Comparisons of GDP Over Time and Across Space: The State of the Art

GUIDO FERRARI

University of Florence, School of Statistics, Viale Morgagni, 59, 50134 Firenze, Italy. E-mail: ferrari@disia.unifi.it

ABSTRACT

In this paper, the GDP comparisons over time and across space are analyzed in an integrated time-space approach, to highlight the duality of the topic that allows a comprehensive treatment of the issue. The properties of the real GDP ratios are discussed, including the deflation and the exchange rate use, as well as the GDP PPPs and the International Comparison Program (ICP) framework.

Keywords: Real Gross Domestic Product, International comparisons, GDP PPP, Deflation, Exchange Rate.

Comparaciones del PIB a lo largo del tiempo y en el espacio: El estado del arte

RESUMEN

En este trabajo, las comparaciones del PIB a lo largo del tiempo y en el espacio se analizan en un enfoque de tiempoespacio integrado, para destacar la dualidad del tema que permite un tratamiento integral del problema. Se discuten las propiedades de los coeficientes del PIB real, incluida la deflación y el uso del tipo de cambio, así como las PPP del PIB y el marco del Programa de Comparación Internacional (PCI).

Palabras clave: Producto interno bruto real, Comparaciones internacionales, PIB PPP, Deflación, Tipo de cambio.

JEL Classification: E00, E31, C43

Artículo recibido en octubre de 2017 y aceptado en diciembre de 2017 Artículo disponible en versión electrónica en la página www.revista-eea.net, ref. ə-36110

1. INTRODUCTION

Comparing global or per capita Gross Domestic Product (GDP) over time and across space is still a very debated matter, also because of the controversy about the interpretation of GDP as an indicator of well-being (see, for example, Costanza *et al.*, 2009; Press, 2011; Jones and Klenow, 2016)¹.

The comparison is performed through GDP ratios. In time domain, with the objective of measuring the GDP rate of increasing or decreasing. Indeed, GDP growth rate is regarded as a measure of the economic growth of a country, or region, or whatever else space (Henderson et al., 2012; Boundless Economics, 2016); or, even, of a cross section of countries (Barro, 1991). According to Amadeo (2017), the GDP growth rate is the most important indicator of economic health. Every government, both at a national and a sub-national scale, but also the international organisms, institutions and bodies (WB, 2016; IMF, 2017), the media (Khan, 2014), as well as the unions of countries, such as the European Union (EU), need to have this crucial information for their economic policy decisions (Lequiller, 2005). This interest is confirmed by Rambaldi et al. (2006), although they focus on a state-space real income comparison. These decisions regard both the domestic economy and the international relationships, which heavily riverberate on the economic and social life of the population. Enterprises and corporations too, namely the managers and the Chief Excutive Officers (CEO)s, although working at the microeconomic level, need to know this macroeconomic information². In space domain, the purpose is to check whether the GDP of a country (or whatever space) is higher or lower than that of another country. This issue revealed nearly a century ago, introduced by the need of comparing the GDPs of various countries in a framework of the United Nations (UN) and EU international aid to underdeveloped and developing countries, subsequently extended to transition countries, and of subsidies to EU countries and regions, and in view of the contributions due by each member country of UN or EU for supporting the community life and activity. It has been adequately approached after the World War II, basically thanks to the research carried out at Pennsylvania University by Kravis et al. (1978), Summers and Heston (1988; 1991), Heston and Summers (1996), opening the way to the International Comparison Program (ICP) and the Penn World Tables (PWT). As

¹ In a sense, strangely enough if one thinks of the considerable path that has been traveled and the achievements got in the field. The discussion fuels misunderstandings as well, even in the definition itself of the aggregate that some, including economists and economic statisticians, define as the "total production in a given time period", instead of "the total market value of all final goods and services produced in an economy".

² Comparing GDP over time is tantamount to the problem of a person wishing to know whether his current salary is higher or lower than that of 10 years before, or to know the actual value of a sum of money inherited 20 years ago, or having the need to adjust the rent of the house he lives in as required by law, and so on.

a result, comparisons of GDP per capita in developed and developing countries have been undertaken (Maddison, 1983), as well as international differences in growth rates have been investigated (Nuxoll, 1994).

The purpose of this paper is to retrace retrospectively and prospectively the path of GDP comparisons over time and across space, identify and discuss the weak points, and sketch the future development lines. Specifically, at Paragraph 2, GDP time comparisons will be analyzed and discussed, whereas at Paragraph 3 the same will be done for GDP space comparisons; in the Conclusion at Paragraph 4, the threads of the arguments carried out will be pulled.

2. GDP TIME COMPARISONS

2.1. Real term comparisons

For the sake of better approaching the topic, let's denote current price GDP as $GDP_{tt} = \sum_{i=1}^{n} p_{it} q_{it}$, where GDP_{tt} indicates GDP at time *t* evaluated at the price of time *t*, *n* the number of goods and services that compose the GDP, p_i the price of item *i* and q_i the quantity or volume of item *i*. Similarly, we denote constant price GDP as $GDP_{t0} = \sum_{i=1}^{n} p_{i0} q_{it}$, where it's easy to see that GDP_{t0} indicates GDP at time *t* evaluated at the prices of time *0*.

