S. Sergeev
(Statistics Austria)

Possible approaches for the linking of the regions at the BH and aggregated levels for the ICP 2011

The ICP is carried out from the 1980 Round in accordance with the Regional principle. This is due to the situation that some "ICP "Regions" are in the fact the international organizations (EU, OECD, CIS, etc.) which need the regional PPP results (usually without the impact of non-members) for the internal purposes and due to the fact that the Regions or sub-Regions comprised usually more or less homogeneous sets of the countries. The last circumstances allow to produce shorter and more comparable item lists and, in effect, more reliable regional results.

In accordance with the regional organization of the ICP, the regional results are calculated firstly and these regional results are linked in such way that the within-regional results are unchanged. There are several approaches for the linking of the regions. The ICP 2005 used the Ring comparison and between-regional PPPs[1] at the BH as well as at the aggregated levels, to link the Regions in the Worldwide comparison and to keep simultaneously the within-regional results as unchanged.

Taking into account the experience of the ICP 2005 and the respective TAG recommendations, the Global Core List approach should be used for the ICP 2011. A common set of products is included in all regional products lists. All countries can attempt to price these overlapped products. PPPs calculated on the basis of these price data are linking factors between regions. Regional representativity of price data and robustness of the results should be substantially increased relatively the Ring comparison. What impact has this decision on the computational procedures[2] for the linking? Some considerations on different possible methods are given below.

General computational approaches

Two possible computational approaches to obtain the Global results with fixity (the combination of regional data sets with linking factors) can be used *(both for BH-PPPs as well as for aggregated PPPs):*

A. Regional approach (RA) [ICP 2005 approach]

Calculation of between-regional PPPs

Country's data for the linking are combined in the Regional input data sets. The calculations are done for the Regions as some Super-Countries.

---

[1] The concept of the between-regional PPPs were proposed by E.Diewert – see Diewert, 2004.
[2] Main non-computational problematic points for the linking are the following.
   a) Different Regions use different methods for:
       - Non-market services: Productivity Adjustment for CoE (some Regions – Yes; some other – No)
        - Education (some Regions–Output approach; some other–Input costs approach)
       - Rents (some Regions – Price appr.; some other – Quantity appr. or reference Volume index)
       - Different level of details for the regional GDP classification
   b) Different computational methods in the Regions.
If Fixity of Regional results is requested then full inter-Regional comparability is impossible in the World comparison due to factors mentioned above but it is necessary to attempt to reduce maximally their impact.

Main features:

- Straightforward approach with explicit use of fixity of regional PPPs

- Some rules should be established to form the regional input data sets

- Additional efforts are necessary to allow that each country within a region can contribute equally to the between-regional linking factors[3]

- Probably, less flexible if the regional methodologies are substantial different (e.g. Productivity Adjustment in Non-market services – Yes / No)

B. Country Approach with Redistribution / Re-indexation (CAR)

All countries participate individually in the Global calculation. Regional Totals (Volumes or PPPs) from the Global comparison are redistributed / re-indexed in accordance with the Regional ratios (Volume shares or PPPs)[4]

Main features:

- Each country treated equally in the Global comparison

- Fixity of regional results is obtained by an indirect (two-stage) approach

- Probably, the differences in the regional methodologies will have some lower impact on the (in)comparability of the World results with fixity [see Heston (2010)]

I. Possible approaches for the linking at the BH level

Input data

- Regional BH-PPPs (to keep the regional fixity)
- Country's prices (with the indication of importance of products) for items from the Global Core List

A. Regional approach (RA) [ICP 2005 approach]

Calculation of between-regional PPPs

Country's data for Linking are combined in the Regional input data sets. The calculations are done for the Regions as some Super-Countries.

The main features:

- Straightforward approach with explicit use of fixity of regional PPPs

---

[3] If the concept of the between-regional PPP is used then it is logically that each region should be treated symmetrically. However the combination of Regional input data in this case should also follow this principle - each country within a region should be treated symmetrically during this process.

[4] The shares of a country in Real expenditure of a region or the World is nothing more as Volume indices "Country / Region" and "Country / World". The shares of a Region in Real expenditure of the World is nothing more as Volume index "Region / World". Therefore the product of the share of a Region in the World and the share of a country within this Region is Volume index "Country / World":

VI "Region / World" x VI "Country / Region" = VI "Country / World".
The same concerns the PPPs.

- Some rules should be established to form the regional input data sets (Probably not all collected data for the Core list should be used. For example, if only few countries in a region priced a product[5] then – Should this product be included in the between-regional calculations?)

- Less flexible if the regional methodologies are substantial different (e.g. Productivity Adjustment in Non-market services – Yes / No)

B. Country Approach with PPP-Re-indexation (CAR-PPP)

All countries participate individually in the Global calculation.
Regional PPPs) from the Global comparison are re-indexed in accordance with the Regional ratios (PPPs)

PPP „Country / World_T" =

PPP „Country / Region_T" * PPP „Region_T / World_T"

[Reg. Comparison]   [Global Comparison]

The main features:

- Fixity of regional results is obtained by an indirect (two-stage) approach

- The possibilities to change input country data are rather limited. Input country data for the Core List items are included in the Regional calculations and these are validated and fixed (much) earlier. If we assume that a country X from a Region has high inconsistencies for some ICP items (outliers) in the Global comparison then – Should we correct these data in the Global comparison? If "Yes" then - Should be corrected this data in the Regional comparison? If "Not" then – Should be these questionable data be excluded from the respective Regional comparison? [*The Regional approaches is more flexible in this aspect, at least, technically. It is not necessary to include all collected price data in the Regional sets.*]

- Probably, the differences in the regional methodologies will have some lower impact on the (in)comparability of the World results with fixity

TAG ICP 2011 decided that at the BH level the between-regional CPD-PPPs should be calculated:

(a) *Country's prices submitted for the Global list are divided by the country regional BH-PPPs, to convert these into common Regional numeraire.*

(b) *Between-regional BH PPPs "Regional numeraire / World numeraire" are calculated on the basis of price set from (a) by the weighted CP(R)D method[6]*

This is a straightforward approach with explicit fixity of regional PPPs:

PPP „Country / World Numeraire" =

PPP „Country / Regional Numeraire" * PPP „Regional Num. / World Num."
[Regional Comparison]   [Global Comparison]

---

[5] See the specific China case in the ADB 2011.
[6] This method can be named as the RIPD method (**R**egion – **I**mportance – **P**roduct - **D**ummies).

The between-regional approach was selected but one point was still open - **What should be input data for the calculations?**

Two possible versions:

(a) *To use **individual country's prices** (submitted for the Global list and divided by the country regional BH-PPPs)*
Advantages:
- Whole inter-country price variation is taken into account
- Easier define Regional importance / representativity of products
Disadvantages:
- Regions are treated non-symmetrically (Regions with higher no. of countries have higher impact on the CPD coefficients – world prices)[7]

(b) *To use **Regional average (GM) prices** (submitted for the Global list and divided by the country regional BH-PPPs) as **regional sets** for the calculation of **b**etween-regional BH PPPs "Regional numeraire / World numeraire"*
Advantages:
- Straightforward principle "One Region – One set of data"
- Symmetrical treatment of Regions
Disadvantages:
- A part of inter-country price variation is lost
- Regional importance / representativity of products should be defined specially (By the RC ?)

More detailed analysis of technical aspects of the original Diewert's approach *(the use of country's prices vs. the use of regional average prices)* is made below.

