### 2019

# **IARIW-HSE**

Special IARIW-HSE Conference "Experiences and Future Challenges in Measuring Income and Wealth in CIS Countries and Eastern Europe" Moscow, Russia, September 17-18, 2019

## **PPP-Based Stratification of CIS-EU/OECD Economies**

Andrey Kosarev

Paper Prepared for the IARIW-HSE Conference Moscow, Russia, September 17-18, 2019 Session 5: Purchasing Power Parity and Input-Output Time: 8:30 – 10:00, September 18

## PPP-based stratification of CIS–EU/OECD economies

#### **Modern computations of PPPs**

The Global International Comparison Program (Global ICP) is a unique worldwide statistical project in which almost all countries participate. The implementation of this project is regulated by the UN Statistical Commission and coordinated by the World Bank (e.g. see [8], [9], [10], [11]). The full cycle of required work – from collecting primary information to calculating purchasing power parities (PPPs) – takes several years, and until recently the comparison cycles were mainly carried out at intervals. Up to the present time two full-scale Global ICP cycles have been conducted, based on data for 2005 and 2011, and the work on the third global cycle, based on data for 2017, is being completed currently. CIS countries participated in both cycles of the Global ICP (2005 and 2011) and are participating in the current third cycle. A regional comparison cycle was also conducted in the CIS region based on data for 2014 – the CIS ICP 2014. The CIS Statistical Committee acts as a regional coordinating agency for PPP-based comparisons.

Within the framework of the Global ICP, all countries of the world are divided into six main regions. Five regions were formed on a geographical basis (Africa, Asia, the CIS, Latin America and the Caribbean, Western Asia) and one region on an institutional basis (a joint group of EU/OECD countries)<sup>1</sup>. Each region carries out the full amount of required work which results in the calculation of the purchasing power parities for the countries within the region. Then the World Bank combines the regional results into a single global PPPs system by means of summary calculations (e.g. see [12]).

For 2014, global comparisons were not conducted, but separate regional comparisons were undertaken and the results for two regions – the CIS and the EU/OECD – are available. For the CIS countries, the purchasing power parities to the Russian ruble and for the EU/OECD countries, PPPs to the Euro and the US dollar were calculated. In order to combine these two groups of results, the CIS Statistical Committee jointly with World Bank experts, developed a special procedure called a "partially-multilateral comparison" (PMC-procedure) (e.g. see [5]). Its application made it possible to obtain a methodologically harmonized set of PPPs to the US dollar for a fairly large group of economies (54 countries) —the combined group of CIS, EU and OECD countries (CIS – EU/OECD). Detailed results have been published, e.g. in [1].

This group of countries is quite diversified: large economies are well presented as well as developing economies which until recently were referred to as "transitional" ones. The possibility of conducting comparisons on a uniform statistical basis within such a group is of great interest. Stratification, formation of more homogeneous subgroups provides a tool that allows one to review the development trends in similar economies when drafting economic policies.

<sup>&</sup>lt;sup>1</sup> The choice of the EU/OECD region is due to the fact that PPP-based comparisons in this group of countries have been conducted for a long time, and by the start of the Global ICP, interactions between countries have already been established here, including exchange of the methodological and organizational solutions that support the management needs of the EU and OECD governing bodies.

#### **Moving Linear Segment procedure**

There are many methods designed to sub-select more homogeneous groups within a large set of points. One approach is based on assessing the similarity of the economic structures of the countries under comparison. Similarity factors, in turn, can be computed using various techniques. One of the effective solutions for computing the similarity factors was proposed by S. Sergeev (e.g. see [6]) and was used to support the grouping of countries in [5].

Besides selecting groups of structurally homogeneous economies it is interesting to subselect homogeneous groups based on specific economic tendencies peculiar for each group. When the same indicator in different groups has different trends, the subdivision of the general set of points into such groups gives grounds for an independent analysis of this indicator. This allows a better prediction of the respective indicator's changes by knowing which group the economy in question belongs to. Subdivision of the general set of points into homogeneous groups from the point of view of their internal tendencies ("tendencyhomogeneous") can be based on the analysis of a cross-section scatter plot. If such groups exist in reality, then they can be selected out of the general set of points – each of them is characterized by its own tendency of the indicator in question.

