



[02.03]

Government Collective Services

Compensation and Productivity Adjustments

DRAFT

ICP Book

Global Office

3rd Technical Advisory Group Meeting

June 10-11, 2010

Paris, France

Table of Contents

Background.....	3
1.0 Treatment of Comparative Resistant Services in the 2005 ICP.....	4
1.a The Final 2005 Approach.....	4
1.b Effects of Productivity Adjustment on the Global Comparison	7
2.0 Experiments with Priced and Non-priced Services	8
3.0 Other Methods of Dealing with Non-priced Services up to the 1975 Round.....	12
3.a Health Personnel.....	12
3.b Education.....	13
3.c. Collective Government Expenditures on Personnel.....	15
4.0 Linking Issues for Regions in 2005	15
4.a Estimates of Overlapping Price Levels	16
5.0 A Proposal for Linking in 2011 and Conclusions.....	17
5.a Estimates of Productivity by Country	17
5.b Productivity Adjustment	18

Government Collective Services - Compensation and Productivity Adjustments

Background

This paper addresses the questions of how the ICP measures the output flowing from expenditures that are not directly purchased by final users. These are often referred to as non-priced or comparison-resistant services in the ICP literature, the main headings being education, health and collective government expenditures provided by governments or non-governmental organizations (NGOs). We begin in Section 1 with a summary of how the estimation was done in the 2005 ICP within each region. Where output is not market priced, one way to derive PPPs for the inputs of labor in these expenditure headings is to make salary comparisons for different skill and occupational employment categories based on international standard definitions. Using these salary comparisons as the equivalent of PPPs, then dividing expenditures by the PPP provides indirect quantity comparisons of these services. However, the method can produce improbable volumes for low salary countries such as Tajikistan, Yemen or Cambodia. The reason for this in part is because salaries are lower in lower-income countries because personnel are accompanied by less capital and are therefore producing less output than in higher-income countries. Lower productivity also results because there is little pressure on supervisors to use staff efficiently when their salaries are so low and the value of their output is so hard to measure, the case in government and education especially. This problem was recognized early on in ICP work when there were fewer countries but was even more evident in the 2005 ICP when the range of salaries across the 146 countries was much larger.

In the report on the 1975 ICP Kravis, Heston and Summers (KHS, 1982)) Chapter 5 was devoted to these comparison-resistant services. Two points from this chapter will be noted, the first dealing with evidence on whether the PPPs of priced services can be used to approximate the PPPs for comparison-resistant services. This is section 2 while Section 3 discusses some of the methods considered to deal with this problem in the early ICP rounds. Section 4 deals with how the regions were linked in the 2005 ICP and considers whether the linking could have taken into account the different methods used in the regions. Section 5 concludes with a discussion an alternative that might be considered for the 2011 ICP for improving the comparability across regions.

1.0 Treatment of Comparative Resistant Services in the 2005 ICP

In preparation for the 2005 ICP the EU-OECD countries had already had lengthy discussions on how to make comparisons for non-priced services. These discussions were informed by earlier experience in linking associated countries into the EU-OECD comparisons. This experience is summarized in Serguev (1998) and will not be separately treated here. In the end the EU-OECD countries agreed to base PPPs for comparison resistant services in 2005 on salary comparisons, stratified by type of occupation and skill. As the Commonwealth of Independent States (CIS) were linked to the EU-OECD region through Russia, this meant that 50 of the countries were already committed to basing PPPs on salary comparisons. Given this constraint on the global comparison it was decided initially to follow the same procedure in the other regions.

1.a The Final 2005 Approach

The schedule for the South American region was ahead of other regions and they followed the method used in the EU-OECD. However, as the salary information was reviewed in the Asian Region, the initial results, even after extensive validation of the data, did not appear to yield plausible value measures. In particular the indirect quantities for poor countries like Laos for headings like collective government, greatly exceed those in Hong Kong or Singapore. Similarly, in Western Asia the initial results produced volumes of comparison resistant services for Yemen that were much larger than those of its oil rich neighbors.

The Asian Region was the first region to introduce productivity adjustments for the headings of publicly provided education, health and collective government. The method is described in Appendix D of the final report of 2005 ICP (World Bank, 2008).¹ The approach uses a Cobb-Douglas production function as a method to adjust for productivity differences across countries in a similar manner to the 1975 ICP for collective government, education and health services as described in Section 2 below.²

¹ In its Preliminary Report the Asian Development Bank (July, 2007, Appendix D) the method is also described and some quantitative illustrations of the effects of the adjustments are provided. The Final Report on the ADB website provides an abbreviated description.

² In anticipating some discussion in Section 2, the 1975 ICP used a small sample of countries with estimates of health capital to obtain adjustment factors, while the 2005 ICP used estimates of capital per worker for the whole economy. Another difference is that ICP 2005 applied the adjustment factor to salaries and then used these to obtain indirect quantities. ICP 1975 adjusted quantities to obtain indirect PPPs.