If one wishes at measuring the rate of variation of GDP from, say, year *t*-1 to year *t*, and takes the ratio of the two current price figures, $\sum_{i=1}^{n} p_{it} q_{it} / \sum_{i=1}^{n} p_{it-1} q_{it-1}$, in other words, he compares them, he would obtain a misleading information, as the result incorporates both the volume variation and the price variation, we could roughly define "GDP inflation". For example, a government records a 3% GDP increase by using this ratio, congratulates itself for the good result, just decides to allocate funds to alleviate poverty, and then discovers that the GDP inflation, measured by the price index of the GDP basket, was 3.2%, meaning that actually the volume of the GDP, or the real term GDP, instead of increasing, did decrease.

This means that time comparisons of GDP should be carried out at constant price, or in real terms, in order to isolate the price effect and express the GDPs of every year at the same price level. The importance of GDP real term comparisons is stressed, for example, by Ahmad *et al.* (2003), and Scott (2013). On the methodology for calculating the real GDP growth rate, see for example, Glen (2017) and UNICEF (2017). The issue of having real term GDP figures has been solved by the National Statistical Offices (NSO)s, which since long time elaborate them, besides the GDP current price figures. Thus, the ratio $\sum_{i=1}^{n} p_{it-1} q_{it-1} q_{it-1} q_{it-1} q_{it-1} q_{it-1} q_{it-1}$ gives the correct response, that is, the variation of the sole volume of what has been produced additionally in one year. Of course, the above can be generalized: multiple GDP comparisons 0/t, 0/t-1, 0/t-2, t-2/t, t-1/t, t-2/t-1, and so on, as well as the reciprocals, are possible.

If one takes a basis for comparisons, say time 0, and makes all the constant price GDP fixed base ratios up to time t, he obtains: $\frac{GDP_{10}}{GDP_{00}} = \frac{\sum_{i=1}^{n} p_{i0}q_{i1}}{\sum_{i=1}^{n} p_{i0}q_{i0}}, \dots, \frac{GDP_{s0}}{GDP_{00}} = \frac{\sum_{i=1}^{n} p_{i0}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}}, \dots, \frac{GDP_{s0}}{GDP_{00}} = \frac{\sum_{i=1}^{n} p_{i0}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}}, \dots, \frac{GDP_{t-10}}{GDP_{00}} = \frac{\sum_{i=1}^{n} p_{i0}q_{it-1}}{\sum_{i=1}^{n} p_{i0}q_{i0}}, \dots$

It is easy to see that these GDP ratios are transitive, due to the invariance of the base GDP. So, moving base ratios can be obtained by dividing each ratio by the pevious one: $\frac{GDP_{s0}}{GDP_{00}}$: $\frac{GDP_{s-10}}{GDP_{00}} = \frac{\sum_{i=1}^{n} p_{i0}q_{is}}{\sum_{i=1}^{n} p_{i0}q_{i0}}$: $\frac{\sum_{i=1}^{n} p_{i0}q_{is}}{\sum_{i=1}^{n} p_{i0}q_{i0}} = \frac{\sum_{i=1}^{n} p_{i0}q_{is}}{\frac{DP_{s0}}{GDP_{s0}}} = \frac{\sum_{i=1}^{n} p_{i0}q_{is}}{\frac{DP_{s0}}{GDP_{s0}}}$ It easy to see that transitivity implies time reversibility: $\frac{GDP_{s0}}{GDP_{00}} = \frac{1}{\frac{GDP_{s0}}{GDP_{s0}}}$. Indeed, from transitivity condition $\frac{GDP_{r0}}{GDP_{00}} \cdot \frac{GDP_{s0}}{GDP_{r0}} = \frac{GDP_{s0}}{GDP_{00}}$, if we put r=s, we obtain $\frac{GDP_{s0}}{GDP_{00}} \cdot \frac{GDP_{s0}}{GDP_{00}} = \frac{GDP_{s0}}{GDP_{00}}$.

2.2. Deflation

There are countries where constant price GDP is not elaborated. Similarly, at a sub-national scale, constant price GDP is not available. In this events, should GDP comparisons be needed, constant price GDP figures can be obtained from current price GDP figures. As the problem consists of obtaining $\sum_{i=1}^{n} p_{i0} q_{it}$ from $\sum_{i=1}^{n} p_{it} q_{it}$, the most immediate and trivial thing that would come to mind is to divide $\sum_{i=1}^{n} p_{it} q_{it}$ by the ratio $\frac{\sum_{i=1}^{n} p_{i0} q_{it}}{\sum_{i=1}^{n} p_{i0} q_{it}}$, which is but a GDP Paasche type price index.

This heuristic way of approaching the issue demonstrates that constant price GDP can be obtained through a price isolation procedure known as "deflation", using a price index called "deflator". The problem is that GDP Paasche type price indexes are not available since they are not elaborated anywhere. Nonetheless, when needed, GDP deflation is obtained by scholars and stakeholders in national and regional micro-macro modeling, generally by taking the Consumer Price Index (CPI) as the deflator, thus introducing a bias due both to the fact that the CPI basket is different from that of GDP - being just a subset - and the fact that the CPI is universally worked out as a Laspeyes type index. Sometimes, according to circumstances, the ratio between the current price GDP and the constant price GDP of another country, as much as possible homogeneous to the concerned one, is used as the deflator. Indeed, this ratio is an implicitly obtained GDP Paasche type price index: this is the reason why it is called "GDP implicit deflator" or also "GDP implicit price index". Needless to stress, this imputation

is all the more approximate the more countries differ each other and even more if it is used at a regional scale³.