For example, analysis of the scatter plot relating GDP per capita and other PPP-based macroeconomic indicators shows several important tendencies for the combined ICP group of CIS–EU/OECD economies. The ratios reflecting consumer behavior, GDP structure, as well as some other indicators have stable tendencies to change in line with the GDP per capita level. It should be noted, that not only common tendencies for the whole group of economies can be mentioned, but sub-tendencies can be argued too: weaker and stronger economies have different angles of inclination of detected common tendencies.

In theory, several fine and effective sophisticated methods, e.g. like LOESS/LOWESS (e.g. see [2]), might be considered as being appropriate. But for the purposes of a practical macroanalysis, the use of very subtle methods based on complex calculations is not always necessary. In developing proposals for economic policy, it is critically important to be able to reasonably assess the tendency of changes that will result from the proposed possible measures of economic policy. To do this, it suffices to sub-select [tendency-homogeneous] groups using linear approximation. Within the whole set of countries, it is enough to subselect groups, each of which is characterized by its own, linearly approximable tendency for the analyzed indicator. In addition, the use of sophisticated calculation methods can often be unjustified due to the lack of required statistical data.

A special "moving linear segment" procedure (MLS-procedure) is proposed below for identifying "structural break points" separating different groups within the ICP set in the CIS-EU/OECD comparison.

The MLS procedure assumes that the analyzed set consists of multidimensional points, and if all the points are ordered by one of the coordinates – conditionally referred to as a "key coordinate" – then the values of some other coordinates will also show a tendency to ordering (in an ascending or descending order). Those coordinates are correlated with a key coordinate. This correlation may have different parameters in different parts of the ordered source set. Each part of the original set where the coefficients of dependence of the coordinate under study on the key coordinate differ significantly from the other parts, can be considered as a [tendency-homogeneous] group of points within the general set.

So, in the ICP set of CIS – EU/OECD countries sorted in ascending order by the "key" PPPbased indicator – GDP per capita – one can sub-select groups, each of which has its own tendency of the PPP-based macro-indicator in question which depends on the key indicator changes. Considering that the set under consideration is not too large and includes 54 points it seems reasonable to check the possibility of sub-selecting three tendency-homogeneous groups. Conditionally: weaker economies ("low-income countries"), mid-level economies ("medium-income countries"), and the most powerful economies ("high-income countries").

The MLS procedure includes the following steps:

1) The whole set of the multidimensional points under consideration is ordered by one of the coordinates – a key indicator which can be any indicator available for all points (for example, when analyzing the potential relationship of certain macro indicators with the GDP per capita level, the latter acts as a key indicator);

2) On the ordered set, a linear relationship between the indicator in question and the key indicator is estimated (for example, the dependence of the price level index (PLI) on the GDP per capita). For this, within the general set a sequence of segments with the same number of points is established. Each subsequent segment differs from the previous one by a shift of 1 point. According to practices, it is better to choose the number of points in segments taking into account the number of groups that a researcher expects to single out within the whole set (i.e., when searching for three homogeneous groups, it is effective to aim at the size of segments which is approximately one third of the total number of points). It is convenient to have an odd number of points in a segment so as to refer each segment to its central point. Calculations of linear regression describing the dependence of the indicator in question on the key indicator are performed for each segment. The result is a sequence of linear regression equations each of which is based on the same number of points and is shifted relative to the previous one by 1 point – a "moving linear segment" sequence.

3) The set of the estimated linear regressions is reviewed to identify more homogeneous groups. If in reality within the general set of points there are groups that have different coefficients in the linear regression equation describing the relationship between the key indicator and the indicator in question, then this can be noticed on the chart. Within the obtained set of linear regression segments, such groups demonstrate similarity in respect to regression angles. The point on the key indicator axis, on both sides of which there are groups of regressions with different prevalent angles, represents an assumed boundary between the adjacent groups of points: these adjacent groups of points demonstrate a different relationship between the key indicator and the indicator in question and different trends in the studied indicator are characteristic depending on the change in the key indicator. To detect such a structural break point in the dependency of the indicator in question on the key indicator, the ratio of the prevailing regression angles to the right and left of this point is used:

sbc(i) = (a(j+1) + ... + a(j+k)) / (a(j-1) + ... + a(j-k))

j = (k+1), , (J-k)

*sbc(j)* structural break coefficient at the point *j* 

- a(j) linear regression angle for the segment with the center at the point j
- *k* a number of the linear regression segments to the right and to the left of the point *j*
- J a total number of points in the given set

The result is a sequence of *sbc(j)* values for the indicator in question.