Following the notation of the final report (World Bank, 2008, Appendix D) output (Y) is a function of labor (L) and capital (K), with labor and capital coefficients being α and $(1-\alpha)$, the constant returns to scale production function is given in (3) below.

$$(3) Y = cL^\alpha K^{1-\alpha}$$

The productivity adjustment involves estimation of the relationship of output per worker to accompanying inputs, in this case labor. (3) may be rewritten as in

$$(4) Y / L = c(K / L)^{1-\alpha}$$

We do not know α but the 2005 ICP but appendix D says,

“Similarly empirical studies have found values of α to be in the range of 0.5 to 0.7 for low- to high-income economies, respectively.” (World Bank, 2008, p. 180)

Assuming these values hold across the 2005 ICP countries then grouping countries into low, medium and high-income countries combined with their capital output levels allows estimation of their labor productivity.

Because data were not initially requested from countries there was no information on capital stock or labor specific to education, health or collective government to fall back on. It was heroically assumed that L was the population aged 15-64, which implies that labor force participation rates for that age cohort are the same across the wide range of 2005 countries. The next data problem was K , the capital stock information; it was available for some countries in Asia, but only a few in Africa and Western Asia. Again a very strong assumption had to be made, namely that for the few countries available, the range of capital output ratios of 2.5 to 3.5 applied to all countries. These values rose from low to high-income countries so values of 2.5, 3.0 and 3.5 were assigned to low, medium and high-income countries.

The next problem was to group countries within Asia, West Asia and Africa into the three income groups. Initial groupings were proposed in Regional meetings and typically agreement was reached without change. This meant each country had an assumed capital output ratio and value of $(1-\alpha)$, the inputs necessary to estimate (4). In applying the assumed values the country specific value of c , obtained from an initial EKS GDP (Ikle for Africa) aggregation, was substituted into (4) above for each country in the

regional comparison.³ This method provides an adjusted estimate of output per worker that might lead to moving countries into a different income group, in which case another iteration was performed. A non-iterative illustration is provided in Table 1, which includes 7 countries in Africa, Asia and West Asia where a productivity adjustment was implemented. It also includes Russia, which took part in both the OECD and CIS regional comparisons and the Kyrgyz Republic, which only took part in the CIS comparison. Along with Bolivia and Chile, these 4 countries made no adjustment of salaries when estimating the PPP for staff in converting collective government, health and education expenditures to volumes of output.

The countries in Table 1 have been assigned per capita GDP groups in column 1. The per worker GDP from 2005 is given in column 2, which uses the ILO labor force measure as the denominator while column 3 uses the proportion of population age 15-64. The ILO measure is a census, not employment number, but it does illustrate differences in reported participation rates among countries. The difference between columns 2 and 3 is low in Laos and very high in Yemen where few women are reported in the labor force. Applying the 'age 15-64' labor definition used in the 2005 ICP for productivity adjustments, GDP per worker level with HK as 100 is given in column 4. Hong Kong was taken as the high-income country in Asia for which no productivity adjustment was made.

Using equation (4), the GDP per worker figure in column 4 adjusted for the capital per worker group and associated labor coefficient is given in column 5 with HK = 100 for the countries for which no adjustment was made. For the remaining 7 countries, the values of column 4 have been repeated in column 5. (should the numbers be the same?)As can be seen for the 4 countries, the adjustment for capital itself has a relatively small effect (from 6.6 to 8.8%) on labor productivity. The larger effect for low and medium income countries is between countries with and without a productivity adjustment. This effect is also clear in column 5. For Nepal, Yemen, Kyrgyz, India and Laos, the volume of the 3 compensation headings would be under 10% of what they would have been with no adjustment; for Bolivia and China between 10-15%, and for Chile, Russia, and South Africa between 25% and 40%.

³ In this chapter we use EKS for the aggregation method used in the 2005 ICP Report interchangeably with GEKS, an increasingly used term, acknowledgement that the method originated by Gini.

1.b Effects of Productivity Adjustment on the Global Comparison

These are clearly large adjustments. It should be made clear that one cannot infer from column 5 the effect on GDP per capita of the countries. This depends on their actual salary PPPs and expenditure shares on the compensation headings. And to gauge the impact on Global Comparisons, the exact aggregation procedure makes a difference within and between regions.⁴ However, some insight may be gained from looking at the ICP 2005 Report, in particular Tables 3 through 6 that provide 19 sub-aggregates of exchange rate and PPP converted per capita and total expenditures. From these sub-aggregates, three aggregates of PPP converted and one aggregate of exchange rate converted expenditures were calculated:

1. HECG1 = Education, Health and Collective Government Expenditures;
2. HECG2 = 1 above from national accounts (the same as exchange rate converted) expenditures.
3. PPP converted Domestic Absorption ($DA1 = GDP - Exports + Imports = Actual Consumption + Gross Domestic Fixed Capital Formation + Collective Government$); and
4. $DA2 = \text{Sum of 13 (12?) headings of Actual Consumption} + \text{Collective Government} + \text{Gross Fixed Capital Formation (Sum of 4 components)}$

The HECG1 and HECG2 are shares of comparative resistance services to domestic absorption using PPP conversions from the 2005 ICP and shares at national prices. Table 2 provides a comparison of the ratios of PPP converted shares to shares at national prices. Countries are grouped by the size of the ratio and region. The average ratio is quite large, with the share at PPPs being 2.33 times larger than national shares that are in turn 19.2 % of GDP across the 146 countries. As discussed below, this somewhat overstates the effect of the PPPs used to convert comparison resistant services. But first, the striking differences across the regions in Table 2 should be noted. The CIS and Asian regions have average ratios well over 3 with Africa also high. West Asia and South America are similar and still near 2, while half the OECD countries, largely high income, are below 1.5 with many less than 1.1. There is a noticeable split in Western Asia with 4 of the Gulf countries being comparable to their well off OECD counterparts.

The reason these ratios can be misleadingly high is that the denominator shares in national currencies must sum to 1.0, but not so the numerator shares. The ratio of DA2 to

⁴ For example, the effect within a region would depend on whether EKS or Ikle (harmonically share weighted G-K) was used. In the global comparison it would depend on whether EKS was performed on the regions, or the individual countries.

DA1 in Table 3 illustrates the non-additive character of indexes like EKS. As can be seen the average of these ratios tends to be greater than 1.0 in the 2005 ICP with CIS and Asia being greater than 1.2, the OECD less than 1.1 and the remaining regions between 1.1 and 1.2. But 25 countries have ratios greater than 1.25 and the ratio can approach 2.0 as is the case of Mozambique, with 5 countries exceeding 1.5. The country information underlying Tables 2 and 3 are provided in Table 4. The ratios greater than 1.0 in Table 3 are the effect of comparison resistant services shown in Table 2. The correlation, R^2 is .48 and the slope is .10 over the 146 countries in Table 4.

The main conclusion from Tables 2 and 3 is that the estimates for comparison-resistant services in 2005 produce large increases in PPP converted shares for low income countries even in regions where productivity adjustments were made, like Asia. In regions with lower income countries but where no productivity adjustments were made, like CIS, it appears that there were even larger increases in shares of PPP converted comparison resistant services compared to shares at national prices. For these reasons, efforts are underway to improve the methods used in the 2011 ICP, and the general outline of an improved approach to treatment of education has been recommended for all regions, though the final form of this method is not yet available. We turn now to the question of whether there is any guidance for the 2011 round that can be gained from experiences in rounds of the ICP before 2005.

2.0 Experiments with Priced and Non-priced Services

The report of the 3rd round of the ICP for 1975 by Kravis, Heston and Summers (1982), referred to as (KHS), presented a comparison of the PPPs of priced commodities, priced services like shoe repair, haircuts and dry cleaning with non-priced services based on the 16 countries in the 1970 ICP, phase II. Of the 151 basic headings of expenditure 10 were comparison-resistant services and 21 were priced services comprising respectively 11 and 16 % of total expenditures on GDP. The relationships between the price level of priced services (P_p) and market priced commodities (P_c) and per capita GDP (all indexed to US=100) in logs from (KHS, p.137) are given below with standard errors of the coefficients in parentheses.

$$(1) \ln P_p = 2.351 (.321) + .414 (.090) \ln PCGDP [R^2 = .57 \text{ RMSE } .322 \text{ n}=16]$$

$$(2) \ln P_c = 3.317 (.128) + .266 (.036) \ln PCGDP [R^2 = .78 \text{ RMSE } .129 \text{ n}=16]$$

To illustrate the meaning of equations (1) and (2) consider country A with a per capita GDP 90% of the US and country B with a per capita GDP 5% of the US. The Price level of priced services for country A will be 3.3 times that of country while for

commodities the multiple between country A and B is only 2.2. It follows that the ratio of PPPs of priced services to PPPs of commodities rises with per capita GDP. One explanation of this relationship is the Balassa-Samuelson effect. It is based on the non-tradability of most priced and non-priced services and the assumption that productivity in services grows more slowly than commodities.⁵ Labor is usually mobile within countries so that wage rates for similar skills will be comparable whether in the production of traded commodities or non-traded services.⁶

But it also turns out that the PPP for non-priced services rises more rapidly with per capita GDP than priced services so that the ratio of price levels for comparison resistant to priced services also rises with GDP. The average ratio of prices for comparison-resistant to market services for Kenya and India, the lowest of the 16 countries in 1970, is 32.5 and for France and Germany, the countries just below the US (at 100 by definition) it is 86.5. As will be discussed below, the results of the 2005 for 146 countries displays a similar relationship

How does one explain this relationship? Is there less physical capital per worker in the education, health and collective services than in other services? Does physical capital per worker rise more rapidly in these non-priced sectors than in priced services? Or as is more likely, are both effects evident? And is labor monitoring less in low wage countries compared to high wage countries? If so it would lead to larger productivity differences between priced and non-priced services in low wage countries? We do not have the data to answer these questions. However, since the 2005 ICP exhibits a similar relationship it is worth noting the adjustments considered for the 1975 ICP.