3. GDP SPACE COMPARISONS

3.1. Real term comparisons

Comparing GDP in space is a somewhat more complicated matter than it is in time domain. Once again for the sake of approaching the topic in an understandable easily way, let's suppose that we want to compare the GDPs of countries. To begin with, we assume that the countries under comparison are two, country A and country B. Similarly to what we have done in time domain, we can denote the GDPs of the two countries, at their own prices, as follows: $GDP_{AA} =$ $\sum_{h=1}^{k} p_{hA} q_{hA}$ and $GDP_{BB} = \sum_{s=1}^{g} p_{sB} q_{sB}$ respectively. A difference with respect to time domain appears at once: while in time comparisons the GDP basket is the same at any time, in space comparisons this is not, because the two countries have different production, consumption and investment structures, and consequently, GDPs have different item composition. Differently than in time domain, we cannot speak about GDP constant price figures: each country has its own GDP expressed at country's prices. If we imagined to have the power, we would arrange to obtain $GDP_{AB} = \sum_{h=1}^{k} p_{hB} q_{hA}$ and $GDP_{BA} = \sum_{s=1}^{g} p_{sA} q_{sB}$, where the former denotes the GDP of country A evaluated at the prices of country B, and the latter the GDP of country B evaluated at the prices of country A.

Similar to what is done in time domain, the comparison would be undertaken using the ratio of the two GDPs at the same prices. A further peculiarity emerges: while in time domain a natural sorting does exist in the sense that it's natural considering the ratio between the GDP of one year and the GDP of the previous year (even though in certain cases it may be worth inverting the timing), in space domain there is no natural sorting: each of the two countries can be put at the numerator. Thus, if we refer to country A, the comparison would be expressed by the ratio $\frac{GDP_{AB}}{GDP_{BB}} = \frac{\sum_{h=1}^{k} p_{hB} q_{hA}}{\sum_{s=1}^{g} p_{sB} q_{sB}}$, and by the ratio $\frac{GDP_{AA}}{GDP_{BA}} = \frac{\sum_{h=1}^{g} p_{AA} q_{AB}}{\sum_{h=1}^{k} p_{hA} q_{hB}}$, and, due to the above mentioned lack of a natural sorting of the two countries, if we refer to country B, by $\frac{GDP_{BA}}{GDP_{AA}} = \frac{\sum_{h=1}^{g} p_{SA} q_{SB}}{\sum_{h=1}^{h} p_{hA} q_{hA}}$ and $\frac{GDP_{BB}}{GDP_{AB}} = \frac{\sum_{h=1}^{k} p_{hB} q_{hB}}{\sum_{s=1}^{g} p_{SB} q_{sB}}$. The first and third ratios are GDP Laspeyres type indexes, the second and the fourth are GDP Paasche type indexes. Since they are not base reversal, the comparison is not univocally determined. This drawback can be overcome by putting a GDP

³ We have compared global GDPs. Obviously, it is possible to compare GDP per capita. This occurs and it is indeed very useful in particular contexts, for example, when we are interested in comparisons aiming at understanding the time development of a micro phenomenon by eliminating the population weight.

Fisher type ratio
$$\left(\frac{GDP_{AB}}{GDP_{BB}} \cdot \frac{GDP_{AA}}{GDP_{BA}}\right)^{\frac{1}{2}} = \left(\frac{\sum_{h=1}^{k} p_{hB} q_{hA}}{\sum_{s=1}^{g} p_{sB} q_{sB}} \cdot \frac{\sum_{h=1}^{k} p_{hA} q_{hA}}{\sum_{s=1}^{g} p_{sA} q_{sB}}\right)^{\frac{1}{2}}$$
. The GDP Fisher type index is base reversal, i.e., $\left(\frac{GDP_{AB}}{GDP_{BB}} \cdot \frac{GDP_{AA}}{GDP_{BA}}\right)^{\frac{1}{2}} = \frac{1}{\left(\frac{GDP_{BA}}{GDP_{AB}} \cdot \frac{GDP_{AB}}{GDP_{AB}}\right)^{\frac{1}{2}}}$.

therefore the comparison is univocally determined. It should be stressed that actually the above argumentation is but a mental exercise, as we do not have these kinds of real GDP figures since it's impossible to evaluate the items of one country at the prices of another country. Continuing with our exercise, we could take a GDP basket of goods and services common to both countries as a sample on which to calculate for each country the GDP at the prices of one of the two, thus eliminating the price level difference. This way, the comparison is univocal; indeed, if the prices are those of country B, we have $\frac{GDP_{AB}}{GDP_{BB}} = \frac{\sum_{h=1}^{k} p_{hB} q_{hA}}{\sum_{s=1}^{g} p_{sB} q_{sB}}$ and $\frac{GDP_{BB}}{GDP_{AB}} = \frac{\sum_{h=1}^{k} p_{hB} q_{hB}}{\sum_{s=1}^{g} p_{sB} q_{sA}}$ and base reversibility is ensured. The same if we take the

prices of country A.

The issue complicates further when the comparisons are regarded in a multiple framework, that is, when more than two countries are involved in comparison. In this case, it would be possible to take a GDP sample basket of goods and services common to all the *n* countries, and to evaluate the items at the same prices, which ensures transitivity, which in turn ensures base reversibility. Indeed, if the countries are three, A, B, and C, and the prices are those of country C: $\frac{GDP_{AC}}{GDP_{CC}} =$ $\frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jC}}; \frac{GDP_{BC}}{GDP_{CC}} = \frac{\sum_{j=1}^{m} p_{jC}q_{jB}}{\sum_{j=1}^{m} p_{jC}q_{jC}}; \frac{GDP_{AC}}{GDP_{CC}}: \frac{GDP_{BC}}{GDP_{CC}} = \frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jC}}: \frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jC}} = \frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jC}}: \frac{GDP_{BC}}{GDP_{CC}} = \frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jC}}: \frac{GDP_{BC}}{GDP_{CC}}: \frac{GDP_{BC}}{GDP_{CC}} = \frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jC}}: \frac{GDP_{BC}}{GDP_{CC}}: \frac{GDP_{BC}}{GDP_{CC}}: \frac{GDP_{BC}}{GDP_{CC}}: \frac{GDP_{BC}}{GDP_{CC}}: \frac{GDP_{CC}}{GDP_{CC}}: \frac{GDP_{CC}}{GDP_{CC}$ $\frac{GDP_{AC}}{GDP_{BC}} = \frac{\sum_{j=1}^{m} p_{jC}q_{jA}}{\sum_{j=1}^{m} p_{jC}q_{jB}}.$

