

**Economic and Social Council**Distr.: General  
18 December 2014

Original: English

---

**Statistical Commission****Forty-sixth session**

3-6 March 2015

Item 3 (i) of the provisional agenda\*

**Items for discussion and decision: International  
Comparison Programme****Report of the Friends of the Chair group on the evaluation of  
the 2011 round of the International Comparison Programme****Note by the Secretary-General**

The present document is the first part of a report prepared by the Friends of the Chair group on the evaluation of the 2011 round of the International Comparison Programme (ICP) assessing the scope, activities and lessons learned. The mandate of the Friends of the Chair group received from the Statistical Commission in 2014 is described in section I. A description of the background and organizational setting against which the 2011 round was conducted is provided in section II. Section III covers governance issues as well as technical and methodological aspects. The procedure for the completion of part II of the report is set out in section IV, and the recommendations are provided in section V. Points for discussion are provided in section VI. Part II of the evaluation report will be submitted to the Statistical Commission at its forty-seventh session, in 2016. It will provide analysis, conclusions and recommendations on the areas subject to further evaluation by the Friends of the Chair group, such as the frequency with which the ICP exercise would take place, a discussion of issues of methodology if a more frequent periodicity of ICP is decided by the Commission and an examination of relevant financing issues required for establishing a solid and sustainable international programme and the relationship of such programme to capacity-building efforts. The Statistical Commission is invited to discuss and approve the recommendations included in the present report containing part I of the evaluation.

---

\* [E/CN.3/2015/1](#).



## **Report of the Friends of the Chair group on the evaluation of the 2011 round of the International Comparison Programme**

### **I. Mandate given to the Friends of the Chair group by the Statistical Commission and working methods of the group**

1. The Statistical Commission, at its forty-fifth session, in March 2014, agreed to establish the Friends of the Chair group to carry out the evaluation of the 2011 round and prepare an initial report for review at the forty-sixth session of the Commission, in 2015. The Commission appreciated the interest of Austria and India in serving as co-chairs and that of other countries in participating in the group and requested the group to take into account regional experiences in its work, such as the recommendations made by the Advisory Group on Statistics. The Friends of the Chair group, consisting of representatives of participating countries, was established to represent the country scope and with the expectation of making a proposal on the desirability of a new round. Specifically, the Friends of the Chair should seek to deliver:

(a) An evaluation of the effectiveness of the global, regional and national levels of governance in the implementation and conduct of the programme;

(b) A review of the technical and methodological programme, including the ICP handbook and book and the ICP kit;

(c) An evaluation of the regional results.

2. In June 2014, the Friends of the Chair group received a draft framework for conducting the ICP review. The first meeting of the Co-chairs, Statistics Austria and the Ministry of Statistics and Programme Implementation of India, along with the secretariat of the Statistical Commission and the World Bank Development Data Group, was held on 17 June 2014 via virtual connection. The meeting concerned five topics:

(a) Introduction and meeting objectives;

(b) Composition of the group;

(c) Terms of reference of the group;

(d) Support role of the Development Data Group and the Statistics Division of the Department of Economic and Social Affairs;

(e) Timetable for the evaluation.

3. It was agreed that, as for the ICP 2005 evaluation, the methodology of the ICP 2011 evaluation should include a survey of the opinions of major stakeholders, the participating countries and the major users. The Development Data Group circulated the survey questionnaires used in the ICP 2005 evaluation as starting point for developing the questionnaires for the 2011 evaluation.

4. The Friends of the Chair group decided, upon request by the Commission at its forty-fifth session, to divide the evaluation report in two parts: part I would focus mainly on governance issues and methodological innovations and improvements. Part II, which will be submitted to the Statistical Commission in 2016, would cover the question of frequency of ICP rounds, methodological issues linked to a higher

frequency and financing issues. It would also include an overall assessment of ICP 2011 based on the results of a user survey and draw on the experience from countries participating in ICP 2011. Part II would conclude by providing recommendations on the future of ICP.

5. Part I of the evaluation report reflects the view from the perspective of the main stakeholders involved. It is based on reports from the organizations that are part of the governance structure, such as the Executive Board and the ICP Global Office in the World Bank, and on ICP 2011 reports and publications of the regional agencies. One major source of information are the answers to a questionnaire addressed to the main stakeholders in August 2014. Further information was gathered through interviews with selected major stakeholders.

6. Information on governance performance as experienced on the global and regional levels and a description of the technical and methodological innovations and improvements in ICP 2011 are provided in section III below. The further procedure leading to the final evaluation report of the Friends of the Chair group, to be submitted to the Statistical Commission in March 2016, is set out in section IV. Preliminary concluding remarks based on the evaluation exercise done thus far by the group are provided in section V. Points for discussion are provided in section VI.

## **II. Background and organizational setting**

7. ICP is a worldwide statistical exercise aimed at estimating purchasing power parities (PPPs), to be used as currency converters to compare the macroeconomic indicators and the economic situation of countries around the world. The primary purpose of ICP is to generate relevant PPP data to convert the gross domestic product (GDP) and its sub-aggregates, reported in national currencies, into a common currency that equalizes the purchasing power of the currencies. By using PPPs as conversion factors, the resulting comparisons enable the measurement of the relative social and economic well-being of countries, monitoring of the incidence of poverty, tracking of progress towards the Millennium Development Goals and the effective targeting of programmes effectively. PPPs also assist international markets by identifying the relative productivity and investment potential of different countries.

8. ICP is both technically and operationally complex. Its success is measured by the extent of improvement in the quality of the price and national accounts data. The basic principle is that all participants agree jointly on the methodology, workplan and timetable. The governance arrangements need to provide an environment conducive to the success of ICP, both in terms of providing technically sound data and as an international partnership with participants from national, regional and international agencies working together effectively.