The original approach proposed by E.Diewert assumes that **country's prices within a regional set are regarded as "outlet" prices**[8]) This looks as a natural way; however this approach has some problematic points. The main problematic point is non-symmetrical treatment of the regions due to **different no. of the countries from different Regions participating in the Global comparison.** The Regions with higher no. of Ring countries have higher impact on the CP(R)D estimations and respectively, in the final effect, on the ICP results. To demonstrate this statement more clearly, it is better to present the CPRD method not as a regression procedure

---

[7] The present point does not concern quality of supplied data but the purely quantitative aspect. The symmetry means her that the impact of a region on the Global results should not depend on no. of countries belonging to the Region. Pacific islands or Africa can have much higher impact of the Global CPD by the 2005 BH approach than LA simply due to the fact that they can have higher no. of prices in the Global CPD.
If we assume that there are differences in the quality in input data then it is desirable to take these differences into account. However to do this we should elaborate the criteria to evaluate the quality of input data and the procedures – how this factor can be included in the calculations.
We know that not only the Regions are very different concerning the quality of input data but also the countries within the regions are very different in this aspect. Why we treat the countries symmetrically in the Regional comparisons? If we have no such criteria and procedures then we should assume that all countries and Regions are equal [*presumption of innocence* ;-)].
[8] It is possible to calculate by this approach not only the CPRD PPPs but also the EKS PPPs. In this case both types of binary Jevons GM-PPP (*- and -*) should include all possible inter-regional binary PPP for all countries in the regions. For example, if we calculate the PPP between Reg I (countries A, B, C, D) and Reg II (countries E, F, G) then all 12 possible pairwise binary PPPs are included in the Jevons-GM (A/E, A/F, A/G, B/E, B/F, B/G, ..., D/E, D/F, D/G).

but as a specific kind of the G-K method (geometric G-K with implicit weights) – see Sergeev (2005). International prices can be presented in the following form - if individual country's price data of Ring Countries is used then average international prices are calculated as the following:

$$(1) \quad \pi_i \; = \; \prod_{r=1}^{R} \{ \prod_{j=1}^{N(r)} \; [(P_{ijr}/\gamma^{Z_{rij}})/PP\mathrm{P_r}]^{q_{ijr}} \}^{1/_r\Sigma q_{ijr}} \; =$$

$$= \; \prod_{r=1}^{R} [(AvP_{ir}/\gamma^{Z_{ir}})/PP\mathrm{P_r}]^{N(r)_r\Sigma q_{ijr}/_r\Sigma q_{ijr}} \qquad i = 1, 2,..., M$$

$AvPr_{ir}$ – average (GM) price of item i for a set of Ring countries from the Region r;

$^{N(r)}_r\Sigma q_{ijr}$ – implicit weight (quantity) for Region **r** within item i (which depends in on no. of regional Ring countries priced this item).

The term (1) contains a specific kind of Gerschenkron effect: more no. of prices from a Region leads to higher impact of this Region on international average prices. Additionally the inter-regional PPPs depend also on no. of prices collected for an item within a Region. If item "**i1**" was priced in 10 Ring countries in a Region and item "**i2**" – only in five countries in the same Region then item "**i1**" will have a higher impact on the CPRD coefficients and respectively on the PPP for this region the item "**i2**". It is not always desirable because the no. of priced items has no a direct relation with expenditure weights and very often simple items (with low expenditure weights) are priced more frequent because it is easier to find and identify them.

From a general point of view, the Global inter-regional comparison should be carried out principally in a symmetrical way "One Region - One set of input data" => no Region should have more impact than other in the Global comparison simple due to higher no. of participating countries.

Is it possible to treat the regions symmetrically and simultaneously to work with data for individual countries?

Chapter 14 of the ICP Manual (page 20) indicates one possible way:
> "*If equal weighting of regions is considered desirable, it is possible to assign equal weight to each region while working with data for individual countries. Individual prices need to be weighted to take account not only of the differing numbers of prices reported by countries but more importantly of the differing numbers of countries between regions. The frequency-weighted CPD or CPRD method[9] in which the prices in a country are weighted inversely to the number of prices they report can be used to deal with the first factor. It ensures that each country receives equal weight within each region. However, the issue here is to ensure that each region receives equal weight. In this case, the prices in a country also need to be weighted inversely in proportion to the number of countries per region. Weights can easily be devised to take both factors into account*".

It seems that there can be significant problems by this approach. First of all, it is not desirable that the countries within the Regions with low no. of prices have equal

---

[9] The frequency-weighted CPD method was used in the first phase of the ICP. See Chapter 5 of "*A system of international comparisons of gross product and purchasing powe*r", by Kravis I. B., Kennessy Z., Heston, A. and Summers R., The Johns Hopkins University Press, 1975. See also Heston, Aten (2002).

impact[10], secondly (most important) - this method does not balance the impact of regions in the calculation of average international prices for products. One simplified example: let us have two mini-Regions: Reg I (4 countries) and Reg II (2 countries) and a BH with two products (PS1 and PS2) which were priced in countries as the following:

	Reg I	Reg II
PS1	4 prices	1 price
PS2	2 prices	2 prices

The countries from Reg I with 6 prices totally receive weight = 1/6; from Reg II with 3 prices totally receive weights = 1/3. The average international price for PS1 will be based on the prices of Reg I with weight (4 * 1/6) = 4/6 and on the prices of Reg II with weight (1 * 1/3) = 1/3; i.e. there is an obvious preference for Reg I in the calculation of average ring price for product PS1. *Vice versa*, the average international price for PS2 will be based on the prices of Reg I with weight (2 * 1/6) = 2/6 and on the prices of Reg II with weight = (2 * 1/3) = 2/3; i.e. there is an obvious preference for Reg II for product PS2. Formally, there is a "balancing" (a higher weight of Reg I for PS1 and vice versa a higher weight for Reg II for PS2. However, this mechanical balancing distorts (biases) international prices for both products PS1 and PS2 and this is rather a disadvantage.

The most simple, understandable and straightforward method, to avoid (or to decrease) the non-desirable impact of the factors mentioned above, is **the calculations of inter-Regional BH-PPPs on the basis of average regional prices calculated as GM** (use of GM guarantees invariant results) **of country's prices in numeraire regional currencies**.

The use of average regional GM prices leads to a symmetrical treatment of the Regions ("One region – one set of input data"). The World average prices calculated by this approach are "neutral" in the Regional aspect (they do not depend on no. of countries in the regions). It is visible from the term (1) which can be presented by this approach as follows:

$$(2) \quad \pi_i = \prod_{r=1}^{R} \{[(AvP_{ir} / \gamma^{Z_{ir}}) / PPP_r]^{q_{ir}}\}^{1/_r \Sigma q_{ir}}; \quad i = 1, 2,..., M$$

More concretely this approach can be described as follows:

- **average regional prices** are calculated as GM (use of GM guarantees invariant results) of respective country's prices in numeraire regional currencies;

- the following options can be used for the **attribution ("averaging") of the asterisks** (indication of representativity) to average regional prices:

● a formal (mechanical) approach: for example, at least half of the countries should have asterisk for a product,

● use of formal rules is simple but it seems this is not the best way. **The attribution of the asterisks * should be done in the consultation with the**

---

[10] See, A. Heston and B. Aten "Linking Country Groups in International Real Product and Purchasing Power Comparisons, page 33-34 (http://siteresources.worldbank.org/ICPINT/Resources/linking4.doc).

respective Regional co-coordinators[11]. This approach is based on a more objective basis because the Regional co-coordinators should have the knowledge - Is a product representative for their Regions as a whole or not?

Probably, the actual problem with the Regional representativity will be not so significant. The Table below show the situation for the ICP Core items for the 2010 Survey "Transportation & Restaurants, Hotels within the Eurostat comparison.