Indeed, the above method of taking common GDP sample basket and evaluating the goods and services at the same prices is what has actually been done in the framework of an approach, called multilateral or multiple approach, used to make real GDP international comparisons, just consisting of choosing a common basket for all GDPs being compared and evaluating the goods and services of their baskets at the same prices.

A remarkable example of this approach is represented by the Geary-Khamis (GK) rationale. The GK multilateral transitive GDP comparison is based on a common GDP basket for all the *n* countries involved in the comparison. The price p_{ih} (i=1,...,n; h=1,...,k) of each good and service in the n countries is converted into a common currency called "Purchasing Power Standard" (PPS) by means of a conversion factor w_i $(i=1,...,n)^4$. For each good and service, the arithmetic mean of the price $w_i p_{ih}$, weighted with the quantity q_{ih} , is taken: $z_h = \frac{\sum_{i=1}^n w_i p_{ih} q_i}{\sum_{i=1}^n q_{ih}}$ (h = 1, ..., k), which is the average price of good or service h in the n countries. The conversion factor is in turn defined for each country as the arithmetic mean of the relative variation of the average price z_h/p_{ih} weighted with the value $p_{ih}q_{ih}$: $w_i \frac{\sum_{i=1}^h z_h q_{ih}}{\sum_{i=1}^n p_{ih} q_{ih}}$ (h = 1, ..., k). The GK transitive GDP comparison is defined by the ratio $GK \ GDP_{ii} = \frac{w_i}{w_{ii}}$, for any $i \neq i'$.

Another approach, called bilateral or binary, consists of choosing common GDP baskets in pairs, making comparison and making them globally multiple.

The Gini-Elteto-Koves-Szulc (GEKS) method is the key example of this rationale. The idea underlying the GEKS binary transitive GDP comparison is that of transforming GDP Fisher type index imposing to all the couples of comparisons squares condition: а least $\min_{\sum_{s=1}^{n} \log GEKS \ GDP_{i,s}GEKS \ GDP_{s,i'}} \sum_{s=1}^{n} \left(\log \ GEKS \ GDP_{i,s}GEKS \ GDP_{s,i'} - \right)$ log $F GDP_{i,s}F GDP_{s,i'}$, where *i* and *i'* are two countries under comparison and s denotes all the other countries involved in the comparison. As transitivity implies GEKS $GDP_{i,s}$ GEKS $GDP_{s,i'} = GEKS GDP_{i,i'}$, the above least squares condition can be re-written: $\min_{\sum_{s=1}^{n} \log GEKS \ GDP_{i,i'}} \sum_{s=1}^{n} (\log GEKS \ GDP_{i,i'} - \sum_{s=1}^{n} (\log GEKS \ GDP_{i,i'}))$ log $F GDP_{i,s}F GDP_{s,i'}$)². By deriving the above function with respect to $logGEKS GDP_{i,i'}$ and taking the first order conditions, after some simple mathematics, the transitive *GEKS GDP*_{*i*,*i*} = $\prod_{s=1}^{n} (F GDP_{i,s}F GDP_{s,i'})^{\frac{1}{2}} =$ $\left[\left(F \ GDP_{i,i'} \right)^2 \prod_{s \neq i,i'}^n F \ GDP_{i,s} F_{GDPs,i'} \right]^{\frac{1}{n}}$ comparison is obtained.

The above mentioned are two paradigmatic situations. Actually, for the GDP international comparisons other methods have been proposed. Since long time, the problem of the choice of the methodology is approached and solved in the International Comparison Program (ICP) framework, where the GDP comparisons are conducted under the GDP Purchasing Power Parities (GDP PPPs) denomination.

3.2. Using the Exchange Rate

A solution that has been resorted to in the past to obtain real GDP figures for international comparisons, and still used, is deflating, to evaluate all the GDPs

⁴ This conversion must be done even when the countries have the same currency, to take account of the differences in price levels.

at the same prices. The price index used in this case is the exchange rate. Actually, it is about acting on the aggregate as it is and expressing it in the same currency as the other aggregate. The currency normally used is the US\$.

The exchange rate is not a suitable GDP deflator. Basically, it's a price index based on the common product basket of the trade balance. This is a somewhat small sample and above all, it is not representative of the GDP basket. Moreover, it is subject to speculative maneuvers of the financial market and to the decisions and constraints of the politics. Therefore, it is a biased tool for representing the true price level ratio of the concerned countries (Ferrari and Zhao, 2016).

As a result, as stressed by Castles and Henderson (2005), from the pure economic point of view, "making cross-country conversions of GDP at actual exchange rates, rather than using PPP converters, causes the gap between rich and poor countries to appear as much greater, chiefly because the prices of non-traded goods are typically lower in poor countries".