4) It is clear the if *j* is not a point of a structural break, the angles of the linear regression segments to the left and to the right of *j* are close in terms of quality, and in the limiting case sbc(j) = 1. If sbc(j) > 1, then linear segments to the right of this point on average have larger angles than those to the left of it (similarly, if sbc(j) < 1, then linear segments to the right of this point on average have smaller angles than to the left of it). If sbc(j) increases with the

increase of *j*, it means that the slope of the linear regression segments increases (similarly, the opposite is true). It is very important to detect local extremes in the obtained sequence *sbc(j)*, if they exist. To the left and to the right of each extremum *sbc(j)* the relations between the indicator in question and the key indicator are qualitatively different. The value of the key indicator in such point is a benchmark for breaking the whole set into homogeneous (in terms of the indicator in question) groups. All groups have their own tendencies in changing the studied indicator depending on the changing of the key indicator, so we call them "tendency-homogeneous" groups.

The proposed MLS procedure is outlined briefly in its most general form, since this paper focuses not so much on a detailed discussion of technical solutions as on the presentation of general characteristics of the stratification of the CIS – EU/OECD economies based on PPP data.

#### Stratification in the context of Harrod-Balassa-Samuelson effect

Use of PPPs in the economic analysis enables us to detect and examine many important social and economic patterns. Among them there is a well-known Harrod-Balassa-Samuelson effect, (the HBS-effect, e.g. see [7]). The growth of GDP per capita is accompanied by the growth of the ratio of the currency exchange rate to PPP (Price Level Index, PLI). A number of studies also show that this general tendency may have different features in different economies, i.e. the rate of growth of the price level index with the growth of per capita GDP depends on which income group the economy in question belongs to (e.g. see [4]). The analysis shows that such stratification can be carried out for 54 CIS–EU/OECD countries that participated in PPP-based comparisons from 2014 data. A chart illustrating the correlation of the GDP per capita (GPDpc) and PLI for these countries as a whole is presented in *Fig. 1*.



Fig.1. 2014 ICP CIS–EU/OECD economies: PLI and GDP per capita.

In the given set of data, three groups of economies can be distinguished with different rates of PLI growth in relation to the GDPpc growth. Applying the MLS procedure mentioned above, one can obtain a set of linear segments of the dependence of PLI growth on GDPpc growth shown in *Fig. 2*.



*Fig.2. 2014 ICP CIS–EU/OECD economies: PLI and GDP per capita, moving linear segment.* 

As one can see from the chart (*Fig. 2*), there are reasons to assume that there are at least three groups of economies with different relations between PLI changes and GDPpc changes. The turning points are approximately in the centers of the circles drawn at the chart. By applying the MLS procedure, we can obtain a sequence of values for the *sbc (j)* criterion. Local extrema in this sequence indicate structural break points, i.e. those the points to the left and to the right of which the dependence of PLI on GDPpc in a generalized form has different slopes.

The analysis shows that among 54 countries of the 2014 ICP CIS – EU/OECD, three groups of economies can be distinguished, so to speak: "low-income", "middle-income" and "high-income", each of which has its tendency in the relation between PLI and GDPpc.

For weaker economies in 2014 ("low-income"), whose 2014 PPP-based GDP per capita did not exceed approx. USD 28.5 thousand, there is a rather moderate tendency for PLI to increase with the increase of GDPpc: by approximately 0.0016% with an increase in GDPpc by 1 USD (t = 8.99, R = 86.6%).

For "middle-income" economies, whose GDPpc, according to the same calculations for 2014, ranged from approximately USD 28.5 thousand to USD 41.5 thousand, there is a tendency for PLI to increase faster with the increase in GDPpc – by approximately 0.0031% with GDPpc increasing by 1 USD (t = 9.61, R = 89.5%).