E10-2 / Overall38 VR1: ICP Items		No. of prices	No. of *
European PPP exercise			
Date: 14/03/2011	**Transportation and ReCaHo**		
Itemcode	Short description	No. of prices	No. of *
11.07.11.1.01.aa=ensw	ICP core list: Motor car, diesel, AUDI A4, SB	38	25
11.07.11.1.01.fa=ensw	ICP core list: Motor car, diesel, HYUNDAI Santa Fe, SB	33	11
11.07.11.2.01.aa=ensw	ICP core list: Motor car, petrol, <1,200cc, CHEVROLET / DAEWOO Matiz, SB	12	5
11.07.11.2.01.bb=ensw	ICP core list: Motor car, petrol, <1,200cc, CITROËN C3, SB	27	19
11.07.11.3.01.fb=ensw	ICP core list: Motor car, petrol, 1,200 - 1,699cc, FORD Focus, SB	35	28
11.07.11.3.01.ka=ensw	ICP core list: Motor car, petrol, 1,200 - 1,699cc, 75 HP, PEUGEOT 207, SB	34	30
11.07.11.4.01.da=ensw	ICP core list: Motor car, petrol, 1,700 - 2,999cc, FORD Mondeo, SB	31	15
11.07.11.4.01.eb=ensw	ICP core list: Motor car, petrol, 1,700 - 2,999cc, HONDA CR-V, SB	35	19
11.07.11.4.01.ha=ensw	ICP core list: Motor car, petrol, 1,700 - 2,999cc, MERCEDES-BENZ C Class, SB	34	19
11.07.11.4.01.ia=ensw	ICP core list: Motor car, petrol, 1,700 - 2,999cc, NISSAN Qashqai 4WD, SB	34	24
11.07.11.4.01.mb=ensw	ICP core list: Motor car, petrol, 1,700 - 2,999cc, VOLKSWAGEN Passat, SB	29	19
11.07.11.5.01.cb=ensw	ICP core list: Motor car, petrol, >3,000cc, MERCEDES-BENZ M Class, SB	36	12
11.07.11.5.01.da=ensw	ICP core list: Motor car, petrol, >3,000cc, VOLKSWAGEN Touareg, SB	18	8
11.07.12.1.01.ba=ensw	ICP core list: Motorcycle, HONDA CBR 125 R , SB	34	22
11.07.12.1.01.cb=ensw	ICP core list: Motorcycle, HONDA Hornet 600, SB	34	22
11.07.13.1.01.ca=ensw	ICP core list: City bicycle for men, WKB	31	16
11.07.13.1.01.da=ensw	ICP core list: Bicycle for children, 7 gears, WKB	29	15
11.07.21.1.01.ai=ensw	ICP core list: Summer tyre, 185/65 R14 (86) T, WKB	38	35
11.07.21.1.01.aj=ensw	ICP core list: Summer tyre, 185/65 R15 (88) H, WKB	38	36
11.07.21.1.02.ea=ensw	ICP core list: Car battery, BOSCH S4, SB	34	24
11.07.21.1.02.ga=ensw	ICP core list: Car battery, 55 Ah, WKB	36	27
11.07.22.1.01.aa=ensw	ICP core list: Diesel, normal, BNR	38	36
11.07.22.1.01.ba=ensw	ICP core list: Petrol, 95 octane, BNR	38	38
11.07.22.1.01.bb=ensw	ICP core list: Petrol, 98 octane, BNR	31	24
11.09.32.1.01.ca=ensw	ICP core list: Tennis balls, multipack, DUNLOP, SB	35	24
11.09.32.1.01.cb=ensw	ICP core list: Tennis balls, multipack, WILSON, SB	36	28
11.09.32.1.02.fa=ensw	ICP core list: Badminton racket, YONEX, SB	4	0
11.09.32.1.02.ga=ensw	ICP core list: Table tennis racket, WKB	36	30
11.11.11.1.01.ap=ensw	ICP core list: Grilled fresh sea fish (seasonal variety), main course, GMC-R, BNR	35	26
11.11.11.1.01.bd=ensw	ICP core list: Spaghetti Bolognese, main course, MODC-R, BNR	38	34
11.11.11.1.01.be=ensw	ICP core list: Pizza Margherita, MODC-R, BNR	37	32
11.11.11.1.01.bh=ensw	ICP core list: Breaded escalope, pork/chicken/turkey, main course, MODC-R, BNF	37	29
11.11.11.1.01.bq=ensw	ICP core list: Sweet and sour pork/chicken, main course, MODC-R, BNR	36	28
11.11.11.1.02.cd=ensw	ICP core list: Beef steak, main course, self-service restaurant , BNR	25	15
11.11.11.1.02.cp=ensw	ICP core list: BIG MAC, MC DONALD'S, BNR	34	32
11.11.11.1.02.cw=ensw	ICP core list: French fries, street stand, BNR	33	17
11.11.11.1.03.dd=ensw	ICP core list: Ice cream, industrially made, tea room/coffee house, BNR	36	30
11.11.11.2.01.ba=ensw	ICP core list: White house wine ("open" wine), glass, MODC-R, BNR	35	31
11.11.11.2.01.de=ensw	ICP core list: Draught beer, domestic, large, at the table, pub/bar, BNR	38	34
11.11.11.2.01.eb=ensw	ICP core list: BAILEYS' ORIGINAL IRISH CREAM, glass, at the table, pub/bar, BN	38	32
11.11.11.2.01.ec=ensw	ICP core list: Whisky, JOHNNY WALKER, Red Label, glass, at the counter, pub/t	37	33
11.11.11.2.02.gb=ensw	ICP core list: Orange juice, at the table, bar/café, BNR	38	38
11.11.11.2.02.gc=ensw	ICP core list: COCA COLA / PEPSI, at the table, bar/café, BNR	38	37
11.11.11.2.02.gf=ensw	ICP core list: Espresso, small cup, at the table, bar/café, BNR	38	37
11.11.11.2.02.ja=ensw	ICP core list: Mineral water, still, bottle, take-away, BNR	37	33
11.11.21.1.01.ba=ensw	ICP core list: Hotel - Category 2, Capital, weekend, BNR	32	26
11.11.21.1.01.bd=ensw	ICP core list: Hotel - Category 2, Capital, weekday, BNR	37	35
11.11.21.1.01.db=ensw	ICP core list: Hotel - Category 3, Capital, weekday, BNR	38	31

[11] There is an opinion that the procedure of the attribution of asterisks at the Regional level is much more subjective than at the country's national level. It seems that the objectivity of attribution of asterisks at the national level and the subjectivity of this at the Regional level are overestimated. Is the treatment of representativity even within the countries straightforward per se? It is doubtful that it is possible to attribute fully objectively the "average national" asterisks. There are general recommendations, some guidelines, etc. on these points but all these are not strict exactly defined rules. Therefore some conventions for the attribution of the asterisks are inevitable. For example, if an item is defined as a cluster of similar brands and a country X indicated that there are 2 representative brands and 1 non-representative or 3 representative and 4 non-representative then - What should be indicated for the representativity at the end for whole item in these cases? This problem exists also if the price collection is done by several country's regions (the especially for such large and heterogeneous countries like China, India, Russia, USA): for example, if an item was defined within a country as representative in 3 regions and as non-representative in 2 regions or as representative in 3 regions and as non-representative in 4 regions - What should be indicated for the representativity in these cases? This problem exists also for the type of shops / outlets because the ICP uses presently the concept of average price and not individual prices for outlets.

It is visible that the most of the ICP items should be considered as representative for the EU region – these were considered as representative practically by 80-90% of the ECP countries. One can assume that such situation will be also in other regions. So, the loss of data is rather an imaginary but not real danger. Further, numerous experiment showed that the impact of * is generally much lower than quality of prices (robustness of prices).