The discussion about using the exchange rate to deflate GDP in international comparisons has known a renewed fervor in recent years still by Henderson (2015) who returned on the point of the best procedure for comparing real GDP as a still-unsettled issue in economics. This has fueled a debate attended by Henderson (2016) himself, Summers (2016), and Baneth (2016) who concluded that there can be no single 'true' method for intercountry comparisons of real product.

3.3. The International Comparison Program (ICP)

All the activity concerning the real GDP international comparisons – conducted using ratios named GDP Purchasing Power Parities (GDP PPP) - has been undertaken in the framework of the ICP (WB, 2013). Thus, the history of the comparisons proceeds parallel to the history of the ICP. To quote Eurostat-OECD (2006), the origins of international volume comparisons of GDP can be traced back to the experimental comparisons carried out by the Organisation for European Economic Cooperation (OEEC) in the 1950s, involving 9 countries. The reasons given for using GDP PPPs are well known, and stand in the commonly recognized inadequacy of the exchange rates to measure the space price level differences. In the late 1960s, the research work was continued, initially under the programme for International Comparisons of Output and Productivity (ICOP) initiated by Angus Maddison at Groningen University in the Netherlands.

Responsibility for the ICP was shared by the United Nations Statistical Division (UNSD) and the University of Pennsylvania, with the latter taking the leading role. The research was a co-operative effort involving many institutions and persons in the countries participating in the project. Both the Ford Foundation and the World Bank (WB) made major financial contributions. The first director

of the project was Irving Kravis of the University of Pennsylvania (Eurostat-OECD, 2006). The project's initial task was to develop a methodology which would serve for a comprehensive system of GDP international comparisons based on parities.

The next task was to test the methodology by making actual comparisons. Thus, the first three rounds or phases of the ICP - 1970, 1973 and 1975 - were essentially experimental in character. Comparisons were set up for a small group of countries representative of different income levels, social systems and geographical regions. In Round I, comparisons of GDP were made for ten countries for 1970. In Round II, the ten Round I countries were joined by six others. Comparisons covering all sixteen countries were made for 1970 and for 1973. Round III comparisons covered 34 countries and had 1975 as the reference year. The results of these and subsequent ICP comparisons were published with the US as reference country and the international dollar as numeraire (Eurostat-OECD, 2006).

After Round III, there were three major developments. First, the ICP became a regular part of the work programme of the UNSD, with the University of Pennsylvania advising on methodological issues. Second, Eurostat started to play an increasingly important role, organising the comparisons for the EU, providing technical and financial assistance to regional comparisons in Africa and encouraging the OECD to become involved in the work. The third and most significant development was the regionalisation of the ICP.

The need for regionalisation became evident in several respects during Round III. It was clear that, as the number of participating countries increased, a highly centralised scheme of organisation was no longer feasible, especially as there was no one international body in a position to manage it. Early results from Round III had indicated that countries within regions tended to form more or less homogeneous subsets and it was recognised that there were operational advantages in grouping countries according to their geographical proximity. Round III comparison for the EU had demonstrated that a comparison tailored to meet the specific requirements of a region need not jeopardise the larger comparison of which it is a part (Eurostat-OECD, 2006).

Regionalisation placed a greater share of the work on the regional organisations of the UN, namely: the Economic Commission for Europe (ECE), the Economic Commission for Latin America and the Caribbean (ECLAC), the Economic and Social Commission for Asia and the Pacific (ESCAP) and the Economic and Social Commission for Western Asia (ESCWA). This left UNSD at the centre to co-ordinate the regional comparisons and to ensure that they could be linked in a global comparison. Two rounds of the ICP were completed after regionalisation: Round IV which covered 60 countries and had 1980 as the reference year; and Round V which covered 64 countries and had 1985 as the

reference year. A third round, Round VI, covering 83 countries and having 1993 as the reference year, was started but not completed. Reports were published for the regional comparisons carried out in Africa, the Middle East, Asia and Europe.

At the global level, the regional distribution of countries that participated in the six rounds of the ICP shows a limited and uneven coverage. The results are published on average six year after the reference year.

After Round VI failed to produce a world comparison, the UNSD decided in 1997 that the ICP should be thoroughly reviewed before any further round was attempted. A consultant was recruited to establish whether or not the ICP should continue and, if it should, what improvements were required and how these improvements should be brought about. The consultant's report was presented to the UNSD in 1999, containing several remarks and suggestions, including that of not sanctioning a new round until at least the management and resource issues had been addressed.

The UNSD response to the report was to ask the WB - the de facto global coordinator of the ICP since 1993 - to propose a strategy for a comprehensive solution to the deficiencies identified by the consultant. The WB, in consultation with other interested agencies, drew up an implementation plan for a new round of the ICP.

The reference year for the new round is 2005. Data collection was scheduled to start towards the end of 2004. Regional comparisons were scheduled to be organised by the African Development Bank, the Asian Development Bank, ESCWA, ECLAC, Statistics Canada, the Interstate Statistical Committee of the Commonwealth of Independent States (CISSTAT), the State Committee of the Russian Federation on Statistics (Goskomstat-Russia), Eurostat and the OECD. A global office has been established at the WB to provide overall co-ordination and to ensure technical and procedural uniformity across the regions (Eurostat-OECD, 2006). The last ICP round was carried out in 2011.

One of the regional comparison programmes that resulted from the regionalisation of the ICP was the European Comparison Programme (ECP). It was launched at the 27th plenary session of the Conference of European Statisticians (Geneva, June 1979). The ECE is responsible for the ECP and publishes the results of its comparisons, but the actual comparisons are organised by other agencies. This is because the ECP is an amalgam of independent comparisons involving different groups of countries. The Eurostat-OECD PPP Programme is central to the ECP and brings with it coverage beyond Europe through the inclusion of non-European OECD Member Countries.