9. ICP dates to the late 1960s, starting with a small project comprising 10 countries in 1970. Later ICP rounds were conducted in 1975, 1980, 1985, 1990 (partial) and 1993. Since 1990, the World Bank has had the role of global coordinator of ICP. The 2005 round of ICP marked a turning point in the programme, with the aim of meeting criticisms levelled at ICP and quality problems. The 2005 round was completed with the publication of the results based on price and other data collected for 146 countries. The 2005 round built on an extensive partnership between many international, regional and national agencies.

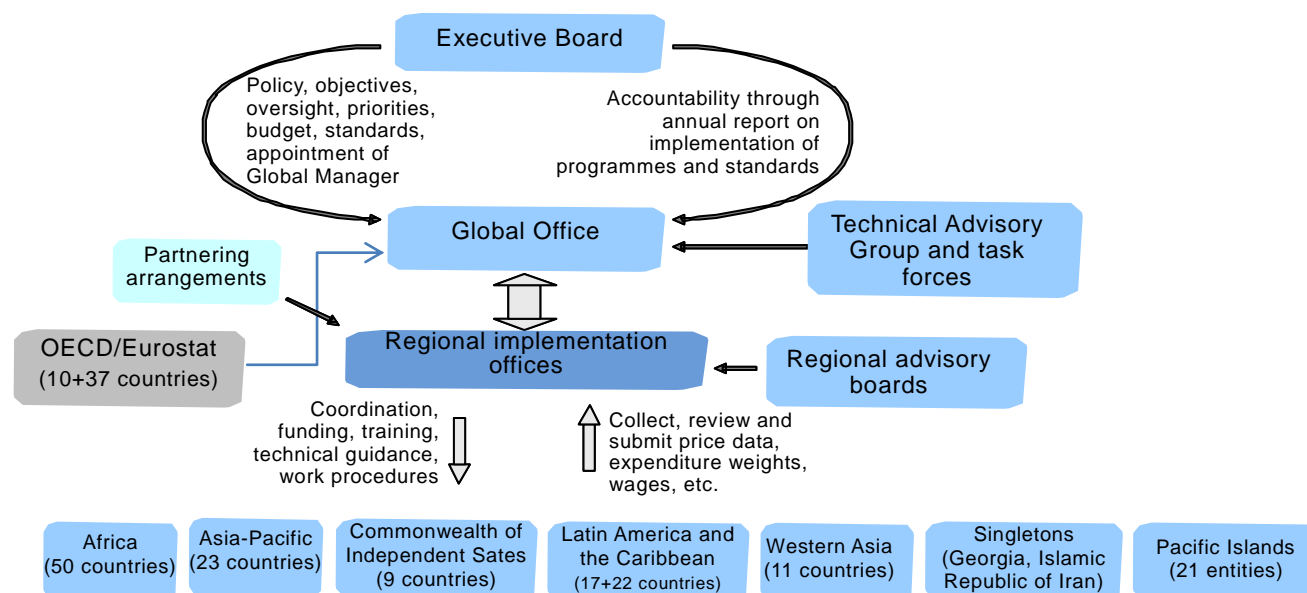
10. At its thirty-ninth session, in February 2008, the Statistical Commission welcomed the successful completion of the 2005 round and, in order to maintain the momentum, requested that preparations for the next ICP round begin immediately, with a target year for data collection of 2011 (see decision 39/103). Following the successful completion of the ICP 2005 round under World Bank management, the Statistical Commission requested the Bank to take on the global coordination of the 2011 round, which had to leverage the successful implementation of the 2005 round. This, together with a concerted effort by international and national statistical agencies, meant that the 2011 round should be better planned, managed and coordinated. The most important requirements for organizing the 2011 round were the development of a clear and broadly agreed governance structure and a clear work programme. The proposed governance structure built on: (a) the mandate provided by the Statistical Commission; (b) the most effective elements of the governance structure of the 2005 round; and (c) recommendations of the Friends of the Chair group in its ICP 2005 evaluation report. The work programme was discussed and reviewed by the ICP interim executive board meetings in June and October 2008 on the basis of the comments of its members. The ICP interim executive board presented a report to the Statistical Commission at its fortieth session, in February 2009, and the main elements of the 2011 ICP governance framework were endorsed by the Commission. This framework was later expanded to serve as the reference document for the institutional set-up of the Programme and the partnership arrangements between the Global Office on the one hand and the Organization for Economic Cooperation and Development (OECD) and Eurostat on the other, as well as the regional organizations that were expected to coordinate activities at the regional level.

11. Significant progress was made in preparing for the 2011 round between the fortieth and forty-first sessions of the Commission. During this period, the ICP governance structure was put in place: the global manager was hired in April 2009 and the new Global Office was established. The Executive Board, the Technical Advisory Group and the regional coordinating bodies were set up and held their first meetings in September and October 2009. The ICP Global Office, hosted by the World Bank, worked permanently to broaden the scope of the 2011 programme, streamline quality-assessment processes, improve the relevance of PPP statistics to poverty, ensure the sustainability of PPP delivery and enhance statistical capacity-building activities related to the generation of ICP basic data, with a specific focus on price statistics and the implementation of the System of National Accounts.

12. The figure below shows the overall governance structure of ICP 2011 and the partnering arrangements with the regional agencies and the OECD/Eurostat programme, the latter being a relatively autonomous programme. The Executive Board was the decision-making and strategic body. The Global Office, being the secretariat of the Executive Board, was responsible for the overall coordination of ICP. The Technical Advisory Group and task forces (the Computation Task Force, the Validation Expert Group and the Results Review Group) provided research and advice on technical issues and assisted in the computation and analysis of the results (in close communication with the Global Office). The regional offices served as regional coordinators, coordinating and supporting the national statistics institutes in implementing the comprehensive survey programmes. Regional advisory boards were established for several regions.

## Governance structure of the International Comparison Programme, 2011

Under the auspices of the Statistical Commission



13. One of the biggest challenges related to the scope of the 2011 programme was the accommodation of the drastic widening of coverage to as many as 199 countries or economies: 50 from Africa; 23 in Asia and the Pacific; 9 in the Commonwealth of Independent States (including the Russian Federation, which also participated in the OECD/Eurostat programme); 17 in Latin America; 22 in the Caribbean; 12 in Western Asia (including Egypt and the Sudan, which also participated in the Africa programme); 21 Pacific island countries and territories (including Fiji, which also participated in the Asia-Pacific programme); 47 in the OECD/Eurostat programme; and the economies of Georgia and the Islamic Republic of Iran.