From a general point of view, the between-Regional linking should be carried out principally in a symmetrical way "one Region - one set of input data" - no Region should have more impact on overall results as other[12]. Obviously the averaging of prices of countries within the regions can't eliminate in a general case the differences in no. of average prices and in no. of representative prices in the Regions within the linking. This depends, first of all, on the quality of the established core item list and, secondly, on the price collection for this list in the countries. Nevertheless **the use of average (GM) Regional prices - one set of input data per Region – seems to be preferable from theoretical point of view (no. of Ring countries from a region should not play any role) as well as from practical point of view (more understandable and transparent)[13].**

In the meantime the concept "Representativity" was changed on the concepts "Importance". It seems that if TAG **moved from Representativity** (some assumptions about different PLI for representative and non-representative products reflecting in the coefficient R) **to Importance** (some assumptions about different expenditure weights for important and less important products) then **we should move to the weighted CPD instead of the CPRD[14]**. This is in accordance with the logic of the change of the concept and this should decrease the negative impact of incorrect allocation of * (Y.Dikhanov indicated that the CPRD method is very sensitive to the incorrect allocation of asterisks- *"Because representativity is one of the dimensions in the regression, along with country and product names [being on the right hand side of the regression expression], an error in representativity would be comparable to an error in country or product name. In other words, it would be like confusing tomatoes with potatoes, or Hong Kong with Butan."*). If the concept "Importance" is accepted then asterisks are just a form of weighting below the BH level and even if the weights are bad for some countries they may still be better than nothing. The experiments show that mistakes in weights have much less impact on the PPP than mistakes in prices. From the technical point of view, the Coefficient R is calculated within the regression and based fully on input data. If we use weighted CPD then the ratio of weights for expenditure weights for important and less important products is a parameter which we introduce ourselves – this is no regression parameter anymore and this should reduce Randomness. Hopefully experiments with the weighted CPD will show this.

TAG ICP decided that the between-regional PPPs (*concrete computational method is still open but the general regional approach was approved*) should be calculated at

---

[12] Exactly this principle „One region – one set of input data" is used in the approach proposed by E.Diewert for the calculation of aggregated between-regional PPPs.

[13] The discussions of the author with S.Ahmad during the ICP 2005 works and the discussions with A.Heston during the preparations of the ICP 2011 showed that these experts inclined earlier also on the modificaton of the original Diewert's approach (one Region – one set of data).

[14] E. Diewert was even earlier in favour of the weighted CPD (he proposed to used the ratio 3:1) – see Diewert (2004, page 43)

the BH level. However **the recent development in the ADB Region leads to the situation that the CAR method can be considered once again**.

China has decided to price all **601** Global Core items, while the rest of the 23 participating economies in Asia would collect prices for about **390** Global Core List products. Additional core items priced by China National Bureau of Statistics (NBS) would not be included in the validation process at the regional level as no other country in the region would price them, and but will be utilized in the inter-regional linking in a global setup.

If additional **211** items (priced by China only) are included in the calculation of between-regional PPPs then it means that the Asian between-regional PPPs (used for whole ADB Region) will depends heavily on China's price data. However it would be also very strange that additional China's price data is not utilized in the inter-regional linking.

It is not an accident that SUMMARY OF DECISIONS of ICP 2011 Executive Board Meeting, (New York, February 20, 2011) say:

"*The Board requests the TAG to submit by May 2011, through the Global Office, a technical note on the regional and/or global PPP impact, if any, of China's pricing all 601 items from the Global Core list, instead of the sole subset of it which all economies in the Asia and Pacific region jointly agreed to include in their price surveys*".

How this problem[15] can be solved?
If the Country Approach with Redistribution (CAR) is used at the BH level (Eurostat-OECD approach) then the impact of China additional price data is reduced for the ADB as a whole (this will concern mainly China itself) but China will have more links with the countries from other regions and China's PPP in the global exercise will be more robust. The Country Approach has generally more "degree of freedom" concerning the use of country's input data for the between-regional linking[16].

On other side, if we come back to the country approach by the CP(I)D method then international CPD prices will also gravitate to the prices of the Regions with higher

---

[15] Probably, additional prices supplied by China will really not affect drastically the ADB results in the Global comparison, at least, at the GDP level but this can be checked only after whole price data will be available. Nevertheless there is beside the technical aspect an other aspect – the organisational aspect. The ADB decided to include in the ADB list only 390 Core products from 601. Indirectly this means that they considered 211 products as non-representative for the Asian Region. Because all Regions should price not only representative but also non-representative products, to obtain equi-representativity then one can believe that the ADB considered these 211 products as extremely non-representative and excluded these from the list. The RC should supply the Regional input data for the linking to the GO and participate actively in ther validation. Therefore is very important to know in advance:
- What is the ADB position on China's additional price data (outside ADB comparion)?
- Do ADB and other Asian countries agree hat their results in the Global comparison depend on additional China's data?
[*The size of numerical differences does not play significant role by such rather political decisions*].
[16] As it was indicated by P.Konjin and F.Koechlin (message of 29.03.11), the actual problem for the between-regional approach is much more general than China's case. Generally one should decide: If a Region has N countries then – How many countries should price a core product, to include this product in the Regional data set for the linking? If the importance indicator is used then the problem is more complicated –How should be related no. of country's prices in a region with their importance indications, taking into account that each region should price not only own representative products but also representaive products from other regions, to allow the euqi-representativity / equi-importance.

no. of the countries (a kind of the Gerschenkron effect at the BH level). If we do not want this then we should use **the global EKS method for the countries (CAR-EKS)[17] also at the BH level**.

II. Possible approaches for the linking of the regions at the aggregated levels

To obtain Global aggregated results with fixity, the regional data sets should be presented in a standard classification:

- Linking BH factors and Regional BH-PPPs

- Aggregated Regional PPPs

- Aggregated Regional Volume indices (Volume shares)

- BH NA data in national currencies

A. Regional approach (RA) [ICP 2005 approach]

Calculation of between-regional PPPs
Country's data for Linking are combined in the Regional input data sets. The calculations are done for the Regions as some Super-Countries.

The main features:

- Straightforward approach with explicit use of fixity of regional PPPs

- Some rules should be established to form the regional input data sets

- Additional efforts are necessary to allow that each country within a region can contribute equally to the Regional input sets

- Less flexible if the regional methodologies are substantial different

(e.g. Productivity Adjustment in Non-market services – Yes / No)

B. Country Approach with Redistribution / Re-indexation (CAR)

All countries participate individually in the Global calculation.
Regional Volumes or PPPs from the Global comparison are redistributed / re-indxed in accordance with the Regional ratios (Volume shares or PPPs).

The main features:
- Each country treated equally in the Global comparison

- Fixity of regional results is obtained by an indirect (two-stage) approach

- Probably, the differences in the regional methodologies will have some lower impact on the (in)comparability of the World results with fixity

---

[17] This would be an analogue of the present Eurostat-OECD approach with the fully symmetry of all Regions. Of course, many bilateral PPP (like USA-Albania or Iceland – Tajikistan as well as the links via 3rd countries (like "USA – Canada via Zambia) are not very sensible but the same situation will be in the EKS Global aggregation if the Country approach will be adopted.

II.1 EKS method for the calculation of between-Regional aggregated PPPs used in the ICP 2005 and proposed modifications

The between-regional PPPs for the aggregation were proposed by E.Diewert (Diewert, 2004 and 2008). This approach was described in the Ch.15 of the ICP Manual and was used in the official ICP 2005. The Diewert's concept allows to keep the intra-regional PPPs as fixed figures and simultaneously to obtain the between-regional PPPs in a compact and elegant form.

Generally the Diewert's regional approach was straightforward. Nevertheless this original Diewert's approach used in the official ICP 2005 had two drawbacks:

1) First of all, original Diewert approach is not invariant relatively the choice of the Regional numeraires

2) Secondly, the countries within the Regions are treated in a non-symmetrical way in the Global comparison. This was not in the spirit of the EKS method officially adopted for the ICP 2005.