Six rounds of the ECP were conducted between 1980 and 1999/2000. Before 1996, the ECP covered two groups of countries. Group I consisted of the countries that were participating in the comparisons organised by Eurostat and the OECD for their joint programme. Group II consisted of countries from

central and Eastern Europe that were participating in the comparisons coordinated by Statistics Austria specifically for the ECP at the request of the ECE.

Group I comparisons were multilateral comparisons, based on a common basket of goods and services, with each country being compared directly with each of the other participating countries. Group II comparisons were bilateral comparisons with each country compared directly with Austria. Each bilateral comparison was based on a different basket of goods and services. Quality adjustments were made when it proved impossible to find strictly comparable goods and services. Adjustments were also made for differences in the productivity of producers of non-market services such as general public administration, health and education. Comparisons between Group II countries were made through Austria. Comparisons between countries in Group I and countries in Group II were also made through Austria.

The 1996 comparison was a turning point for the ECP. It had three important features. The first was that a third group of countries, Group III, was included in the ECP. The group was made up of the member countries of the Commonwealth of Independent States (CIS) - except Ukraine - plus Mongolia and Turkey. The Group III comparison was a multilateral comparison. It was managed by the OECD with the assistance of CISSTAT, Goskomstat-Russia and the State Institute of Statistics of Turkey. Countries in Group III were compared with countries in Group I and Group II through Austria. The second feature of ECP 1996 was that the comparisons in Group II were organised as a single multilateral comparison and not as a set of bilateral comparisons with Austria. The third feature was the start of the break-up of Group II. The Czech Republic, Hungary, Poland and the Slovak Republic – countries which originally had participated in Group II comparisons - moved to Group I. The Russian Federation and Slovenia, while remaining in Group II, also participated in the Group I comparisons on an experimental basis. By ECP 1999/2000, Group II no longer existed. The results of all ECP comparisons carried out between 1980 and 1996 were published with Austria as the reference country and the Austrian schilling as numeraire. After 1996, results have been presented with the EU as reference country and the euro as numeraire (Eurostat-OECD, 2006).

The first official comparison conducted by Eurostat was for 1975. It covered the nine countries that were EU Member States at that time. Subsequently, until 1990, Eurostat carried out comparisons every five years, in 1980, 1985 and 1990. These comparisons were principally for EU Member States and countries like Greece, Portugal and Spain that were in line for EU membership. But countries falling into neither of these two categories - such as Israel in 1980, Austria in 1980, 1985 and 1990, and Switzerland in 1990 - also participated. Austria's participation was occasioned by the need to provide a "bridge" between the EU

Member States and other countries in ECP Group I and the eastern and central European countries in ECP Group II. Throughout these first twenty years, the Programme was under the direction of Hugo Krijinse Locker.

After 1990, Eurostat adopted the "rolling benchmark approach" and started making annual comparisons. At the same time, the number of countries covered rose from 14 in 1991 to 19 in 1994 after which it stayed stable until 1999. The increase was due to the continuing enlargement of the EU and to countries of the European Free Trade Association (EFTA) harmonising their statistics and statistical programmes with those of EU Member States. In 1999, the number of countries covered by Eurostat comparisons rose to 31 (Eurostat-OECD, 2006).

Encouraged and assisted by Eurostat, the OECD began organising comparisons for those OECD Member Countries that were not already included in Eurostat comparisons in the early 1980s. The first comparison had 1980 as the reference year, but data collection did not start until 1983 and was limited to four countries. Two of these - Japan and the United States (US) - were major non-European economies of particular interest to the European Commission (EC) because of their trade and investment links with the EU. The results of this retrospective exercise were combined with those from the Eurostat comparison, thereby covering 18 of the OECD's 24 Member Countries.

Thereafter the OECD worked closely with Eurostat to effect comparisons for 1985 and 1990. When put together, the Eurostat and OECD comparisons covered 22 OECD Member Countries in 1985 and all 24 OECD Member Countries in 1990. It was during this time that formal agreements between the two organisations were made to establish the Eurostat-OECD PPP Programme. Eurostat and the OECD agreed to co-ordinate the data collections in two different groups of countries with the object of combining the data sets of the two groups in a single comparison.

From the beginning, Eurostat has published the results of comparisons with the EU as reference country and an artificial currency unit - the Purchasing Power Standard or PPS - as numeraire. (The PPS is, in effect, the "average" of the currencies of all EU Member States.) The OECD published the results of the 1980 and 1985 comparisons with the US as reference country and the international dollar as numeraire because coverage of OECD Member Countries was incomplete. From 1990, when all Member Countries started to participate in the Programme, the OECD has published results of comparisons with the OECD as reference country and OECD dollars as numeraire. (The OECD dollar is the "average" of the currencies of all OECD Member Countries.)

As regards methods for calculating the parities, prior to 1980, the ICP employed the Country-Product-Dummy (CPD) method to calculate GDP PPPs at the basic heading level and the Geary-Khamis (GK) method to aggregate the basic heading GDP PPPs up to the level of GDP. These methods were not

generally accepted at the time and their advantages and disadvantages relative to other methods were a subject of debate among experts. Regionalisation allowed the use of alternative methods. Eurostat and OECD comparisons have always used the GEKS method to calculate parities at the basic heading level.

For the 1975 comparison, Eurostat and OECD used an aggregation method other than GK, which, instead, was used for 1980 and 1985 comparisons.