In KHS some illustrative and informative calculations were made that take into account the relative importance of the comparison resistant services in expenditures. In place of equal productivity PPPs based on salary comparisons, the PPPs for priced services and the PPPs for all other basic headings were used in separate G-K aggregations. The effects on per capita income for the averages of the 1st and 2nd quartiles of the 16 countries for 1970 are as follows. If the published GDP per capita of the first quartile average (Kenya, India, the Phillipines and Korea) is 100, then using the PPPs of priced services the index is 87. That is, if staff in education, health and collective government were only as productive as labor in priced services in their countries, the per capita GDP of the lowest quartile would be only 87% compared to using the equal

⁵ The Balassa-Samuelson effect is based on the fact that commodities are largely tradable whereas services, especially comparison resistant services, are largely non-tradable. See (Heston, Summers and Nuxoll, 1994), where this is tested and where alternative explanations are also cited.

⁶ The tradable/non-tradable distinction is not as tidy as we make it appear. Consumers are tourists and they also travel to get health care in other countries. Construction companies put up projects outside their home countries to cite some examples of how loose can be the distinction.

productivity assumption. For the next poorest quartile (Columbia, Malaysia, Iran and Hungary) the comparable result was 95%, or a 5% reduction in per capita GDP.

There are some constraints on labor mobility, including premiums in larger organizations, but in general workers can move between commodity and service sectors so that wages for similar skills tend to receive similar pay. This makes it plausible to believe that market-priced service PPPs might be a reasonable approximation of the PPP for non-priced services. KHS (1982, p.140) conclude,

“...results for middle and high-income countries are insensitive to the treatment of comparison resistant services. More is at stake for the low-income countries.”

As a result of these findings, several modifications of the equal productivity assumption were introduced in the 1975 ICP for the different headings of comparison-resistant services, which are discussed in Section 3.⁷

How does this result hold up in for the 2005 ICP? First these results cannot be compared directly to the publication because the author is working with a set of 129 basic headings of expenditures and PPPs made available to researchers, a data set that has less detail than used in the ICP 2005 report. Further these PPPs already embody the productivity adjustment used for Africa, Asia and Western Asia so the comparison is not as clear as for the earlier 1970 ICP. The comparisons reported in Table 5 only use 2 EKS aggregations over the 128 basic headings of domestic absorption that is GDP less the net foreign balance and net expenditures of residents abroad. Domestic absorption (DA) is used a) because the net foreign balance and net expenditures of residents abroad has never been handled satisfactorily in history of the ICP; and because DA is a meaningful measure of the bundle of goods and services that countries distribute among various heads of expenditure.⁸

Table 5 is based on 2 EKS aggregations of the 2005 ICP basic headings, the first is similar to the final report except that the aggregation is over all countries and fixity is

⁷ In the 1990s when the OECD began to make comparisons for non-member countries especially in Eastern Europe and the former Soviet Union some experiments were made using priced services PPPs for comparison resistant-services. However, as has been the case in many rounds of the ICP, the number of observations on priced services was not considered adequate.

⁸ To properly handle foreign trade, export and import PPPs should be separately estimated. Because this involves additional price comparisons, typically the ICP has converted the net foreign balance by a common conversion factor, the exchange rate in the 2005 ICP and some earlier ICPs, the PPP for domestic absorption. Use of the exchange rate is not base country invariant and has an asymmetric effect for countries with price levels below the exchange rate as to whether they are running a trade deficit or surplus.

not observed.⁹ The second uses the same aggregation but substitutes the PPPs for priced services for each country for compensation payments for collective government, and publicly provided education and health in place of the PPPs used in the 2005 ICP. Table 5 summarizes the results by region.

Table 5 is a distribution of countries within each region by the ratio of the PPP over DA of the second (using priced services) to first aggregation. If the ratio for a country is greater than 1.0, it means that their corresponding estimate of DA per capita relative to the US will be smaller using the PPP for priced services. As can be seen the use of priced services PPPs would produce smaller estimates of per capita DA for most countries, the average being 17% for all 146. The pattern by regions is plausible with the difference being smallest in the OECD countries, and highest in the CIS countries where no productivity adjustment was made and in Africa, where the base comparison already includes a productivity adjustment. In general, the pattern in Table 5 is consistent with the KHS results of a generation earlier where the biggest effect of this substitution was on lower income countries and regions.