14. The World Bank arranged collaborators in all ICP regions, and the ICP Global Office worked through those collaborators to undertake the regional price surveys and comparisons. In some cases the regional collaborator was a development bank (Asia and Africa) and in some cases a United Nations regional commission (Western Asia and the Middle East and Latin America and the Caribbean). In others, it was an international organization already running an ICP (OECD, Eurostat and the Commonwealth of Independent States). This federated approach worked as follows:

- For the African region, the coordination was carried out by a regional coordinator — the African Development Bank (AfDB), a price statistics team and a national accounts team.
- For the Asia-Pacific region, the regional coordinating agency was the Asian Development Bank (ADB), whose Development Indicators and Policy Research Division, assisted by an economics and statistics officer, was the regional coordinator. A regional advisory board (similar to the Executive Board of the Global Office) was formed, whose members were chosen as a mix of main stakeholders, regional agencies and selected national implementing agencies reflecting subregional representation.

- For the Commonwealth of Independent States region CIS-Stat acted as the regional coordinating agency and Rosstat as a partner organization.
- The Eurostat PPP exercise is governed by the PPP Working Group, which consists of delegates of all participating countries.
- For the non-European OECD countries, no governance structure as such was in place, but the OECD PPP unit coordinated the works and participated in all Eurostat PPP meetings to attempt to have common decisions on methodological points. In addition, OECD organized several PPP meetings for the non-European OECD countries.
- For the Latin America and Caribbean region, the governance was similar to that of the Global Office. A strategy of focal points of the Economic Commission for Latin America and the Caribbean (ECLAC), by country, was adopted, with each staff member in contact with focal points from the assigned countries. In some countries there were two national focal points, because the consumer price index belongs to the national statistical offices and the national accounts to the central bank.
- In the Western Asian region, ICP was headed by the Regional Coordinating Office at the Economic and Social Commission for Western Asia (ESCWA). The Office coordinated the activities of the participating countries, which in turn put in place national teams led by a national coordinator. In addition to its role of supervising the implementation of the programme, the Office also played a part in liaising between the Global Office on one side and the national teams on the other, and transmitted all news and updates. Moreover, a regional executive board was created to keep track of progress.
- The Australian Bureau of Statistics was the regional coordinator for the Pacific islands programme.

15. The results of ICP 2011 were presented at the forty-fifth session of the Statistical Commission, in March 2014. The report, prepared by the World Bank on behalf of the ICP Executive Board, described the computation of the ICP 2011 estimates and the governance activities that took place to support the computation process and to review the preliminary results. The report also referred to the schedule of publication of the final results and to the related communication strategy, as well as to the PPP statistics revision policy. The report also outlined the plans for the evaluation of the ICP 2011 round. While welcoming the World Bank report and the progress in the implementation of the 2011 round, the Statistical Commission supported the two-stage approach for releasing summary results along with key documentation by 30 April 2014, followed by a comprehensive report that would include the full set of results and more comprehensive documentation in the third quarter of 2014.

### **III. Major initial findings**

16. A sound picture of the ICP 2011 round could be drawn from the information provided from ICP documentation, including reports, minutes of meetings of the Executive Board, the technical advisory group and the regional coordinator, and surveys and interviews conducted with the major stakeholders of ICP 2011.

17. The questionnaire distributed to the stakeholders covered a number of subject areas, such as a global versus a regional perspective, the governance framework and the implementation of the roles and responsibilities of the distinct governance bodies, the decision-making process and the communication between the main actors in ICP 2011. The questionnaire (see below) had a focus on the stakeholders closely involved in the ICP 2011 round.

18. In addition, interviews were conducted with selected representatives of the stakeholders. Important additional background information came from the annual ICP reports to the Statistical Commission, the ICP handbook and book, the operational material and guide and ICP quarterly updates and regional reports.

### **Questionnaire distributed to stakeholders**

#### **A. General issues — management**

- General evaluation of ICP 2011
- General governance issues (cooperation at the global ICP level, with the regional coordinator)
- Dissemination of ICP results and access to input data

#### **B. Technical issues related to the preparation and conduct of surveys, the validation of input data, computations, etc.**

- Preparation of product specifications
- Price collection, data entry, intra-country validation
- Inter-country validation, calculation of PPPs under basic headings and their aggregation
- Breakdown of expenditure to basic heading ICP classification and quality of national account data
- ICP toolkit
- ICP handbook, ICP book, operational materials and guide, ICP website and ICP quarterly update
- Capacity-building and integration with national statistical programmes

#### **C. Evaluation of the ICP 2011 methodological issues**

- Global core list versus the ring of the countries
- Principles of establishing the regional and global core lists: representativity versus importance
- Treatment of comparison-resistant areas (like housing rents, non-market services, construction) in the regions and in the global comparison
- Computational aspects of estimating basic heading and aggregated PPPs (within regional and between regional)

19. The major findings are presented below. It has to be noted that these findings are based only on the opinions of the main stakeholders. Countries' views and users' opinions will be covered in part II of the evaluation report.

## **A. General evaluation**

20. There is a general understanding that ICP 2011, with its considerably expanded coverage (from 150 to 199 countries), has brought a much higher acceptance compared to earlier exercises. The publication of ICP 2011 led to an open discussion of pros and cons of using PPPs vis-à-vis exchange rates.

21. It seems that either the arguments for using PPPs are becoming widely accepted or that the arguments against them — their limited availability and reliability — are much weaker than they used to be. The use of PPPs worldwide has increased with their wider availability.