Subsequently, Eurostat proposed the GEKS method to be used to aggregate basic heading parities because it provided volume indexes that were free of the Gerschenkron. The OECD, on the other hand, wanted to retain the GK method, because, while its volume indexes were not free of the Gerschenkron effect, the real final expenditures on which the indexes were based were additive - which GEKS real final expenditures are not - and, as such, more relevant to the type of analysis carried out by OECD economists. Consequently, in 1988 and again in 1989, the UNSD, the OECD and Eurostat jointly convened a meeting of experts to discuss aggregation methods. The experts recognised that comparison results serve many different purposes and that there was no one method of aggregation which can be considered satisfactory for all these purposes. They recommended the calculation and dissemination of two sets of results: one set to be aggregated using the GEKS method, the other to be aggregated using the GK method.

On the other hand, the results for EU Member States being used for administrative purposes - the allocation of structural funds - as well as for economic analysis, Eurostat selected the GEKS results as the official results for the EU. The OECD agreed to publish these first and to publish the GK results a year later. This allowed time for the GEKS results to be accepted as official and avoided any confusion that could arise from the simultaneous dissemination of two sets of results (Eurostat-OECD, 2006).

The debate on GDP international comparisons is still open. It is alimented by the huge, and ever increasing relevance kept by the topic, even more in the present globalization time, which makes increasingly complicated and integrated the economic and social relationship among countries and peoples and claims for even more refined tools for their balanced control and management and as much as possible equalitarian and capable to allow an adequate development to poor countries, to help them and allow to get suitable achievements in welfare and progress.

Besides the above quoted Eurostat-OECD Manual, many have been the contributions to the above debate, with the aim of finding out acceptable and shared solutions. First of all, the series of paper that refer, strictly or in a broader sense, to the activity of the University of Pennsylvania. Just to quote some of them, Kravis *et al.* (1975), Kravis *et al.* (1978, 1982). Also, the academic debate was deep and stimulating and was participated, among the many valuable

contributors, by Hill (1982), Gerardi (1982), Krijinse Locker (1983), Khamis (1984), Kravis and Lipsey (1991), Ferrari and Riani (1998).

According to Feenestra *et al.* (2013), for over four decades, the Penn World Table (PWT) has been a standard source of data on real GDP across countries. Making use of prices collected across countries in benchmark years by the International Comparisons Program (ICP), and using these prices to construct purchasing-power-parity exchange rates, PWT converts GDP at national prices to a common currency -U.S. dollars - making them comparable across countries.

4. CONCLUSION

We have seen that for making GDP comparisons both in time and in space domain there is a need of disposing of constant price, or real, figures.

In time domain, the problem of getting real term GDP figures has been solved by the National Statistical Offices (NSO)s, which, besides the GDP current price figures, provide since long time constant price GDP. In space domain, the GDP figures are obviously at the prices of the concerned places only. In this case, the problem of having real GDP figures might be solved by "deflating" them, that is, by measuring them with the same prices, using an opportune space price index. Such a price index is the exchange rate, and actually deflating by using the exchange rate is a procedure that, although incorrectly, has been widely used, and so still is it in some cases. Nevertheless, the right framework for GDP international comparisons is represented by the ICP, which does not aim at elaborating all the GDPs under comparison at the same prices, but elaborates sample GDP ratios at the same prices, instead - which are then expressed in US\$ called "GDP Parities", and therefore, provides direct comparisons.

But the field remains very steep and slippery, many issues are still open, the debate is alive, also due to some not easily explicable misunderstandings that have arisen.

At a sub-national scale, the NSOs elaborate within country regional real GDP figures that allow within country regional GDP time comparisons. Within country, cross region GDP comparisons, although possible and executed, are not significant, as the regional price levels are not the same. Regional GDP PPP figures are not elaborated by any NSO. Several countries have started programs aimed at elaborating regional households' expenditure PPPs that, as proxy, will allow deflating regional real GDP and making GDP regional comparisons. The ICP does not elaborate sub-national GDP parities; as a consequence, inter country, cross region GDP comparisons are not possible.