One conclusion of this paper is that the use of the PPP for priced services is a possible fallback should there be problems with other methods. It has the advantage of being simple and transparent to implement. It assumes that the relationship between productivity in collective government and in publicly provided education and health services is similar across countries. However, some such assumption is involved in most approaches, so perhaps this is not a major limitation. It also depends on good estimates of the PPPs for priced services, which are difficult to specify leading to less than robust results. In fact it was the inability of countries to collect enough priced service items across countries that led KHS to use alternative methods in the 1975 ICP.

When the OECD considered using priced services in the late 1980s the proposal was abandoned because of data limitations.

Is the coverage of priced services is adequate in the 2005 ICP to justify a recommendation for use of the PPP for priced services as an alternative treatment of comparison resistant services? More research is required to answer this question. Such future research might include; a) the sensitivity of the results to alternative bundles of priced services; b) the implicit volume comparisons for these services relative to alternative estimates; and c) whether the use of priced services might be combined with other methods.

⁹ Fixity is preserving the relationships of countries within a region to those obtained in the regional comparison. Even when fixity is observed, an aggregation across the 5 ICP regions (CIS is included with OECD) produces different results than an aggregation of the 146 individual countries.

3.0 Other Methods of Dealing with Non-priced Services up to the 1975 Round

Having examined the approach in the 2005 ICP to comparison resistant services and the possible use of the PPPs of priced services, we consider some other approaches examined in previous ICPs. This discussion is not exhaustive because of space constraints and because Sergueev (1998) has provided an excellent summary of methods of the early rounds of the ICP and especially those developed from 1975 to 1998 for comparing non-priced expenditure headings for OECD ICP rounds. Some of these methods were only used to link countries in Eastern and Western Europe for which Sergueev provides interesting illustrations. Here the focus will be on some of the methods considered in KHS (1982).

3.a Health Personnel

Because there had been criticism of the treatment of services in the 1970 ICP, much more data was requested for the 1975 ICP in the health field including numbers of medical personnel by type, stock of health capital and a larger number of directly priced physicians services. A number of studies were consulted and a compilation was made of the passing rate of medical students on a standard test in the 32 of 34 ICP countries in 1975 that had medical schools. (KHS,1982, p. 152). The interested reader can explore the many checks that were done on the method eventually adopted. For present purposes three main points of the treatment of health are especially relevant for future ICP considerations.

First, it is extremely useful to collect as much quantity, quality and direct price and wage information as possible. To illustrate, the 1970 ICP only used direct quantity comparisons for medical personnel, which is analogous to 2005 in that equal productivity of personnel was assumed. Initially, the 2005 ICP also was going to assume equal productivity, but the quantities were to be indirectly derived by dividing expenditures on personnel by PPPs estimated from salary comparisons. Certainly one would like to obtain comparable results by these two approaches since both make the same assumptions about productivity. However, this is only possible if both salary and direct quantity information is collected. This same principle would apply to other expenditure groupings including education, collective government, construction, and dwelling services.

Second, in the health field there are many directly priced medical services purchased by households used in the ICP, such as physician visits, Xrays or lab tests. These can be directly compared to the PPPs from salary comparisons of hospital personnel or the PPPs derived from expenditures and quantities of medical personnel.

Further, the 1975 ICP also used direct quantities consumed by households like hospital bed days and measures like physicians or nurses per person, which are slightly closer to output. Ideally, we would like measures of output of medical services and more research is taking place, but it is unlikely to be operational for the 2011 ICP comparisons.

Third, the use of crude production functions to better approximate medical output requires some heroic assumptions, but even Cobb-Douglas functions provide a useful framework for comparisons. The attempt to collect information on the stock of medical capital was not a great success in that only 8 of the 34 countries could provide estimates. However the range of countries included low, middle and high-income countries. Health capital was broken into hospitals and all the rest. These observations were stylized into adjustment factors for health personnel to approximate equivalence in productivity of dentists, nurses and physicians in the 1975 ICP. For countries below 30% of US per capita GDP the divisor for personnel was 1.30, and for countries between 30 and 50%, it was 1.15 (KHS, 1982,143).

The 1975 approach was similar to that used for some regions in the 2005 ICP except that the latter used capital per worker for the whole economy across all publicly provided comparison resistant services including hospitals. In KHS hospital capital and hospital bed days were used in a production function producing similar adjustment factors to those for personnel, and so the same adjustment was done to hospital bed days as described above. In addition the adjustments covered all medical personnel whether or not they were employed in hospitals or were providing market-priced services. One important conclusion of the 1975 ICP treatment of health was that the direct price measures of the PPPs of health services, like doctor visits, rise with per capita income much faster than the PPPs measured indirectly by dividing expenditures by labor input, about which more below.

