix et des questions analogues." *Metron* 4 (1): 3–162.
- . 1931. "On the Circular Test of Index Numbers." *Metron* 9 (9): 3–24.
- Heston, A. 1986. *World Comparisons of Purchasing Power and Real Product for 1980*. New York: United Nations and Eurostat.
- Hill, P. 1996. *Inflation Accounting: A Manual on National Accounting under Conditions of High Inflation*. Paris: Organisation for Economic Co-operation and Development.
- Hill, R. J. 1999a. "Comparing Price Levels across Countries Using Minimum Spanning Trees." *Review of Economics and Statistics* 81: 135–42.
- . 1999b. "International Comparisons Using Spanning Trees." In *International and Interarea Comparisons of Income, Output and Prices*, Studies in Income and Wealth, Vol. 61, ed. A. Heston and R. E. Lipsey, 109–20. Chicago: University of Chicago Press.
- . 2001. "Measuring Inflation and Growth Using Spanning Trees." *International Economic Review* 42: 167–85.
- . 2004. "Constructing Price Indexes across Space and Time: The Case of the European Union." *American Economic Review* 94: 1379–410.
- . 2009. "Comparing Per Capita Income Levels across Countries Using Spanning Trees: Robustness, Prior Restrictions, Hybrids and Hierarchies." In *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, ed. D. S. Prasada Rao, 217–44. Cheltenham, U.K.: Edward Elgar.
- . 2010. "Linking the Regions in the International Comparison Program at the Basic Heading Level and at Higher Levels of Aggregation." Unpublished paper, Department of Economics, University of Graz, Austria, November 23.
- Iklé, D. M. 1972. "A New Approach to the Index Number Problem." *Quarterly Journal of Economics* 86 (2): 188–211.
- Khamis, S. H. 1972. "A New System of Index Numbers for National and International Purposes." *Journal of the Royal Statistical Society Series A* 135: 96–121.
- Kravis, I. B. 1984. "Comparative Studies of National Income and Prices." *Journal of Economic Literature* 22: 1–39.
- Kravis, I. B., A. Heston, and R. Summers. 1982. *World Product and Income: International Comparisons of Real Gross Product*. Statistical Office of the United Nations and World Bank. Baltimore: Johns Hopkins University Press.
- Marris, R. 1984. "Comparing the Incomes of Nations: A Critique of the International Comparison Project." *Journal of Economic Literature* 22: 1, 40–57.
- Sergeev, S. 2001. "Measures of the Similarity of the Country's Price Structures and Their Practical Application." Conference on the European Comparison Program, United Nations Statistical Commission. Economic Commission for Europe, Geneva, November 12–14.
- . 2009a. "Aggregation Methods Based on Structural International Prices." In *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, ed. D. S. Prasada Rao, 274–97. Cheltenham, U.K.: Edward Elgar.
- . 2009b. "The Evaluation of the Approaches Used for the Linking of the Regions in the ICP 2005." Unpublished paper, Statistics Austria, December.
- Szulc, B. 1964. "Indices for Multiregional Comparisons." *Przegląd Statystyczny* 3: 239–54.
- Walsh, C. M. 1901. *The Measurement of General Exchange Value*. New York: Macmillan.