BIBLIOGRAPHY REFERENCES

- AHMAD, N., LEQUILLER, F., MARIANNA, P., PILAT, D., SCHREYER, P. and WÖLFE, A. (2003). "Comparing growth in GDP and labour productivity: Measurement issues". *Statistics Brief*, OECD, No.7, December.
- AMADEO, K. (2017). "What Is the GDP Growth Rate? *The Balance*-US Economy". https://www.thebalance.com/what-is-the-gdp-growth-rate-3306016
- BANETH, J. (2016). Comparing Real GDP Across Countries: Comment and Rejoinder. *Economic Affairs*, 36, pp. 84–88. *doi*: 10.1111/ecaf.12153
- BARRO, R.J. (1991). "Economic growth in a cross section of countries". *The quarterly journal of economics*, 106(2), pp. 407-443.
- BOUNDLESS ECONOMICS (2016). Economic Growth. Lumen Learning. https://courses.lumenlearning.com/boundless-economics/chapter/long-run-growth/
- CASTLES, I. and HENDERSON, D. (2005). "International Comparisons of GDP: Issues of Theory and Practice". *World Economics*, 6(1), pp. 55-66, January.
- COSTANZA, R., HART, M., POSNER, S. and Talberth, J. (2009). "Beyond GDP: The Need for New Measures of Progress". *The Pardee Papers*, No. 4, Boston University, January.
- EUROSTAT-OECD (2006). Eurostat-OECD Methodological Manual on Purchasing Power Parities. Luxembourg.
- FERRARI, G. and RIANI, M. (1998). "On Purchasing Power Parities Calculation at the Basic Heading Level". *Statistica*, LVIII, 1.
- FERRARI, G., and ZHAO, Y. (2016). "Exchange Rate and Purchasing Power Parities in Renmimbi Internationalization: A Look with Statisticians' Eyes". *Journal of Chinese Economics*, 4(1), pp. 62-82.
- FEENSTRA, R.C., INKLAAR, R. and TIMMER, M. (2013). "The Next Generation of the Penn World Table". NBER Working Paper No. 19255, July. http://www.nber.org/ papers/w19255.pdf
- GERARDI, D. (1982). "Selected Problems of Inter-Country Comparisons on the Basis of the Experience of the EEC". *Review of Income and Wealth*, 28, 4, December.
- GLEN, J. (2017). "How to Calculate Growth Rate of Real GDP". *Investor Guide.com*. www.investorguide.com/article/15785/how-to-calculate-growth-rate-of-real-gdp-d1412/
- HENDERSON, D. (2015). "Comparing Real GDP Across Countries: The Issues". Revisited. *Economic Affairs*, 35(2): 286-298. doi.org/10.1111/ecaf.12128.
- HENDERSON, D. (2016). "Comparisons of Real GDP Over Time and Across Space: Taking the Subject Further". *Economic Affairs*, 36: 373-378. *doi*: 10.1111/ecaf.12203
- HENDERSON, J.V., STOREYGARD, A. and WEIL, D.N. (2012). "Measuring economic growth from outer space". *American Economic Review*, 102(2), pp. 994-1028, April. doi: 10.1257/aer.102.2.994
- HESTON, A. and SUMMERS, R. (1996). "International price and quantity comparisons: potentials and pitfalls". *American Economic Review*, 86(2), May.
- HILL, P. (1982). Multilateral Measurements of Purchasing Power and Real GDP. Eurostat
- JONES, C.I. and KLENOW, P.J. (2016). "Beyond GDP? Welfare across Countries and Time". *American Economic Review*, 106(9), pp. 2426-2457, September.

- KHAMIS, S.H. (1984). "Application of Index Numbers in the International Comparisons and Related Concepts" *Review of Income and Wealth*, 30, 1, March.
- KHAN, M. (2014). "What does GDP really tell us about economic growth?" *The Telegraph*. 15 October.
- KRAVIS, A., KENESSEY, Z., HESTON, A. and SUMMERS, R. (1975). A System of International Comparisons of Gross Product and Purchasing Power. Baltimore: The John Hopkins University Press.
- KRAVIS, A., HESTON, A. and SUMMERS, R. (1978). *International Comparisons of Real Product and Purchasing Power*. Baltimore: The John Hopkins University Press.
- KRAVIS, A., HESTON, A. and SUMMERS, R. (1982) *International Comparisons of Real Product and Purchasing Power*. Baltimore: The John Hopkins University Press.
- KRAVIS, I.B., HESTON, A.W. and SUMMERS, R. (1978). "Real GDP per capita for more than one hundred countries". *The Economic Journal* 88.350, pp. 215-242. IMF (2017). *World Economic and Finacial Survey.*
- KRAVIS, B.I. and LIPSEY R.E. (1991). The International Comparison Program. Current Status and Problems. In International Economic Transactions: Issues in Measurement of Economic Research, Hoper, P. and Richardson J.D. eds., pp. 437-468.
- KRIJINSE LOCKER, H. (1983). "On the Estimation of Purchasing Power Parities on the Basic Heading Level". *Review of Income and Wealth*, 29.
- LEQUILLER, F. (2005). "Is GDP a satisfactory measure of growth?" OECD Observer, January.
- MADDISON, A. (1983). "A comparison of levels of GDP per capita in developed and developing countries, 1700-198". *The Journal of Economic History* 43.1, pp. 27-41.
- NUXOLL, D.A. (1994). "Differences in relative prices and international differences in growth rates". *American Economic Review*, 84, December, pp. 1423-36.
- PRESS, E. (2011). The Sarkozy.Stiglitz Commission's Quest to Get Beyond GDP. The Nation, May. https://www.thenation.com/article/sarkozy-stiglitz-commissions-quest-getbeyond-gdp/
- RAMBALDI, A.N., PRASADA RAO, D.S. and DORAN, H.E. (2006). "Constructing internationally comparable real income aggregates by combining Sparse benchmark data with Annual National Accounts Data. A state-space approach". *Working Papers and Studies*, Eurostat, June.
- SCOTT, A.W. (2013). "GDP: Does It Measure Up?" *Economic Research*, Federal reserve Bank of St. Louis, May.
- SUMMERS, L. H. (2016). "Comparing Real GDP Across Countries: Comment". *Economic Affairs*, 36, pp. 221–223. doi: 10.1111/ecaf.12184
- SUMMERS, R. and HESTON, A. (1988). "A new set of international comparisons of real product and price levels estimates for 130 countries, 1950-1985". *Review of Income and Wealth*, 34(1), March, pp. 1-25.
- SUMMERS, R. and HESTON, A. (1991). "The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988" *Quarterly Journal of Economics*, 106(2), May, pp. 327-68.
- UNICEF (2017). How to Calculate the Annual Growth Rate for Real GDP. https://www.fool.com/knowledge-center/how-to-calculate-the-annual-growth-rate-for-relg.aspx
- WB (2016). GDP Growth. World Bank National Accounts Data.

WB (2013). Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program-ICP. Washington, DC: World Bank. DOI:10.1596/978-0-8213-9728-2.