22. The methodology applied in ICP 2011 is a significant improvement over that used in the 2005 round. ICP 2011 has put ICP on a firm methodological basis as a result of the introduction of approaches such as the global core list and the application of major innovations such as the ICP toolkit. Specifically, the broad documentation of metadata and the further development of ICP operational guides and handbooks has contributed much to the knowledge of staff conducting the work. A certain challenge arose from the cumulative effect of the two ICP rounds (2005 and 2011), which took the ICP from a one-time “snapshot” created by each solitary benchmark into a kind of time series-like environment with the requirement of time consistency.

23. One clear lesson from the 2011 round is that a six-year interval between rounds is too long. This perception is expressed by almost all stakeholders (even before asking any users' opinions) and is highly interrelated with the acceptance of ICP results, which would be needed at least every two to three years, with extrapolations to annual results. At the moment, at the global level and with the exception of the OECD/Eurostat programme, there is no mechanism to ensure this. Part II of the evaluation report will provide a more in-depth analysis of the issue of the frequency of and intervals between ICP rounds.

## **B. Governance structure**

24. The responses of the parties involved allow the conclusion that the governance structure has generally worked quite well. The structure of interaction between the global and the regional levels, with the Executive Board as the central decision-making and strategic body and the Global Office and Global Manager as the executive units, has been widely proven to be appropriate for handling the main challenges and problems of the ICP exercise. The new structure and the management initiatives taken have turned ICP into an efficient global system.

25. The federated approach to governance, with collaborators at the regional level — albeit with some inefficiencies at certain stages and in some regions — was a good solution. However, not all regions signed a memorandum of understanding or had a clear partnering arrangement with the Global Office, which means a lack of strict agreements on methodologies and timetables among the stakeholders already



involved at the outset of the process. This sometimes led to long discussions and produced considerable delays in the submission of input data and in the finalization of regional and global results. Therefore, frequent additional meetings organized by the Global Office were necessary to calibrate procedures and methodology.

26. One clear lesson learned from the ICP 2011 round is that the computational methods must not be changed in any way once results are known to countries. Similarly, countries should not be able to opt out of the comparison exercise after seeing the results. Both issues would call into question the credibility and reliability of the ICP exercise as a whole. Memorandums of understanding or other written agreements, including on the methodology and procedures, should be signed by all actors at the launch of the ICP exercise.

27. Concerning operational aspects, such as the development of the global core list, survey materials and national accounts materials, the cooperation between the various players worked well. The contribution of regional coordinators in global meetings and the Global Office's contribution in regional meetings enriched the discussions and ensured a smooth progress.

28. A number of supporting activities provided by the Global Office, Technical Advisory Group and World Bank deserve special mention, as they contributed much to the success of the programme:

- (a) The provision of technical assistance and methodological guidance (offering knowledge and expertise);
- (b) The efficient organization of meetings;
- (c) The prompt resolution of software problems (e.g. the introduction and further development of the ICP toolkit);
- (d) Coordination efforts in finalizing the results;
- (e) The flexible provision of financial support.

29. The exchange of data and communications between the global and regional levels worked better than in the 2005 ICP round. Nevertheless, certain problems emerged. Communication generally went from the Global Office to countries via the regional coordinators through the exhaustive ICP website and e-mail communications. However, a feedback loop to get information back from the countries to the Global Office was missing. This also relates to the submission of metadata on national annual averages, national accounts and survey frameworks. A specific communication system/framework could be established that is more secure, transparent and timely (similar to the Eurostat system).

30. The ICP regions had different levels of ICP experience. Respectively, the contribution of the regional coordinators to regional coordinator and Technical Advisory Group meetings was different. Regions with less ICP experience in some cases had difficulty in following the approach recommended by the Global Office.

### **Executive Board**

31. The Executive Board, as had been the case in ICP 2005, acted as the central decision-making and strategic body of the ICP 2011 round. Key international organizations (the International Monetary Fund, OECD, Eurostat and the Statistics Division) had a permanent seat, and several national statistical institutes (mainly from

big economies) were represented on a rotating membership. Executive Board members were appointed on the basis of their deep statistical knowledge and experience in developing supranational statistical indicators and programmes supporting capacity-building, following the rules adopted by the Statistical Commission. The World Bank was present for similar reasons and owing to its role as the host agency. Country representation was drawn to ensure a balanced representation of regions, country-specific economic structures and statistical capacity. Regional coordinators were present to reflect the federated approach to developing PPPs first regionally and second globally.

32. According to those principles, one could say the Executive Board for ICP 2011 was well balanced in terms of regional representation, country size, capacity and the mix of permanent and rotating members. The Board in general had a mix of strategic and technical minds and had sufficient experience to oversee the programme.

33. However, with 25 to 35 members (including 16 from international organizations) the Board was too large to conduct its business effectively. Too large a group risks losing focus and seeing fragmentation into small subgroups advocating for specific objectives or concerns rather than contributing to the overall objective. The number of Executive Board members should thus be strictly limited in order to make it an effective decision-making body.

#### **Global Office and Global Manager**

34. The institutional setting of the ICP 2011 Global Office and Global Manager was as follows: the Global Office acted as the executive office of the Executive Board and was responsible for the overall coordination of ICP. It was located at the World Bank's headquarters in Washington, D.C. The Global Office reported, through the ICP Global Manager, to the Director of the Development Data Group of the World Bank. On matters related to the execution and implementation of the ICP mission, its policy, programmes, priorities and standards, the Global Manager acted within the directives provided by the Executive Board and within the framework of the work programmes and budgets approved by the Board. Thus, elements of the day-to-day management of the ICP Global Office were cleared by the international statistical programmes cluster manager (e.g. budget allocations from the World Bank's budget and the ICP trust fund, staffing actions, performance assessments, travel, etc.). This reporting arrangement allowed for synergies and the sharing of experience between the ICP Global Office and other statistical capacity-building activities provided by the Development Data Group (also under the responsibility of the international statistical programmes cluster manager). This was an implicit additional contribution that the World Bank made to ICP.

35. The Executive Board/Global Office/Development Data Group cooperation model was operating efficiently, as all parties well understood their respective roles and responsibilities. However, the fluctuation of persons in relevant positions of the Global Office was felt to have disrupted the workflow somewhat.