The author of this notice demonstrated the invariance of the original Diewert approach – see, for example, Sergeev (2009b). The problem was the choice of Regional numeriares for the calculations of the Regional Totals. **When one calculates the Regional GDP Total in a regional numeraire (a reference country) then the choice of the regional numeraires changes the structure of Regional Total GDP**. Speaking roughly, the use of a country as numeraire means that the regional GDP evaluated by price structure of this country.

To obtain the clear invariance in Diewert's approach, the author of this paper proposed **to operate with the average Regional numeraires** (like PPS in Eurostat comparison) instead of the regional country's numeraires (like HKD or UK Pound). This is the standardized presentation of PPPs with the base "Region = 1"[18]. This approach does not change the BH results but this is important for the further aggregation above the BH level and the problem with the invariance of the aggregated results will be avoided.[19]

The next problematic point of the Diewert's aggregation method is the kind of regional expenditure data (weights). The situation is ambiguous - the kind of BH-weights can depend partly on the method selected for the aggregation in the worldwide comparison:

a) If the **G-K** method is selected then the original Diewert's proposal – to use Regional expenditure in an (invariant) numeraire is unambiguously straightforward

---

[18] See for more details, S. Sergeev (2001c) „Presentation of results of international comparisons (Some considerations about the scaling procedures)"; Consultation on the European Comparison Programme (Geneva, 12-14 November) http://siteresources.worldbank.org/ICPINT/Resources/Selection_of_numeraire.doc

[19] Y. Dikhanov proposed to use varying base countries and taking a GM of the results. This approach is possible but is very cumbersome. For example, 146 countries from 5 Regions participated in the ICP 2005 (Eurostat-OECD-CIS - 55; Africa – 48; ADB – 23; LA – 10; WA -10). Total no. of variants - 6 072 000 ((6 072 000 = 55*48*23*10*10). It means that more than 6 mio. versions with different base-countries in the Regions had to be calculated. Of course, it is not problematic technically for modern computers. There is more serious problem - Would be all of these 6 mio. versions meaningful? If the GM of the results of all versions is taken at the end, then - Why not to use the Regional Average GM Numeraires from the beginning?

and is fully in spirit of the G-K method: aggregated regional expenditure data of in regional numeraire should be used as "super country" data of Regions.

b) If the **EKS** method is selected then there is a dilemma:

- to use **plutocratic weights** (as proposed initially by Diewert)

or

- to use **democratic weights** (regional average shares calculated from country's shares**)**?

The use of plutocratic weights (aggregated regional expenditure as a "super country") is not in spirit of the EKS method because the Regional expenditure would be dominated by large countries and small countries (as Luxembourg, Malta, Honkong, Singapore, etc.) would have no impact on the Regional data. The large countries will have higher impact on regional weights and this is not in accordance with the EKS principles. Additionally, it is unclear – why the countries are treated symmetrically in the Regional comparisons and non-symmetrically in the World comparison. Therefore a modification for the calculations of the "Between-regional PPPs using the EKS method", where the regions and the countries within regions are treated symmetrically is desirable.

The use of democratic weights allows to calculate the "between-regional aggregated PPPs" using the EKS method where the regions and the countries within regions are treated symmetrically. The "Between-regional aggregated EKS-PPPs" are calculated on the basis of matrix of between-regional aggregated Fishers PPPs using appropriate aggregated regional input data. To calculate the between-regional binary Fisher's PPPs, between-regional BH-PPPs and the regional BH expenditure (weights) are needed. The between-regional BH-PPPs are obtained within the Global comparison. The situation with the regional BH expenditure is less straightforward - use simply the sum of values for the countries (even in a common regional currency) would be not in accordance with the EKS principles. However it is easy to solve this problem taking into account that the Laspeyres and Paasche PPPs do not need absolute expenditure but the weights (shares) can be used instead. It is sufficient to calculate the average regional BH-shares as simple arithmetic mean from national shares (percentages) of the countries belonging to this Region (all countries - large and small - will contribute equally in this case). If region A has C(A) countries then average shares of the <u>Region A</u> are calculated as arithmetic mean from the values[20]:

$$w_{Aj} = \frac{e_{Aj}}{\sum_{i=1}^{n} e_{Aj}}$$

$$j = 1, 2, ..., C(A)$$

and if Region B has C(B) countries then average shares of the <u>Region B</u> are calculated as arithmetic mean from the values

$$w_{Bk} = \frac{e_{Bk}}{\sum_{i=1}^{n} e_{Bk}}$$

$$k = 1, 2, ..., C(B)$$

---

[20] If a region has only two countries then average regional shares will be the same as in Tornqvist method, if more than two countries – then the same as by the Walsh method.

The EKS procedure is be applied further to the set of between-regional Fishers PPPs (where each regional country is included in an equal way) to obtain the between-regional EKS-PPPs. So, all Regions and the countries within the Regions will be treated fully symmetrically by this approach in a natural transparent way in accordance with the underlying EKS principles[21] and democratic weights do not impose a set of common prices[22] to add up the quantities within a region (no bias).

An illustration of the proposed method is given below (Table 1) on the basis of the example from an initial version of Chapter 13 of the ICP manual.

The Chapter 15 of the ICP Manual describes one other possibility based on binary between-regional Fishers PPPs which are obtained with the intermediate use of within-regional EKS-PPPs as it was proposed by R.Hill[23]. A set of transitive aggregate PPPs by this approach is derived from a set of the binary PPPs which are calculated as GM of PPPs between all possible pairs of countries belonging to different regions (this is like country's prices are recalculated to regional numeraires by within-regional EKS-PPPs). However the use of data of individual countries by the R. Hill's proposal does not bring any advantage (neither conceptual nor technical) for the calculation of the between-regional PPPs. Conceptually the R.Hill method is less transparent and not straightforward – if country's input data are used then the standard country approach (so called, universal approach - all countries together) is much simple and understandable (in principle, between-regional PPPs are not necessary by this approach). From the technical side, if the calculation will be expanded on the whole set of the countries then this is a cumbrous procedure because this does not avoid the calculations of numerous bilateral between-country PPPs[24]. Obviously modern computers do these calculations very quickly but it is unclear - Why one needs the between-regional aggregated PPPs in the case when the between-country PPPs have been calculated? The EKS-PPPs can be calculated by an universal approach (all countries from all Regions - country approach) and a simple procedure can be used after this for the fixity of the between-country PPPs obtained within the regional comparisons.

---

[21] Y.Dikhanov indicated that if there are "bad countries" then there are problems with "plutocratic" as well as with "democratic" weightings: if "bad countries" are large countries then "plutocratic" weighting is worse, if small countries then – "democratic" weighting is worse. This is true but this is rather a general problem for all approaches – how to treat the countries with lower quality of input data. For example, the Ikle method uses the country's expenditure shares as the weights. Therefore the use of data of "bad countries" like SRI or TAJ in the Ikle aggregation will have the same higher negative effect as by "democratic" weighting.

[22] If we use democratic regional weights then Regional (GM) average numeraires seem to be not necessary - between-regional BH-PPPs are transitive and democratic regional weights do not depend on regional BH-PPPs.

[23] See R.J.Hill, T.P.Hill (2007).