Finally, for the strongest ("high-income") economies, whose GDPpc, according to the same calculations based on data for 2014, exceeded USD 41.5 thousand, an extremely moderate PLI growth with GDPpc increasing is observed – by approximately 0.0002% with GDPpc increasing by 1 USD (t = 1.23, R = 29.3%). We can actually say that for the economies in a high-income group, PLI is on average at a high level and does not have a distinct tendency to increase as GDP per capita increases.

*Fig. 3* presents the linear approximations of the dependence of PLI on GDPpc for all three selected tendency-groups.



*Fig.3. 2014 ICP CIS–EU/OECD economies: 3 tendency-groups based on the relationship between PLI and GDP per capita.* 

Thus, the application of the MLS procedure allows us to split the whole set of 2014 ICP CIS– EU/OECD countries into three tendency-groups that differ in terms of relationship between PLI and GDPpc changes. Such stratification is important for developing economic policies: depending on which tendency group the analyzed economy belongs to, there are grounds for a more accurate forecasting of price and exchange rates changes in relation to expected changes of GDPpc.

#### Stratification in the context of Engel's law

Engel's law (e.g. see [3]) states that as income rises, the proportion of income spent on durable goods and services increases while the proportion of income spent on food falls. The general trend in the structural change of consumer expenses according to Engel's law is well observed at the chart (*Fig.4*): for 54 countries of 2014 ICP CIS–EU/OECD the ratio of clothing and footwear expenses to food and non-alcoholic beverages expenses increases on average with the increase of GDPpc.



*Fig.4. 2014 ICP CIS–EU/OECD economies: ratio "expenses on clothing & footwear to expenses on food & non-alcoholic beverages" and GDP per capita.* 

Analysis based on the MLS procedure allows one to obtain a set of linear approximations for the moving (by GDPpc) segment presented in *Fig. 5*.



*Fig.5. 2014 ICP CIS–EU/OECD economies: ratio "expenses on clothing & footwear to expenses on food & non-alcoholic beverages" and GDP per capita, moving linear segment.* 

The resulting set of linear segments can be divided into three groups that differ from each other by the trend in the changes of the consumer spending structure (in this case, the ratio of clothing and footwear expenses to food and non-alcoholic beverages expenses) with income changes. The structural break points are located approximately in the centers of neighborhoods shown at the chart. The coordinates of these points can be estimated using the MLS procedure.

The group of low-income economies (whose 2014 PPP-based GDPpc did not exceed USD 28.0 thousand) is characterized by a smaller change in the structure of consumer expenses in relation to the increase in income: a growth of GDPpc by 1000 USD on average corresponds to an increase in the ratio of clothing and footwear expenses to food and non-alcoholic beverages expenses by 0.0067 (t = 3.88, R = 61.3%).

In the group of middle-income economies (PPP-based GDPpc from USD 28.0 thousand to USD 34.0 thousand), the tendency in the context of Engel's law is almost three times stronger: an increase in GDPpc by 1000 USD on average corresponds to an increase in the ratio of clothing and footwear expenses to food and non-alcoholic beverages expenses by 0.0180 (t = 4.03, R = 68.9%).

Finally, for high-income economies (whose 2014 PPP-based GDPpc exceeded USD 34.0 thousand), there is a weak tendency towards the change the consumer spending with the income growth: an increase in GDPpc growth by USD 1000 on average corresponds to an increase in the ratio of clothing and footwear expenses to food and non-alcoholic beverages expenses by 0.0046 (t = 2.24, R = 42.3%).

In general, the identified trends are presented in Fig. 6.



Fig.6. 2014 ICP CIS–EU/OECD economies: 3 tendency-groups based on the relationship between the ratio "expenses on clothing and footwear to expenses on food and non-alcoholic beverages" and GDP per capita.

Thus, the tendency described in Engel's law has different strength for low-, middle- and highincome economies. This is important to take into account when predicting the effects of economic policies for various groups of countries.

Engel's law can be also illustrated by another structural indicator. The ratio of expenses on household furnishings, equipment and maintenance to expenses on food and non-alcoholic beverages also increases as income rises (*Fig. 7*).