3.b Education

Capital stock measures for education were harder to obtain in the 1975 ICP, only 3 countries supplying data. Further the role of physical capital in producing educational services is more ambiguous than in health services. For these reasons no adjustment for capital was attempted in KHS (1982). The several cross-country studies based on common tests in local languages circa 1970 revealed that,

“...the home variables were relatively less important and the school inputs (relating to inputs such as teachers and library facilities) more important in explaining school

performance in developing countries than in developed countries.” (KHS, 1982, p. 156)

However, the state of the art of cross-country testing methodology did not appear robust enough in 1975 to be used as a tool in making volume comparisons across countries.

As with health, both direct quantity and price data were collected including number of students and teachers at the primary and secondary levels. In addition salary data for teachers for 5 levels of teacher education were collected; a CPD was run providing estimates of relative salaries across ICP countries adjusted for relative education level of teachers. These direct PPPs based on salaries were then compared with indirect measures derived from dividing expenditures by teacher quantities. As in the case of medical personnel there is a systematic relationship between direct and indirect PPPs with per capita income. The latter rise faster than the former so the ratio of direct to indirect PPPs declines as per capita GDPs rise. Based on this relationship a downward adjustment was made to the direct quantity estimates that decreased from 80% for the lowest of 4 per capita income groups, 32% for the 2nd and 12 % for the 3rd groups all relative to the top group of per capita income. (KHS, p. 159)

These adjusted teacher quantities represent the inputs to education and for want of any other measure would also represent the output. Another output measure is the number of students taught. There are many problems with the pupil measure too. What is the relationship of class size to educational output? Pupils registered, the most typical number available, is not necessarily the number who are attending and eventually graduating. In the end the square root of the product of the adjusted quantity of teachers and the number of pupils was taken as the output of primary and secondary education. No adjustment was made for 3rd level education. Also, it may be noted that in the 1975 ICP, administrative expenses on education were excluded from the total education expenditures when dividing by the above quantity to derive indirect PPPs. Finally, the method to be adopted by the OECD countries in their 2011 comparisons uses test information to make an adjustment (in fact, small) to the measure of output, the number of primary and secondary students.

This method for the OECD countries for 2011 may prove more difficult when extended to the 180 countries in 2011. While this is along the lines of what was done in the ICP prior to 2005, the experience of the KHS was that information on salaries of teachers and their average years of training would also be useful. There is every reason to believe that quantity measures of students in primary and secondary education may

overstate the output of the education sector for low-income countries, with the degree of overstatement declining with per capita income.

3.c. Collective Government Expenditures on Personnel

As in the 1970 ICP a questionnaire was distributed to countries to obtain information on capital in government. However, the response in 1975 was not sufficient to form a basis for any adjustment. The eventual choice was to use the same adjustment factors that were used for health personnel, namely a 30% adjustment for the lowest group of countries and 15% for the next group with per capita GDPs between 30 and 50% of the US.

The principal conclusion to emerge from the review in this section is summarized in Figure 1 (Figure 5-1 of KHS). On the horizontal axis we plot the per capita GDP of a country relative to the United States, labeled r and the vertical axis the price level of comparison resistant services. The direct price level measure at lower incomes begins at a higher point than the price level indirectly measured by dividing expenditures by quantities. Or a price level estimated from salary comparisons. However the direct measure rises more slowly with incomes so at moderately high incomes both measures yield comparable estimates of the price level. This appears to be a pattern in other difficult to measure sectors like imputed rents for homeowners. When rents are directly measured holding constant amenities of housing they are higher than when measured by the quantity approach where only a few elements of quality are captured. If as KHS conclude, Figure 1 captures a general phenomenon, then it provides strong support for collecting as much information as available for hard to compare expenditure headings.

4.0 Linking Issues for Regions in 2005

Section1 covered the ICP 2005 procedures within each region but not the linking of the different regions. The problem first arose in the context of whether all expenditure groups should be linked through the 18 Ring countries that had agreed to undertake pricing of a large number of shop items. Since common questionnaires were used for areas like housing and government salaries, would it be more appropriate to use data from all countries, or just the Ring countries for linking? Because Ring countries did not appear likely to supply better quality data for housing or salaries in some regions, it was agreed to use data for all countries in the Global Office. The OECD was one exception to this pattern because this decision was taken after the Ring country protocol was adopted. And South America was an exception because for budgetary reasons they were the first

region to complete their work and their salary classification was not comparable, so only the Ring countries, Brazil and Chile were used for linking.