[24] One of additional disputable points is the following – Does play some role the differences in the no. of the countries from different regions in the Ring comparison? This aspect can be considered from different points of view. R.Hill & P.Hill (see paper, 2008) believe that this factor does not play any role. They analyse an example for two regions A and B with 4 countries (M=4) and 2 countries (N=2). They wrote: *„Suppose now that the base country in region A is country a1, while the base country in region B is b1. The eight components of the untransitivized price index between regions A and B can now be simplified to the following:*

$$\mathbf{1} \times \quad \quad PF_{a1,b1} \times \quad \mathbf{1}$$
$$PEKS_{a1,a2} \times PF_{a2,b1} \times \quad \mathbf{1}$$
$$PEKS_{a1,a3} \times PF_{a3,b1} \times \quad \mathbf{1}$$
$$PEKS_{a1,a4} \times PF_{a4,b1} \times \quad \mathbf{1} \quad \quad \quad (4)$$
$$\mathbf{1} \times \quad \quad PF_{a1,b2} \times PEKS_{b2,b1}$$
$$PEKS_{a1,a2} \times PF_{a2,b2} \times PEKS_{b2,b1}$$
$$PEKS_{a1,a3} \times PF_{a3,b2} \times PEKS_{b2,b1}$$
$$PEKS_{a1,a4} \times PF_{a4,b2} \times PEKS_{b2,b1}$$

Sergeev's concern arises from the fact that six within-region-A EKS indexes will appear in (1) as compared with only four within-region-B EKS indexes. Does this imply that region A exerts greater influence than region B on the resulting between-region price index? **The answer is no**. *Region A seems to exert more influence in (4) only because more of the region B EKS indexes in (3) equal one and hence drop out of (4)."*

A possible contra-argument is that the indices which are automatically equal 1 do not bring any information in the comparison. No. of such indices by M=4 and N=2 is (M+N)=6. If one wants to minimize the share of them in the total no. of bilateral regional indices (=2*M*N) and, in effect, to maximize the use of input data, then this is obtained if M=N. It means that it is desirable that the regions A and B have equal no. of Ring countries.

Table 1

Calculation of between-Regional PPPs by the EKS method
(input data are from the example done in former Chapter 13)

Expenditures in national currencies

BH	Region A				Region B			Region C	
	Countries				Countries			Countries	
	1	2	3	4	1	2	3	1	2
1	2000	60	1200	9600	380	1000	3000	4500	225
2	500	18	400	2000	70	290	720	1200	67
3	800	35	700	4000	180	315	1260	1500	120
4	300	24	245	3300	70	170	500	800	36
1-4 = C	**3600**	**137**	**2545**	**18900**	**700**	**1775**	**5480**	**8000**	**448**
5	700	31	600	2800	180	430	1080	1500	75
6	700	32	700	2800	130	190	1280	2400	78
5+6 = I	**1400**	**63**	**1300**	**5600**	**310**	**620**	**2360**	**3900**	**153**
Total = GDP	**5000**	**200**	**3845**	**24500**	**1010**	**2395**	**7840**	**11900**	**601**

National shares of expenditures (%)

BH	Region A				Region B			Region C	
	Countries				Countries			Countries	
	1	2	3	4	1	2	3	1	2
1	40.0	30.0	31.2	39.2	37.6	41.8	38.3	37.8	37.4
2	10.0	9.0	10.4	8.2	6.9	12.1	9.2	10.1	11.1
3	16.0	17.5	18.2	16.3	17.8	13.2	16.1	12.6	20.0
4	6.0	12.0	6.4	13.5	6.9	7.1	6.4	6.7	6.0
1-4 = C	**72.0**	**68.5**	**66.2**	**77.1**	**69.3**	**74.1**	**69.9**	**67.2**	**74.5**
5	14.0	15.5	15.6	11.4	17.8	18.0	13.8	12.6	12.5
6	14.0	16.0	18.2	11.4	12.9	7.9	16.3	20.2	13.0
5+6 = I	**28.0**	**31.5**	**33.8**	**22.9**	**30.7**	**25.9**	**30.1**	**32.8**	**25.5**
Total = GDP	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Regional expenditure shares & between-regional BH-PPPs

BH	Region A		Region B		Region C	
	PPP	Av.Sh.	PPP	Av.Sh.	PPP	Av.Sh.
1	1	35.1	2.5	39.2	9	37.6
2	1	9.4	1.6	9.4	12	10.6
3	1	17.0	3.0	15.7	15	16.3
4	1	9.5	2.4	6.8	8	6.4
5	1	14.1	3.0	16.5	15	12.5
6	1	14.9	3.2	12.4	12	16.6
Total-GDP	xxx	100	xxx	100	xxx	100

The RESULTS by EKS-method (with weights in %) for GDP
MATRIX of BINARY LASPEYRES's PPPs
(... currency units of row-country per 1 currency unit of column-country)

	Reg.A	Reg.B	Reg.C
Reg.A	1.00000	.39001	.09163
Reg.B	2.66608	1.00000	.23959
Reg.C	11.50288	4.41811	1.00000

NUMBER of LASPEYRES-PPPs = 0 in the matrix = 0

INITIAL MATRIX of DIRECT BINARY FISHER's PPPs

(1st line: PPP = ... currency units of row-country per 1 currency unit of column-country,
  2nd line: L/P ratio)

	Reg.A	Reg.B	Reg.C
Reg.A	1.00000	.38247	.08925
	1.00	1.04	1.05
Reg.B	2.61458	1.00000	.23287
	1.04	1.00	1.06
Reg.C	11.20436	4.29419	1.00000
	1.05	1.06	1.00

NUMBER of the MISSING VALUES in the Fisher's-matrix = 0

Matrix of PPPs by EKS-method
(... currency units of row-country per 1 currency unit of column-country)

	Reg.A	Reg.B	Reg.C
Reg.A	1.00000	.38273	.08919
Reg.B	2.61278	1.00000	.23303
Reg.C	11.21207	4.29124	1.00000

Between-regional PPPs for GDP

	R. Hill version	S. Sergeev version
A	1.000	1.000
B on A	2.624	2.613
C on A	11.087	11.212

If the democratic average regional weights (% shares) are used in the combination with the between-regional BH-PPPs then it seems that the avearging of Regional numeraires in not necessary (between-regional PPP are invariant relatively the choice of a Region as a numeraire and the weights (shares) are not depend on any regional numeraire. So, the following data can be used to calculate between-regional aggregated PPPs

- democratic regional weights calculated average Regional shares for BHs as simple arithmetic mean from country's BH expenditure shares (like in Walsh or Tornqvist methods)

- the between-regional BH-PPPs

The Regions and the countries within the regions are treated fully symmetrically by this approach. There is no use of a common set of prices within the Regions for the aggregation of any indicator. Respectively there is no additivity problem by this approach and it seems that this method is consistent with the economic approach to index number theory (this is the GEKS method with between-regional PPPs and democratic shares).

II.2 Country Approach with Redistribution / Re-indexation (CAR) - Unrestricted EKS method

This approach was discussed actively during the TAG 2011 meetings.

The use of an unrestricted (i.e. without fixity[25]) approach on all countries first, and then the redistribution / re-indexation of the regional totals (Volumes or PPPs) according to regional results, to obatain the Global results with regional fixity

$$VI \text{ "Region / World" } x \, VI \text{ "Country / Region" } = VI \text{ "Country / World".}$$

    [Reg. Comparison]   [Global Comparison]

$$PPP \text{ „Country / World\_T" } =$$
$$PPP \text{ „Country / Region\_T" } * \, PPP \text{ „Region\_T / World\_T"}$$

    [Reg. Comparison]   [Global Comparison]

is not a new approach.

The redistribution of Volumes (with the G-K method) was used in the ICP 1980 [see Heston (1986)]. The use of whole set of the countries in the aggregation (universal calculation) with further distribution of regional Volumes in accordance with the Regional shares (two-stage approach) was very logical for the G-K method.

The fixity of PPP is the present Eurostat-OECD approach. The concrete Eurostat-OECD procedure is not fully symmetrical because the fixity is needed only for the EU and OECE countries but not for other country's sub-sets. However the main sense and content of the Eurostat-OECD approach is absolutely the same.

---

[25] The regional fixity can have in some cases remarkable impact of the results but, generally, this factor does not play extremely important role: usually several percent points (the revision of NA data brings sometimes much more significant changes.