#### **Technical Advisory Group**

36. The Technical Advisory Group, consisting of international experts, assisted the Global Office with the clarification of conceptual, methodological and technical questions arising during the comparison. The work done by the Group was greatly appreciated by most stakeholders and was mentioned as an important contribution to

the success of ICP 2011. A number of methodological questions were clarified and conceptual innovations prepared. The process of nominating the members generally adhered to the ICP governance framework document, as approved by the Executive Board and endorsed by the Statistical Commission. The nomination procedure was found to be largely acceptable, although not always comprehensible in the sense that Technical Advisory Group experts should have experience with and excellent knowledge of ICP methodology and be independent, without representing specific economies or institutions. The 2011 Technical Advisory Group had 22 members, which was too many. Furthermore, the relatively high number of around 40 persons (members and observers) temporarily attending the meetings of the Group was not perceived as being efficient. The number of members should be limited in order for it to be effective.

37. The new and improved methodologies increased the complexity of the calculations. Therefore, three task forces were formed: the Validation Expert Group to oversee the validation of data provided for the global comparisons; the Computation Task Force to calculate the global results independently from each other and ensure their convergence; and the Results Review Group to review the results in terms of their plausibility and adherence to the agreed-on methodologies and procedures. The Technical Advisory Group and all three groups contributed greatly to the overall quality of the final results and also ensured transparency regarding the overall process.

#### **Regional coordinators**

38. The regional coordinators had the difficult task of coordinating the regional comparisons and ensuring a smooth workflow between the Global Office and the countries. They were responsible for planning the surveys, providing relevant documentation, building methodological capacity, collecting and checking country data and transmitting the regional data to the Global Office in a timely manner. The degree of success in the implementation of ICP varied across regions, depending on the experience and level of expertise of the parties involved.

39. Some regional organizations and national statistical institutions drew on assistance provided by the Global Office and partnering arrangements. Other regions, including the OECD and European Union regions, could successfully rely on their existing advanced comparison methodology and infrastructure.

### **C. Capacity-building**

40. A large majority of the ICP stakeholders confirm that ICP 2011 had wide positive effects on the regional statistical programmes, regarding both price statistics and national accounts. ICP made substantial contributions in building national statistical capacities (institutional and technical) in both larger and smaller countries. Credit is due to many of the regional agencies for their support, both financially and in providing assistance on technical and methodological problems. Participants honour in particular the work done to improve the statistical base for cross-country comparisons, as well as the contribution of the programme to the improvement of the statistical basis for shedding light on poverty problems.

41. Simultaneously, the regional coordinators confirm that the various parts of ICP had a significant impact on the workload of the regional offices and national

statistical institutions. The extra workload was significant for countries that had to price several hundred products outside of their consumer price index baskets and obtain expenditures in a very detailed classification of gross domestic product (GDP). In all regions, substantial efforts were made to integrate ICP into national statistical programmes.

42. An important element of the ICP 2011 round was the adoption of full GDP coverage. This applied to most regions, with the exception of the Pacific region (consisting mainly of smaller countries and territories), which participated in the household consumption part only. For GDP components such as housing and construction, new or alternative methods were developed for which an adequate national statistical basis was not in place. As a result, parts of the survey programme have been conducted using a centrally managed approach, often with limited or no regional or national involvement.

43. Most regions seem to accept the arguments for a central approach for the more technical subject areas covering health products, education, equipment goods, government and housing. The attitude is mainly or wholly positive, but many respondents argue for increasing regional involvement in future rounds.

44. For the 2011 round, capacity-building was an integral part of the continuous process for improving data quality. Regional coordinators provided assistance and support in solving the various problems encountered. In addition, the regional seminars and workshops on prices and national accounts were important forums bringing statisticians together for the presentation and discussion of new methods and operational practice.

45. The Global Office contributed substantially to capacity-building in the ICP 2011 round by preparing the following material:

- *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program (ICP)*, published by the World Bank, which presents the methodology available for international comparisons, the choices made for the ICP 2005 round, the outcome of those choices and the steps to be taken to improve the quality of the data for the 2011 round. To ensure the transparency of the estimation process, the book presents an in-depth examination of the underlying theory, the methods used and the problems encountered. It also offers a forward-looking perspective on the methodology-related steps that can or will be taken to improve estimates in future ICP rounds.
- The ICP Operational Guide, which provides practical guidelines and descriptive illustrations of methodologies. This work is a compilation of operational materials and instructions that can improve the practitioner's understanding of recommended methodologies.

46. Beyond this material, the Global Office provided critical technical support to countries and regional coordinators, often on short notice, relating to: (a) survey methodology; (b) data validation; (c) national account expenditure compilation; and (d) PPP calculation. This support was provided through various missions and regional and global workshops, as well as through other forms of communication.

47. Building on lessons learned from the 2005 round of ICP, the Global Office developed the ICP software for 2011 (also known as the ICP kit) as a comprehensive

set of physically independent but logically integrated software modules to coordinate data collection and data-processing during the 2011 round. The kit used in the ICP 2011 round had three modules: (a) country data entry and validation for household consumption; (b) country data entry and validation for specific surveys; and (c) regional data validation, processing and reporting module. These modules are available in Arabic, English, French and Spanish. The regional module handles set-up and inter-country validation, including diagnostics and index processing, for all price survey data. The country modules for household consumption items and the compensation of Government employees cover set-up, implementation, price averaging and the intra-country validation of prices. The national accounts module helps to process the disaggregation of expenditures and validate them across the countries. The ICP kit was used intensively in some regions (e.g. Western Asia and Latin America). Some other regions (e.g. Asia-Pacific and Commonwealth of Independent States and OECD/Eurostat) used their own regional software. There were several updates of the ICP kit because of problems that arose.