Fig.7. 2014 ICP CIS–EU/OECD economies: ratio "expenses on household furnishings, equipment and maintenance to expenses on food & non-alcoholic beverages" and GDP per capita.

Here, the analysis of PPP-based indicators using the MLS-procedure also gives reason to distinguish three groups, which are characterized by different trends in the growth of the ratio of expenses on household furnishings, equipment and maintenance to expenses on food and non-alcoholic beverages with the growth of income. These structural break points are located approximately in the centers of the circles shown in *Fig. 8*.



Fig.8. 2014 ICP CIS–EU/OECD economies: ratio "expenses on household furnishings, equipment and maintenance to expenses on food & non-alcoholic beverages" and GDP per capita, moving linear segment.

The group of low-income economies (whose 2014 PPP-based GDPpc did not exceed USD 29.0 thousand) is characterized by a moderate increase of the structural indicator with the increase in income: a growth of GDPpc by 1000 USD corresponds to an increase in the ratio of expenses on household furnishings, equipment and maintenance to expenses on food items on average by 0.0082 (t = 3.19, R = 53.0%)

For middle-income economies (USD 29.0 thousand < GDPpc < USD 45.0 thousand at 2014 PPP-based calculations) the tendency is stronger: an increase in GDPpc by 1000 USD corresponds to an average increase in the ratio of expenses on household furnishings, equipment and maintenance to expenses on food items by 0.0180 (t = 4.03, R = 68.9%).

For high-income economies (GDPpc > USD 45.0 thousand at 2014 PPP-based calculations) the tendency is weak and unstable: an increase in GDPpc by USD 1000 corresponds to an increase in the ratio of expenses on household furnishings, equipment and maintenance to expenses on food items on average by 0.0046 (t=1.79, R=40.9%).



Fig.9. 2014 ICP CIS–EU/OECD economies: 3 tendency-groups based on the relationship between the ratio "expenses on household furnishings, equipment and maintenance to expenses on food & non-alcoholic beverages" and GDP per capita.

So, a cross-country analysis of PPP-based indicators supports the extended use of Engel's law and makes it possible to justify the stratification of economies based on the ratio of expenses on household furnishings, equipment and maintenance to expenses on food items.

#### PPP-based stratification as a tool for developing economic policies

The analysis of various PPP-based indicators confirmed the relationship between their levels and change trends and the levels of the key indicator GDP per capita, and also allowed us to estimate the quantitative parameters of such relationships.

The obtained results show that for a number of important indicators it is possible to single out groups of countries that are homogeneous in terms of these trends. Using the GDP per capita as a key indicator, we can divide the total set of countries under consideration in 2014 ICP CIS–EU/OECD into three groups: low-, middle- and high-income economies. For the above examples, the boundaries of these groups are presented in the table below.

Table.1. Boundaries of the homogeneous tendency-groups, combined data of 2014 ICP CIS– EU/OECD

	GDPpc, USD	
	lower break point	upper break point
PLI	28 500	41 500
Clothing and footwear / Food and non-alcoholic beverages	28 000	34 000
Household furnishings, equipment and maintenance /		
Food and non-alcoholic beverages	29 000	45 000

One can see that the estimates of the GDPpc-boundary of low-income groups for almost all of these macro indicators are close – about USD 28.5 thousand. The GDPpc boundaries for high-income groups, in contrast, have a wider spread.

To present the results in the same dimension, the values of linear approximations were normalized by the average measure of all observations of the corresponding indicator. The charts for these normalized linear trends are presented in *Fig. 10*.



Fig.10. 2014 ICP CIS–EU/OECD economies: Trends, denominated by the corresponding average mean, for "Clothing & Footwear / Food & non-alcoholic Beverages", "Household Furnishings, Equipment and Maintenance / Food & non-alcoholic Beverages", "PLI" – in relation to GDPpc

It is significant that the normalized trends practically coincide for all considered indicators in the interval corresponding to low- and middle-income economies. So, at least, the border between less strong and fairly strong economies can be placed at a GDPpc level of about USD 28.5 thousand (in 2014 prices and PPPs). This means that there are reasons to use this grouping for a broader analysis of these countries. It is preferable to determine the group of the strongest economies for each indicator separately, given the wider range of estimates of its lower boundary.