These methods were named in the ICP 2005 Handbook as "Unweighted geometric scaling" (redistribution of PPPs) and "Weighted harmonic scaling" (redistribution of Volumes) due to technical details used here.

What indicator should be fixed – PPP or Volumes? [26] The analysis [Sergeev (2005) and R.Hill (2010), page 14] showed that the results obtained by these different approaches are not symmetrical. Obviously, this question concerns only between-regional comparisons because it is sufficient to fix any indicator (PPP or volume indices) in the intra-regional comparisons and other indicator will be fixed automatically due to the equality : Ie = Ip * Iv [expenditure ratio is the product of price ratio (PPP) and volume ratio].

R. Hill (2010) demonstrated in an explicit form that the fixity of PPPs can be presented as an optimization by altering the multilateral PPP by the minimum least-squares amount necessary to ensure that within-region fixity is satisfied (CAR-PPP).

S. Sergeev (2011) demonstrated that the fixity of the Regional Volume shares (CAR-Volume) has a natural and simple presentation in the terms of minimization of square of deviations of volume shares (weighted LSQ procedure) – see Annex 1.

So, from the point of view of formal mathematical features neither CAR-PPP nor CAR-Volume have a preference. However **the redistribution of regional Volumes should have an obvious contentual preference** because it is declared that the main aim of the ICP is Volume GDP comparison and we should attempt to optimize the obtaining of the Volume indices.

The use of whole set of the countries in the aggregation (universal calculation) with further distribution of regional Volumes in accordance with the Regional shares (two-stage approach) was very logical for the G-K method. However the advantages of this approach are not so clear if the EKS method should be used for the aggregation.

First of all, the EKS calculations in the Global ICP 2011 for more than 180 (each country with each country) would be cumbersome and not very transparent. Many bilateral indices (for the countries with different price and quantity structures) would be not very realistic and useful.

A. Deaton, "*Reshaping the world - The 2005 round of the International Comparison Program*", Princeton University, December 2010 formulated this very claerly: "... *the conceptual difficulties of all exercises of this kind, particularly when making comparisons between countries whose patterns of consumption and relative prices are radically different from one another. It is one thing to make PPP comparisons of France and Germany, or of Kenya and Tanzania, but we are altogether shakier ground when we come to compare Canada with Cameroon, Japan with Senegal, or Bolivia with Tajikistan. Such comparisons are difficult in theory, and subject to a wide margin of uncertainty in practice ...*

It is very doubtful that the comparison "Brazil/LA – Moldova/CIS" is meaningful and will bring useful information. It is not accident that experts mentioned weighting or elimination if there are egregious bilateral indices. Both versions (weighting and

---

[26] In the PPP aggregation software prepared several years ago for ESTAT by the author of this notice (this software is used also in the CIS comparisons) this choice is an option for users.

elimination of non-reliable bilateral indices) are possible but these procedures need weighting systems (some arbitrariness seems to be inevitable) and criteria for the elimination (like LPS). It is very likely that it will be not easy to agree on the following points:

- On which indicator should be based weighting? (coefficient of similarity?)
- Which bounds should be used for LPS if this approach is accepted?
  Should be unique bounds for LSP for all pairs (in this case, some countries like Tajikistan can have very few links at all) or these should be specific for different sets of countries?

All decisions here will be inevitable (at least, partly) arbitrary. However these points should be discussed if we want to use the Global EKS in a rational and efficient form but not in a mechanical way.

General considerations (*not the conclusions*)

1) It is desirable to use the same concept (between-regional approach or country approach) at the both levels: BH level and during the aggregation.

2) The between-regional approach (especially with symmetrical treatment of the regions) is an elegant and straightforward method but, probably, the practical problems can be solved easier (in more transparent form) by the Country approach (CAR approach)

3) If the Country approach is approved for the Global aggregation then the redistribution of regional Volumes (fixity of Volumes shares) has an obvious preference because it is declared that the main aim of the ICP is Volume GDP comparison and we should attempt to optimize the obtaining of the Volume indices

4) If the Country approach with the EKS method is approved for the Global aggregation then the serious consideration on the check of reliability of bilateral links should be done

5) If we moved from Representativity (some assumptions about different PLI for representative and non-representative products reflecting in the coefficient R) to Importance (some assumptions about different expenditure weights for important and less important products) then we should consider the moving to the weighted CPD instead of the CPR(I)D.

References

Diewert, W. E. (2004) "*On the Stochastic Approach to Linking the Regions in the ICP*", Discussion Paper No. 04-16, Department of Economics, University of British Columbia, Vancouver. http://www.econ.ubc.ca/diewert/icp.pdf

Diewert E. (2008) "*New Methodological Developments for the International Comparison Program*", Joint UNECE/ILO Meeting on Consumer Price Indices (Geneva, 8-9 May 2008). The latest version of this paper can be found on the web-site: http://www.econ.ubc.ca/diewert/dp0808.pdf

Diewert E. (2010) "*Approaches to Linking the Regions*", paper for the TAG ICP meeting (World Bank, 17-19 February 2010).

Dikhanov Y. (1994) „*Sensitivity of PPP-Based Income Estimates to Choice of aggregation Procedures*" (IARIW Conference; St.Andrews, New Brunswick, Canada on 21-27 August 1994). A 1997 version of this paper can be found on the WB ICP web-site http://siteresources.worldbank.org/ICPINT/Resources/icppapertotal.pdf

Dikhanov, Yuri (2007), "*Two Stage Global Linking with Fixity: Method 1 (EKS)*". World Bank, Mimeo

Heston, Alan (1986), "World Comparisons of Purchasing Power and Real Product for 1980". United Nations and Eurostat, New York: United Nations

Heston A., Summers R., Aten A. (2001) "*Price Structures, the Quality Factor, and Chaining*". WB / OECP PPP Seminar (Washington; 30.01-02.02.2001)

Heston A. ,Aten B. (2002) "Linking Country Groups in International Real Product and Purchasing Power Comparisons", , World Bank ICP Conference, Washington, D.C., March 11-15, 2002 http://pwt.econ.upenn.edu/papers/heston-atenwb.pdf

Heston A., B. Aten (2009) "*Are All Fishers Equal?*". http://pwt.econ.upenn.edu/papers/AH_BA_april30.pdf

Heston A. (2010) "Background Note on Productivity Adjustments as a Special Case of a General Regional Linking Problem", Paper prepared for ICP Technical Advisory Group, Oct'2010. http://siteresources.worldbank.org/ICPINT/Resources/270056-1255977007108/6483550-1257349667891/6544465-1285250669548/05.01_ICP-TAG04_ProductivityAdjustment.pdf

ICP Manual, Chapters 11-15. World Bank, 2005.

ICP Manual, Chapter 15 "*Linking PPPs and Real Expenditures for GDP and Lower Level Aggregates (Two-stage methods and scalar adjustments)*", World Bank, 2007 http://siteresources.worldbank.org/ICPINT/Resources/270056-1183395201801/icp_Ch15rev.doc

Hill R.J., Hill T.P. "*Regionalization and its Implications for Price Index Construction: The Case of the International Comparisons Program*", School of Economics UNSW, Sydney 2052, Australia, School of Economics Discussion Paper: 2007/28 http://wwwdocs.fce.unsw.edu.au/economics/Research/WorkingPapers/2007_28.pdf

Hill, Robert (2010), "*Linking the Regions in the International Comparisons Program at Basic Heading Level and at Higher Levels of Aggregation*". Paper prepared for ICP Technical Advisory Group.

Koves P. (1983) "*Index Theory and Economic Reality*", Akademia Kiado, Budapest.