#### **D. Quality of data and metadata**

48. The regional agencies responsible for the comparisons in the seven ICP regions (AfDB, ADB, the Statistical Office of the Commonwealth of Independent States, ECLAC, ESCWA and the Australian Bureau of Statistics) shared that responsibility with the national agencies, which collected and validated data within their respective economies. The regional agencies provided the national agencies with methodological and operational guidance. They coordinated and supervised the collection and validation of data within the region in line with the global timetable. They also put together and finalized the regional comparisons and published the results. The Global Office assumed responsibility for ensuring that the seven regional comparisons and the OECD/Eurostat comparison could be combined in the global comparison. The compilation, validation and publication of the global results were also responsibilities of the Global Office.

49. Ensuring the quality and completeness of input data and metadata is a crucial issue for ICP. The Global Office prepared a series of metadata questionnaires to ensure that the 2011 round would have better metadata than the 2005 round. Quality assurance framework checklists were prepared to help collect the information required to evaluate and assess the quality of the submissions. They were available in Arabic, English, French and Spanish. The checklists had to be completed by the countries, the regional coordinator and the Global Office.

50. On the national accounts side, the Model Report on Expenditure Statistics, the national accounts quality assurance questionnaire and the national accounts exhaustiveness questionnaire provided a wealth of metadata on how countries were splitting their GDP data on expenditures into 155 basic headings and on the quality and exhaustiveness of their expenditure estimates. On the price survey side, a survey framework questionnaire was prepared in Arabic, English, French and Spanish. This questionnaire was crucial, providing information on the coverage of the survey across countries.

51. The ICP quality assurance framework was derived from the data quality assessment framework developed by the International Monetary Fund (IMF). The IMF framework brings together best practices and internationally accepted concepts

and definitions in statistics, including those of the Fundamental Principles of Official Statistics of the United Nations and the General Data Dissemination System of IMF. It provides a structure for assessing data quality by comparing a country's statistical practices with best practices. It focuses on the quality-related features of governance of statistical systems, core statistical processes and statistical products. It was developed to introduce rigour, structure and a common language in the assessment of the quality of macroeconomic data. The ICP quality assurance framework covered six topics: prerequisites of quality; assurance of integrity; methodological soundness; accuracy and reliability; serviceability; and accessibility.

## **E. Timeliness and publication of results**

52. The ICP website is the centre for information about ICP. It is a well-organized site providing broad information on the Programme, the entities involved, the activities, the survey programmes and important uses of PPPs. For each region there is a separate page providing information on the region as well as reports on progress and other relevant information. The site contains important documents on the ICP surveys and methods, the ICP book and handbook, the operational guidelines and manuals and research and development information. Information about the progress of ICP and new methods and tools was briefly presented in the ICP Quarterly Update.

53. Transparent information on the various issues was a central element in the processes up to the dissemination of final global results. In the case of the Executive Board, the Technical Advisory Group and the meetings with the regional coordinators, the minutes of each meeting held were made available on the ICP website. The ICP Quarterly Update offered to users a number of relevant articles on the various ICP issues covering governance, in-depth presentations of new elements in the survey programme and some country or regional reports on the experiences.

## **F. Technical aspects and methodologies**

### **Establishing the item lists**

54. In the 2011 round, some areas, such as housing, private health and education, government services, equipment goods and construction, were managed with less involvement from the regions than was the case, for instance, for market consumer products. Generally, this centrally managed approach was widely accepted. Regional coordinators were able to add regional items to the global core list items. Data collection and editing were not centrally managed by the Global Office. There was a three-stage data-editing approach: (a) intra-country validation; (b) regional validation, during which the country's prices were edited and verified by the regional coordinators; and (c) interregional or global validation, during which the prices collected for global core products were edited and verified across all regions by the Global Office.

55. The main reason for the centrally managed approach was the need for the greatest possible worldwide consistency in handling those areas, which was different from market consumer products where region-specific product lists were set up, supplemented by the global core list. The relevant work was managed centrally by the Global Office, but all key decisions were discussed extensively at

Technical Advisory Group meetings, which all regional coordinators attended, and further reviewed for data implications at regional coordinator meetings.

56. For the sake of comparability across regions, but also for transparency, a centralized or at least harmonized approach to housing, non-market services, equipment goods and construction seems in fact to be the better solution. Nevertheless, it must be stated that the global approach was not always well understood in certain regions and that regions have different levels of technical capacity and hence are not all able to participate uniformly. Generally, more active involvement of the regions in all aspects of the work would have been appreciated, for example before the finalization of the lists, so that regional coordinators could make recommendations and/or clarify its representation to the region. Regions should have had more opportunity to tailor the surveys to their needs, under the condition that it remains possible to link the results across regions without loss of comparability.

57. Another approach was taken for creating a global core list for market consumer products. Regional product lists were identified, then the regions decided on the inclusion of the global core list products in those lists.

58. The process of creating global core lists has several limitations and needs to be improved. The regional product specification coding was not as uniform as would be desirable. Not enough was done to use the ICP structured product description as the basis for a registered international product description standard. The global core list definitions were not always very clear, in some cases being too loose. The procedures for the selection of the global core list products (i.e. the number of products per basic heading) were not clear and as a result some basic headings were not covered adequately for some regions. There was a need to develop a special survey for information technology products at the last moment. Not all regions provided sufficient input to its development.

59. The process of developing the global core list specifications could be made more transparent, for example by using common online tools with which all regions can make their proposals. Ideally, the products proposed for the global core list would undergo a preliminary survey to determine their availability and importance in regions and countries before they are included in the survey. This would also help in improving the specifications.

60. Another aspect concerns the “fait accompli bias”: differences in the survey schedules between ICP and the OECD/Eurostat PPP programmes were resolved by lending priority to OECD/Eurostat items in the preparation of the global core lists for household consumption and machinery and equipment. This is because many of their surveys had already been conducted before the ICP field activities were even scheduled. For several basic headings, the global core list was significantly affected by an “OECD/Eurostat fait accompli syndrome”, resulting in a “fait accompli bias”.

61. As a matter of fact, but not surprisingly, there was too much emphasis on the work on consumer goods and services to the detriment of the other surveys. This might be a consequence of the greater complexity of comparing investment goods and construction versus consumer goods and services.