To approximate the trends in the relationship between each indicator and the key indicator (GDPpc) the simplest model was used, only linear functions were estimated. The search was mainly aimed at selecting homogeneous tendency-groups that combine countries according to their economic potential on the basis of the criterion "less strong – reasonably strong – strongest". This approach is justified by several reasons:

First, from the point of view of applying the obtained results for the purposes of predicting the effects of economic policies, the vector of a trend matters more than accurate quantitative estimates. In this regard, linear approximation is an absolutely adequate solution.

Second, for the calculation of a large number of coefficients of a nonlinear model, a significantly larger number of reliable statistical data are needed than are available in the 2014 ICP CIS –EU/OECD data set (as in many other cases of practical macroeconomic calculations).

And third, for obtaining practical solutions there is no need to break down countries according to very detailed categories, since here we are interested in a general trend only.

The above stratification can serve as one of the effective tools for the development of economic policies. One of the factors for correctly forecasting the effects of the expected increase in per capita GDP is belonging of the economy in question to one of the identified groups. This should be taken into account in order to assess the most likely trends of some important indicators – the ratio of the price level and exchange rate, the structure of consumer demand, etc.

The quantitative estimates presented in this paper were obtained using the results of 2014 PPP-based comparisons for the countries of the CIS, the EU and the OECD. In the near future, the results of the PPP-based comparisons on the basis of 2017 data will be published, and all the calculations mentioned above can be repeated with more recent data. With the input of prices, the volume and structure of GDP in 2017, the results will be different, so the stratification parameters of countries for macroeconomic analysis and forecasting purposes will be updated.

#### References

 CIS-STAT: "2014 International Comparison of CIS Countries' GDP Based on Purchasing Power Parities" Interstate Statistical Committee of the Commonwealth of Independent States, Moscow, 2017

2. CLEVELAND William S.: "Robust Locally Weighted Regression and Smoothing Scatterplots", Journal of the American Statistical Association, Vol. 74, No. 368. (Dec., 1979), pp. 829-836

3. KAUS Wolfhard: "Beyond Engel's law – A cross-country analysis", The Journal of Socio-Economics Volume 47, December 2013, Pages 118-134

https://www.sciencedirect.com/science/article/abs/pii/S1053535713001418#!

4. KOSAREV A.E.: "Analysis and Forecasting Based on National Accounts and Balance of Payments: Development of Methods" "Statistika Rossii" [in Russian] Moscow, 2005, ISBN 5-902339-39-1

5. KOSAREV Andrey, SERGEEV Sergey "Linking ICP Regions with Dual Participation Using the 'Partiallymultilateral' Approach (2014 CIS ICP Experience)", Paper prepared for the 35th IARIW General Conference, Copenhagen, Denmark, August 20-25, 2018, <u>http://www.iariw.org/copenhagen/kosarev.pdf</u>

6. SERGEEV S.: "Measures of the similarity of the country's price structures and their practical application" UN Statistical Commission and UN ECE, Conference of European Statisticians, Consultation on the European Comparison Programme, Geneva, 12-14 November 2001

7. TICA Josip, DRUŽIĆ Ivo "The Harrod-Balassa-Samuelson Effect: A Survey of Empirical Evidence" University of Zagreb, Faculty of Economics and Business", Working Paper Series, Paper No. 06-07 (No. 06-7/686)

8. UN Economic and Social Council, Statistical Commission, Forty-seventh session, 8-11 March 2016: "Final report of the Friends of the Chair group on the evaluation of the 2011 round of the International Comparison Programme" United Nations E/CN.3/2016/9

9. UN Economic and Social Council, Statistical Commission, Forty-seventh session, 8-11 March 2016: "Report" United Nations E/2016/24-E/CN.3/2016/34

10. UN Economic and Social Council, Statistical Commission, Fiftieth session, 5–8 March 2019: "Report of the World Bank on the International Comparison Programme" United Nations E/CN.3/2019/14

11. UN Economic and Social Council, Statistical Commission, Fiftieth session, 5–8 March 2019: "Report" United Nations E/2019/24-E/CN.3/2019/34

12. WORLD BANK: "Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program – ICP" World Bank, 2013, ISBN 978-0-8213-9728-2