4.a Estimates of Overlapping Price Levels

In the data available at the Global Office, there were 75 countries with wages, with five Eurostat countries not being identified by name because release of data for the Ring OECD data was all that was approved in the previously mentioned protocol. For each country there were data on the job, the occupational grouping (administration, defense, education and health). In addition a simple grouping by skill designation assigned to each job, and of course countries were assigned to their region. Early estimates suggested that due to the spotty and variable character of the salary observations, it was best to omit defense from the estimation, reducing the number of salary observations from 1876 to 1730. Several of these variables are collinear so for example a CPD cannot be estimated for Region and individual countries, for job and occupational grouping or skill level. Table 6 provides results of a CPD regression where only regional and sectoral coefficients are estimated, but the latter are not shown.¹⁰ Since country variation is not taken into account in Table 6 the explained variance is fairly low. All of the coefficients are significant and the rank of the regions is plausible. Column 2 shows the anti-log of the regional coefficients in Table 6 as a percent of the OECD countries. Roughly speaking salaries for comparable occupations are 14 times higher in the OECD countries than in Africa and at least 5 times higher in the other regions. If expenditures were converted by these PPPs they would greatly raise the estimated quantities of these comparison resistant services. Again, there are no CIS countries in the data set, though wages in several of these countries are quite low compared to EU countries.

When the equation allows for individual country coefficients instead of regions, the explained variance is of course much higher ($R^2 = .97$). With the UK, the base (The 9 OECD countries in the regression average 98% of the UK) Bolivia is only 5% of the UK and many countries in Africa are even lower. Table 6 reinforces the need to follow up the initial steps taken in the 2005 ICP to make some adjustment for productivity.

¹⁰ The ranking of types of employment is not particularly meaningful, since Administration, for example, may have mainly lower level employees, while Education has mainly higher levels in the particular mix occupations sampled within the broad groupings of Administration, Health and Education.

5.0 A Proposal for Linking in 2011 and Conclusions

The linking issue was important for comparison resistant services in ICP 2005 because the productivity adjustment was made in 3 of the 6 regions, leading to likely non-comparability. For a country like Bolivia or Tajikistan, where no productivity adjustment was made, per capita incomes will be less comparable with similar countries in Africa, Asia or Western Asia. The issue is further complicated because the actual aggregation to the global level in 2005 was not done using the 146 countries individually but rather it was done using 5 regional aggregates, the CIS included with OECD. It was not possible in the end to obtain satisfactory regional adjustment factors to improve on a simple Ring country comparison.¹¹ As a consequence, there remain a lack of comparability of low and middle-income countries across different regions in the final ICP report. A somewhat different approach is explored here that uses as much salary information as available to the global office and make a rough allowance for productivity differentials across the regions based on the human and physical capital stock.

5.a Estimates of Productivity by Country

In this concluding section we illustrate with the 2005 data an alternative method of linking based on individual country productivity adjustments using a Cobb-Douglas production function across all the regions. Physical capital, K_p , and labor force estimates, L , are from the Penn World Table and human capital estimates, K_h , from the Barro-Lee Data set.¹² The education measure is average years of primary and secondary school completed. In the equations below the dependent variable is the log of GDP/L and in the first equation only K_p/L is on the right side, while K_h is also included in the second equation.

$$\ln GDP/L = 4.56 + .704 \cdot \ln K_p/L \quad R^2 = .799 \quad n = 105$$

$$\ln GDP/L = 4.87 + .552 \cdot \ln K_p/L + .455 \cdot \ln K_h \quad R^2 = .786 \quad n = 87$$

All the coefficients were significant but for K_h only at the 5% level. The first equation covers more countries, which makes it preferable for the illustration below. In

¹¹ It was possible to estimate from a CPD equation coefficients for the regions. However, these estimates were based on wage comparisons for different occupations much like the regional comparisons and simply reproduced what had been found within regions. That is based on wage and salary comparisons, indirect quantity estimates for low-income regions would be implausibly large.

¹² Sources: pwt.econ.upenn.edu has the labor force data. The physical capital series is preliminary and not yet on the web. In addition country studies of China, India and Indonesia were used to supplement the PWT estimates. The citation for Barro-Lee is: www.cid.harvard.edu/ciddata/ciddata.html. As this Chapter was being written the Barro-Lee data set was updated and it would be useful to see if the new data has broader coverage and more explanatory power. The India and Indonesia estimates are based on Pierre van den Eng (2008) and the China estimate is based on Yanrui Wu (2009).

addition, when the ratio of the estimated to actual value is taken, the standard deviation is significantly larger for the equation including K_h (.0827 versus .065). We conclude that the specification or data problems associated with including human capital in the productivity equation need research beyond this paper to justify inclusion.

5.b Productivity Adjustment

The direction of productivity adjustment across countries is clear. The method adopted in Asia is a method that certainly moves in the right direction in that most agree that output per government, health or education worker is likely to increase with more capital per worker. The approach suggested here is illustrated for the 2005 ICP research data set. The following estimates were made:

First, output per worker estimates were made based upon the above estimating equation for the 106 countries. These estimates were extended to the remaining countries based upon the relationship of the estimated ln of output per worker for the 106 countries to per capita GDP from an initial GEKS estimate. This estimating equation had an R^2 of .465 with sensible coefficients, though if this method were adapted in the future one would want a tighter estimating equation. Applying the equation coefficients to the per capita GDPs of the remaining 40 countries yielded estimates of the ln of output per worker. Another limitation on this illustration of course is that these estimates refer to the whole economy, not labor in comparison resistant services.