#### **Importance versus representativity**

62. The comparability of the products being priced is the fundamental principle underlying the estimation of PPPs. A dilemma facing ICP is that available

comparable products may have significantly different expenditure shares. Since there are no explicit data on expenditures for individual products, indirect or implicit rough weighting is used to obtain “unbiased” basic-heading PPPs. In ICP 2005, the Technical Advisory Group recommended use of the concept of “representativity” elaborated by the European Union and OECD. A “representative” product should reflect the typical price level for a basic heading. Such products are purchased frequently by households and therefore have lower price levels relative to non-representative products. In accordance with this assumption, the extended country product representative dummy (CPRD) method (or Elteto-Koves-Szulc (EKS) method) was recommended by the Technical Advisory Group for the calculation of basic-heading PPPs. The ICP regions attempted to use the “representativity” concept in 2005 but were unable to consistently apply the notion of a representative price level.

63. As a result, the concept was not used in 2005 in the ICP regions (beside European Union/OECD and the Commonwealth of Independent States) or for estimating interregional linking factors. Therefore, the Technical Advisory Group recommended another concept for the 2011 ICP. Countries were asked to classify products for household consumption as either important or less important. Importance was defined by reference to the notional expenditure share of the item within a basic heading. The CPRD method was replaced by the weighted country product dummy (CPD) method.

64. This innovation was not unambiguously accepted by all participants, although in general the concept of “importance” is less ambiguous, simpler to understand and more transparent, but it does not solve all problems for the calculations of “true” (unbiased) basic-heading PPPs if the “typical price level” is ignored.

65. The introduction of the concept of “importance” led to deeper discussions concerning the concept of “equi-representativity”, and therefore it was useful. However the “new” concept did not improve the situation practically and did not decrease the degree of subjectivity for the attribution of the “importance” indicators. In addition, the choice of weights (important versus non-important) is also inevitably subjective. Therefore, further efforts should be focused on additional explanations to countries concerning the entire process. The countries should focus on expenditure shares during the selection of products and should have a good understanding of why the concept of “importance” (significant expenditure shares) is used mainly at the first stage (the selection of products) and the concept of “representativity” (typical price levels) is used mainly during the second stage (price collection, validation and calculation). This allows more conscious validation by the countries and a better evaluation of the results.

## **Rents**

66. Several factors make the comparison of housing expenditures challenging, for instance, the recording of housing expenditure and significantly varying market situations from one country to another. There were substantial efforts from the regional coordinators and the Global Office to improve the methodology as well as input data (prices as well as national accounts figures) in this complicated area. Actual progress was rather moderate, however, the reference volume indices were used for rents in the Asian region (as in ICP 2005). The global specifications for pricing were used in only three regions (Africa, Latin America and Western Asia).



The global linking for OECD/Eurostat and Asia was done in a very rough way by the quantity approach, with the number of dwellings simply adjusted on the basis of three quality indicators (availability of electricity, water and toilets). The available input data were very weak and did not allow a more sophisticated approach. Adopting the quantity approach was making the best of a bad situation. Most likely, the dual approach for the collection of both rental and quantity data needs to be maintained. Improvements in this area should come primarily from improvements in data quality and availability (price and dwelling stock data as well as national accounts data).

### **Government services (health, education and collective services)**

67. The use of a productivity adjustment for the global linking of salaries in government services in ICP 2011 was an obvious improvement. There were, however, several weaknesses: the productivity adjustment calculated using capital-labour estimates for the whole economy for such specific areas as health, education and collective services is a very rough approximation (also, the accuracy of productivity adjustment factors for different countries seem to be different), and not all regions used the productivity adjustment in the regional comparisons (OECD/Eurostat, Commonwealth of Independent States and Western Asia did not use productivity adjustment). In addition, the regions did not use the same methodology (e.g. OECD/Eurostat used the “output” approach for health and education, but other regions used the “input cost” approach). In effect, the results of the countries depend, in some cases heavily, on the regional methodology, thus decreasing interregional comparability. For example, the results obtained by the “output” approach in education are often opposite those obtained by the “input cost” approach with productivity adjustment. No doubt the OECD/Eurostat experience with the “output” approaches in these areas should be analysed carefully by all regions.

68. The challenge with the “output” approach, however, is that it requires input data that are not always available or are of a questionable nature. For example, even the United States of America was included in the OECD/Eurostat comparison of hospital services using the “input cost” approach. In addition, the “output” approach does not need productivity adjustment, but the comparability of the quality of services should be considered. For the time being, the input-based approach with productivity adjustment is most likely the best possible option for ICP. The process for calculating productivity adjustment factors as well as the application of those factors should be streamlined, however.

### **Construction**

69. Construction is a difficult-to-measure area even within a given country because it is site specific and the sites change from month to month. The ICP 2011 approach was a compromise, making the best of less than uniform measurement standards for this tough-to-measure activity. In general, it can be argued that the 2011 input-based approach was the best possible and most achievable option, given the limited availability of data, as it is less costly than some of the other options. On the other hand, the input costs (materials, labour and equipment) do not take into account the outstanding gap among economies for profit, taxation and contractor mark-ups, and therefore do not fully reflect market prices.

70. It is suggested that further research regarding these problems are needed in the next round of ICP and that there should be a greater focus on mark-ups so as to find a better approach for construction. A model-based approach would clearly be better. Ways of funding such an approach on a global basis should be investigated well before the next round.

71. If the input cost approach is kept, the following improvements should be considered:

(a) Better and more detailed specification of items and replacement of a few items deemed to be difficult;

(b) Better coverage of input costs, mark-ups and the like to obtain a better approximation of market prices;

(c) Redesign of the survey questionnaire to ensure that all required data are collected unambiguously;

(d) Ensuring that metadata are collected and recorded better.

### **Global linking**

72. The decision to adopt the global core list approach rather than the 2005 “ring approach” was based on a detailed analysis of the 2005 results. From a global perspective, the application of the global core list was a clear improvement, especially as it ensures more robust linking because it is based on the set of prices for all participating countries.