Kravis, I.B., A.W. Heston, R. Summers (1982), „*World Product and Income, International Comparison of Real Gross Product*", Baltimore, John Hopkins Univ. Press

Prasada Rao D.S. (2001) "*Weighted EKS and Generalized CPD methods for Aggregation at Basic*

Heading Level and above Basic Heading Level" – a paper for the Joint WB / OECD Seminar on PPPs (Washington, 30.01-02.02.2001).

Sergeev S. (2001a) "*Aggregation methods on the basis of structural international prices*" (Joint World Bank - OECD Seminar on PPPs „Recent Advances in Methods and Applications", Washington, D.C.; 30.01-02.02 2001)
http://siteresources.worldbank.org/ICPINT/Resources/ss_WBOECD_v2004full.doc

Sergeev S. (2005a) "*Calculation of the results of the Eurostat GDP comparison with the use of fixity*" http://siteresources.worldbank.org/ICPINT/Resources/FIXITY_PPP_vs_ShVol_SS53.DOC

Sergeev S. (2005b), "The use of weights (indication of representativity) within the CPD and EKS methods at the basic heading level", Paper prepared for ICP Technical Advisory Group.
http://siteresources.worldbank.org/ICPINT/Resources/Use_of_Weights.doc

Sergeev, S. (2009a), "*Aggregation Methods Based on Structural International Prices*", pp. 274-297 in Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications, D.S. Prasada Rao (ed.), Cheltenham UK: Edward Elgar.

Sergeev, S. (2009b), "*The Evaluation of the Approaches Used for the Linking of the Regions in the ICP 2005*", unpublished paper, Statistics Austria, December.

Sergeev, Sergey (2011), "Fixity of Volumes (shares) as the Least absolute deviations (LAD) and Least square deviations (LSQ) procedures." Paper prepared for ICP Technical Advisory Group.

Fixity of Volumes (shares of Volumes) as the weighted Least square procedure (weighted LSQ)

The fixity of the Volume indicators can be presented in the terms of minimization of square of deviations of volume shares - as weighted LSQ procedure. Let us assume that we want to minimize the sum of the square devitaions between country's shares in Word real expenditure obtained without and with fixity. This can be written mathematically in the following form:

Let us have **Sh(A$_j$)_GUnfix** – share of country j from Region A in the World real expenditure obtained in the unrestricted Global comparison [**Sh(B$_k$)_GUnfix** and **Sh(C$_l$)_GUnfix** are the same countries for countries k and l in regions B and C).

To have the fixity of Volume shares, we should take into account the results of Regional comparisons as the constraints. The within-region fixity for Volumes shares constraints can be written as follows[27]:

$$Sh(A_j)\_GFix = \textbf{Sh\_R}_A * Sh(A_j)\_R$$

$$Sh(B_k)\_GFix = \textbf{Sh\_R}_B * Sh(B_k)\_R$$

$$Sh(C_l)\_GFix = \textbf{Sh\_R}_C * Sh(C_l)\_R$$

where

Sh(A$_j$)_GFix – share of country j from Region A in the World real expenditure obtained in the Global comparison with fixity (similar parameters for Regions B & C)

Sh_R$_A$ – (unknown) share of Region A in the World real expenditure obtained in the Global comparison with fixity (similar parameters for Regions B and C)

Sh(A$_j$)_R – share of country j from Region A in the Regional real expenditure obtained in the original Regional comparison (similar parameters for Regions B & C)

If we want to obtain the final volume shares in the global world comparison in such way that these (in total) are deviate minimally (in the terms of square deviation) from volume shares in the original Global comparison (all countries simultaneously) and simultaneously to keep the fixity of the regional shares then we should choice the values of variables **Sh_R$_A$**, **Sh_R$_B$**, **Sh_R$_C$** which **minimize the sum of the square deviations[28]**:

$$_{Na}\Sigma(Sh(A_j)\_GUnfix - Sh(A_j)\_Gfix]^2 + {_{Nb}}\Sigma[Sh(B_k)\_GUnfix - Sh(B_k)\_Gfix]^2 +$$

$$+ {_{Nc}}\Sigma[Sh(C_l)\_GUnfix - Sh(C_l)\_Gfix]^2 \Rightarrow \textbf{min}$$

To equalize the impact of the countries with higher and lower shares within the Regions, the weighting is desirable. **The weighting inversly proportional of shares of Regional shares of countries** - Sh(A$_j$)_R, Sh(B$_k$)_R, Sh(C$_l$)_R seems to be appropriate for this purpose.

---

[27] These constraints are some analogues of the R.Hill constraints for PPPs – see equations (4)-(6) in Hill (2010). Sh_R$_A$, Sh_R$_B$, Sh_R$_C$ – analogues of the parameters α, β, γ in R.Hill approach.

[28] The absolute deviations are more natural for the shares than the squares of deviations. Additionally, these are more understandable and transparent for users.

So, the final function for the minimization is the following:

$$_{Na}\Sigma[Sh(A_j)\_GUnfix - \textbf{Sh\_R}_A * Sh(A_j)\_R]^2/Sh(A_j)\_R +$$

$$+ \,_{Nb}\Sigma[Sh(B_k)\_GUnfix - \textbf{Sh\_R}_B * Sh(B_k)\_R]^2/Sh(B_k)\_R+$$

$$+ \,_{Nc}\Sigma[Sh(B_k)\_GUnfix - \textbf{Sh\_R}_c * Sh(C_l)\_R]^2/Sh(C_l)\_R \Rightarrow \textbf{min}$$

Solving this relatively **Sh_R$_A$**, **Sh_R$_B$**, **Sh_R$_C$**, the following 1st order conditions are obtained:

$$-2* \,_{Na}\Sigma[Sh(A_j)\_GUnfix * Sh(A_j)\_R)/Sh(A_j)\_R] + 2*\textbf{Sh\_R}_A \,_{Na}*\Sigma[Sh(A_j)\_R^2/Sh(A_j)\_R] = 0$$

$$-2*_{Na}\Sigma[Sh(B_k)\_GUnfix*Sh(B_k)\_R)/Sh(B_k)\_R] + 2*\textbf{Sh\_R}_B \,_{Na}*\Sigma[Sh(B_k)\_R^2/Sh(B_k)\_R] = 0$$

$$-2*_{Na}\Sigma[Sh(C_l)\_GUnfix * Sh(C_l)\_R)/Sh(C_l)\_R] + 2* \textbf{Sh\_R}_A*_{Na}\Sigma[Sh(C_l)\_R^2/Sh(C_l)\_R] = 0$$

The 1st equation for variable **Sh_R$_A$** can be simplified to the folloowing form:

$$- \quad _{Na}\Sigma Sh(A_j)\_GUnfix + \textbf{Sh\_R}_A * _{Na}\Sigma Sh(A_j)\_R = 0$$

Taking into account that $_{Na}\Sigma Sh(A_j)\_R = 1$ (sum of country's shares within Regional comparison is 1), variable **Sh_R$_A$** should have the following value:

$$\textbf{Sh\_R}_A = _{Na}\Sigma Sh(A_j)\_GUnfix$$

The same for the variables **Sh_R$_B$** and **Sh_R$_C$**

$$\textbf{Sh\_R}_B = _{Nb}\Sigma Sh(B_k)\_GUnfix$$

$$\textbf{Sh\_R}_C = _{Nc}\Sigma Sh(C_l)\_GUnfix$$

Taking into account that $\Sigma Sh(A_j)\_GUnfix$ is **Sh_GUnFix_R$_A$**, etc., we have the final relations:

Sh_R$_A$ = Sh_GUnFix_R$_A$

Sh_R$_B$ = Sh_GUnFix_R$_B$

Sh_R$_C$ = Sh_GUnFix_R$_C$

It means if we want to minimize the square of deviations between country's Volume shares (unfixed and fixed World results) then the Regional volumes obtained during the global aggregation should be redistributed in accordance with the country's shares obtained in the Regional comparisons.