Second, these output per worker estimates were used to obtain adjustment factors that ranged from 1 to 4 going from high to low income countries. The admittedly very rough groupings were as follows. Output per worker > \$40,000 were assign 1.0, \$30-40,000 1.2, \$20-30,000 1.5, \$15-20,000 2.0, \$10-15,000 2.5, \$5-10,000 3.3 and under \$5,000 4. These factors were used to adjust the parities obtained from wage comparisons to obtain new input parities for the compensation headings of comparison resistant services.¹³

The third step taken is run aggregations using unadjusted and adjusted compensation figures to gauge the impact on the results. While this might eventually prove an effective adjustment for individual countries, that is not the purpose of the

¹³ For the OECD, CIS and South American regions, the basic heading compensation parities were from the research input of the 2005 ICP. For Africa, Asia-Pacific and Western Asia, where the research input compensation parities have already been adjusted, the wage estimating equations discussed in the text were used to derive parities based upon the UK. These could in turn be put on a US base. Some countries in these three regions were not in the wage equation, and so had to be estimated by a short-cut equation based upon per capita income, a method that should be improved.

present exercise. It is already agreed that the OECD and CIS countries will not make any adjustment within their regions for the 2011 ICP. One reason for this are that these regions are already moving to quantity comparisons for educational services and exploring possibilities for the health sector.

Rather the purpose of this exercise is to propose a method to allow regions to retain whatever method they choose for the regional exercise while at the same time linking the regions in a way to permit the adjustment across regions to improve comparability. The results of this exercise are reported in Table 7. Two aggregation methods have been used, both producing similar results, one GEKS and the other a weighted CPD, labeled CPDW. The share of each region of global Domestic Absorption is calculated with and without adjustment for each of the two methods. These shares simply add up the total GDP of each country in a region and divide by the global total. The total DA is also provided for each method. As can be seen, without the productivity adjustment the world total is larger than with adjustment because less output of these services is attributed to many of the lower income countries.

But more importantly the shares of each region are also substantially affected. The OECD share goes up because few of its countries are affected by the adjustment while the shares of all the other regions go down, with Africa by more than 5%. The effect is larger for GEKS than for CPDW. These effects are similar to what would result if one used the PPP for non-priced services for the compensation headings of comparative resistant services. That is, the regional pattern of Table 7 is similar to that of the priced service exercise summarized in Table 5 where the effect would be most for Africa and less for South America and West Asia. However, our judgment is that the variation in the PPP for non-priced services is subject to more error than even the rough productivity adjustment of Table 7.

How would this proposal improve the global comparisons? The adjusted shares provide a regional total that can be distributed among the countries in a region in the same way as the regional comparison. This has a cost in that the comparability across basic heading parities is reduced. However, since greater interest centers on aggregate comparisons across countries the type of adjustment in Table 8 has a major advantage. Because the effects of the productivity adjustment are systematic across the regions, attributing adjusted shares to each region will improve comparability of total DA and GDP per capita across the regions in comparison to doing an adjustment in some regions but not others.

Summing up we propose this as a general method that is likely to be useful for improving the comparability of comparison resistant service volumes across countries

when regions adopt different methods as appears likely in 2011 for collective government, health, and perhaps education and dwelling services. The case of dwelling services is explored in more detail in Chapter 19. The other principal conclusion of this chapter relevant to the 2011 ICP is implicit in Figure 1. Namely indirectly estimated price levels tend to be systematically lower compared to direct price comparisons for low-income countries. Given the heterogeneity of available national statistical bases there is no simple solution to this problem, but there is a simple message. In much the same way that expenditure surveys and commodity flow tables are used as checks on each other in national accounts, collecting more information on the headache headings in the ICP can provide similarly useful checks on PPP and volume estimates.

DRAFT

References

Dean, Edwin, Using Productivity Data For International Comparisons of Output in Nonmarket Sectors, June, 1909

Heston, Alan, Robert Summers and Daniel Nuxoll (1994) "The Differential Productivity Hypothesis and Purchasing Power Parities: Some New Evidence," *Review of International Economics*, October, 1994.

Kravis, Irving, Alan Heston, and Robert Summers (1982), *World Product and Income*, World Bank and Johns Hopkins Press, Baltimore.

Sergueev, Sergey (1998) "Comparison of Non-Market Services at Cross Roads (experience, considerations, proposals)," Item 8, Statistical Commission and Economic Commission for Europe Conference of European Statisticians, Consultation. 3 - 5 June 1998, Vienna.
http://siteresources.worldbank.org/ICPINT/Resources/SS_NMS_ECPEExperience.DOC

Wu, Yanrui (2009), "China's Capital Stock by Sector and by Region," Discussion Paper 09.02, Business School, University of Western Australia.