73. The approach should be further improved, however, by refining the item definitions and ensuring that the concepts, such as “well-known brand”, are used systematically throughout the list. In addition, survey guidelines can be strengthened for some difficult areas, such as health, housing and equipment goods.

74. New computational procedures were introduced in ICP 2011 for the global linking. At the basic-heading level, the weighted CPD method with weights of 3:1 was used to calculate the PPPs instead of the unweighted CPD method, which was used in the 2005 round. At the aggregated levels, a new procedure, the country aggregation with redistribution (CAR), replaced the 2005 super-region method in which linking factors were computed for regional aggregates.

75. A common opinion is that both weighted CPD and CAR methods were improvements and should be maintained in future rounds of ICP in order to ensure greater consistency across rounds. Nevertheless, some further technical improvements and investigations are desirable. This concerns the weighted CPD (depending on further discussion on “importance” and “representativity”) and the CAR-volume approach (for example, the elimination of “weak” links through countries with very different price and expenditure structures).

### **Fixity of the regional results**

76. If one assumes that the regional comparisons have different degrees of accuracy, then the fixity of the regional results can be seen as an advantage. Another benefit of maintaining regional fixity is that the PPPs published by regional organizations are in line with those published by the Global Office. It is important for the credibility of the ICP results to preserve fixity and to have only one set of

results for each region. Some regions use the PPP results for administrative purposes and do not want to see revisions they are not directly responsible for.

77. However, in general terms, it is hard to support the concept of “fixity”. This is essentially an organizational, not a statistical constraint. Maintaining fixity makes the calculation process less transparent and is more labour-intensive. The main problem is not fixity per se but rather the fact that the regions use different methodologies, which decreases the comparability of the global results.

78. Different approaches according to regions should be allowed as long as they are anticipated from the beginning and linking procedures are also established from the beginning, but sufficient effort should be made towards methodological unification. The official ICP results should be published with fixity. However, many users are interested in results that compare countries more directly without imposing the fixity constraint. It would be good to have more experimentation (in addition to the official ICP results) than there has been time for.

#### **IV. Procedure for completing part II of the International Comparison Programme evaluation report**

79. The present report is submitted to the Statistical Commission at its forty-sixth session for its consideration, comments, discussion, approval and advice on further work. After that, the Friends of the Chair will gather additional information from countries participating in ICP 2011 as well as the potential users of the ICP results. Both sets of information will be gathered by a survey questionnaire and additional interviews. This information, together with an in-depth analysis of the possibility of shortening the intervals between individual ICP rounds and the methodological and financial issues attached to greater frequency, will form the basis of the final evaluation report, to be submitted to the Statistical Commission at its forty-seventh session, in March 2016.

#### **V. Recommendations**

80. The 2011 round of ICP was a major step forward in developing a system of calculating PPPs on a global basis. Based on the information collected on the ICP 2011 round the Friends of the Chair group makes the recommendations set out below.

##### **A. Process**

81. It is recommended that the methodology and procedures to be applied during a comparison cycle shall be laid down in a memorandum of understanding or any other form of agreement among all important stakeholders (the Executive Board, the Global Office, the regional coordinator and the national statistical institutes) at the outset. After the methodology and procedures are signed off on, they cannot be changed later in the process.

82. It is recommended that the Executive Board have ownership of the input data as soon as the regional coordinator has finalized the validation and countries have

approved their data. The responsibility for finalizing and publishing the results lies with the Executive Board.

## **B. Governance**

83. It is recommended that the Executive Board continue to act as a strong central decision-making and strategic body. The Board should report to the Statistical Commission as the ultimate owner of the ICP exercise.

84. It is recommended that the number of Executive Board members be limited. International organizations such as the United Nations, the World Bank, IMF and OECD should be permanent members. In addition, the Statistical Commission should nominate four national statistical institutes from among its members to the Executive Board.

85. It is recommended that the Executive Board be supported by a technical expert group. The Board should nominate, on the basis of purely professional criteria, a pool of 10 to 15 experts to assist it on technical and methodological issues. The group can form task forces on specific topics. Technical Advisory Group and task force members are independent experts, not representing any specific region or institution.

86. It is recommended that an institutionalized feedback loop be established between countries/regional coordinators and the Global Office, with a view to improving information flows.

87. It is recommended that the Executive Board be supported by a coordination office (like the Global Office) that would assist it in the preparation of its work and interact with the regional coordinator at the regional level. For the efficiency of the global coordination task, it is essential that positions in the Global Office be adequately and permanently staffed during a cycle of comparison.

## **C. Research agenda/methodological aspects**

88. It is recommended that the following points be implemented as improvement actions or included in a research agenda:

(a) The global core list approach should be applied for global linking and the new computational aggregation procedure — the CAR-volume approach — should be applied, as it ensures more robust global results and better consistency across rounds;

(b) When applying global core lists, broad use should be made of the ICP structured product descriptions and the common online tools in which all regions can make their proposals. Thus the procedures for establishing the global lists and for the validation of global price data would become more transparent and involvement of regions in all aspects of the work would be improved;

(c) For the attribution of the “importance” indicators at the detailed product level, the use of detailed product weights should be investigated;

(d) A centrally managed approach should be applied for areas such as housing, government services, private health and education, equipment goods and construction for the sake of better comparability across regions;

(e) Concerning rents, the dual approach (output prices and quantity approach) should be maintained as long as availability and quality of data (price and dwelling stock data and national accounts data) needed for general use of the price approach are not achieved;

(f) For the comparison of government services, health and education, further improvements should be made in “output”-based approaches and on the increase of the methodological harmonization between the regions;

(g) For construction, a more output-based approach should be targeted, since input costs do not reflect market prices (profits, taxation, contractor mark-ups, etc.).

## **VI. Points for discussion**

89. **The Statistical Commission is invited to express its views on:**

(a) **The recommendations proposed by the Friends of the Chair group;**

(b) **The further procedure for the finalization of the evaluation of the 2011 round of ICP.**