

## ICP Data Access and Archiving Policy



**3<sup>rd</sup> Executive Board Meeting**

**October 18, 2010**

**Washington DC**

## TABLE OF CONTENTS

<b>1. Introduction .....</b>	<b>3</b>
1.1 Background .....	3
1.2 Policy Objective .....	4
1.3 The Core Elements of the 2005 Policy.....	4
<b>2. User Needs .....</b>	<b>5</b>
2.1 General Needs .....	5
2.2 World Bank and Regional Coordinator Needs.....	6
2.3 Issues from a Country Perspective .....	6
2.4 Analysis of the Key Issues .....	7
<b>3. Proposed 2011 ICP Data Access and Archiving Policy.....</b>	<b>7</b>
3.1 Policy Directives .....	7
3.2 Conditions for Approval of Requests.....	8
3.3 Analysis of Policy .....	10
<b>4. Implementation of the Policy .....</b>	<b>12</b>
4.1 Modalities.....	12
4.2 Criteria for Approving Applications .....	12
4.3 Feedback to Countries.....	13
4.4 Treatment of Breaches by Data Users.....	13
4.5 Archival Requirements.....	14
4.6 Procedures for Reviewing Policy .....	14
<b>5. Annexes .....</b>	<b>15</b>
<b>Annex 1: ICP Data Types .....</b>	<b>15</b>
<b>Annex 2: ICP Metadata.....</b>	<b>16</b>
<b>Annex 3: ICP Data Access Principles .....</b>	<b>18</b>
<b>Annex 4: Results from Country Responses to the Data Access Questionnaire .....</b>	<b>21</b>
<b>Annex 5: Template for Request for Data Access and Declaration of Confidentiality .....</b>	<b>23</b>

## 1. INTRODUCTION

### *1.1 Background*

The International Comparison Program (ICP) is a worldwide statistical initiative to collect comparative price data and estimate purchasing power parities (PPPs) for individual countries or economies. Using PPPs instead of market exchange rates to convert currencies makes it possible to compare the output of economies and the welfare of their inhabitants in real terms - that is, controlling for differences in price levels across countries.

The ICP entails large price data collection and detailed national accounts expenditure data compilation. The resulting ICP databases at the participating countries, the regional coordinating agencies, and the ICP Global Office contain valuable price and national accounts information, which could support the empirical research of many international, regional, and national agencies, as well as universities and research centres. However, some countries are concerned about the confidentiality of their price data. They are also concerned about the quality of the ICP data at that detailed level and criticisms they may receive, noting that the ICP was not designed to provide reliable data at the detailed level.

In the 2005 round, the Data Access and Archiving Policy strongly limited access to data below the basic heading level data. Also, in some regions, the memoranda of agreement between the regional coordinating agencies and the countries restricted researchers' access to detailed price data. However, at the end of the 2005 round, it was clear that some of the more important researchers were mostly interested in more detailed data than was available from the 2005 round.

Based on the experience of the 2005 round, the 2011 round will attempt to achieve two fundamental goals in the area of data access: (1) securing the confidentiality and security of data, and (2) ensuring effective dissemination and use of data. A trade-off arises between achieving those two goals. The 2011 round will try to strike a balance between users' needs for access to detailed micro data and participating countries' concerns about this level of access including both confidentiality and quality concerns. However it should be recognized that country's laws on confidentiality cannot be breached under any circumstances.

Also, this new ICP round will focus on collecting, archiving, and providing access to metadata, to increase the quality and utility of the data collected. The limitations in the available metadata for the 2005 round made the ICP data seem like a 'black box' to the users of detailed ICP data.

One of the innovations of this round will be the development of a new methodology to link the regional results. This methodology involves putting in place a Global Core list of products and services that will be priced by all participating countries, including the countries of the OECD/Eurostat PPP program. This will be a unique exercise that will result in a database of comparative price data from more than 170 countries for around 600 products and services. This database of average prices would be extremely valuable to researchers around the world.

Given the proposed new direction of the 2011 round in terms of responding to user's needs by providing wider access to detailed data and increasing the focus on metadata, and given the new innovations of the round that will result in valuable price and national accounts information, there needs to be a policy on data access and archiving to guide practice in these areas. This is the primary purpose of this document. ICP data included under this policy are described in Annex 1.

### ***1.2 Policy Objective***

The core objectives of the ICP Data Access and Archiving Policy are:

1. to provide the ICP Global Office and the Regional Coordinating Agencies access to the detailed data required for their compilation, research and dissemination activities;
2. to maximise the value of the ICP data collection effort by providing user access to more detailed ICP data, under specific conditions, whilst respecting the concerns of participating countries about confidentiality and quality; and
3. to set out the arrangements for archiving ICP data to support future uses of the data including those by Regional Coordinating authorities and participating countries.

### ***1.3 The Core Elements of the 2005 Policy***

The following extract from 2005 ICP Data Access and Archiving Policy is relevant:

*“Two kinds of data access will be allowed. First, the World Bank, regional, and national coordinating organizations may commission research to evaluate the methodology and suggest where improvements can be made for the next round of the ICP. If this work is done by consultants under contract to these organizations or the World Bank, they should be given access to the full set of regional and global prices and quantity data as well as basic heading PPPs and expenditures providing they sign a declaration of confidentiality.*

*The second level of access comes from individuals or organizations outside the ICP framework requesting access for research purposes. This type of access should be allowed since it increases the knowledge base on the uses of PPPs and may provide valuable input for future improvements in data quality.”*

This policy required some interpretation and, with the benefit of hindsight, could have been more clearly specified. The way it was interpreted was that:

1. National average price data at the product level were only available for methodology research undertaken by the World Bank, regional, and national coordinating organizations or commissioned by them;
2. Other approved researchers would only have access to data for the unpublished categories and basic headings.

It should be noted that many countries only made national annual average prices available to the regional coordinators but not to the World Bank.

Access to individual price observations was only possible with the approval of the country concerned.

## **2. USER NEEDS**

### ***2.1 General Needs***

The two most important needs for ICP data are to support poverty analysis and ICP methodological research. The latter was reasonably supported by the 2005 policy whereas the former was not so well supported.

The basic requirement for poverty analysis is national average price data at the product level preferably with some disaggregation by location (especially a capital city/urban/rural breakdown) and store type to better understand the prices paid by the poor and how they differ from national average prices. This implies the sample is designed to provide this level of detail. Data were not available at this level of detail for the 2005 round. The users also require expenditure data at as a detailed level as possible although realistically this will not be available at a more detailed level than the basic heading level.

Another important use is investigating how the ICP works from a statistical compilation perspective with the objective of improving understanding of results as well as improving methods. Research into alternative methods, including aggregation methods for index compilation purposes, requires access to more detailed data. These researchers tend to have strong links to the ICP (possibly as a member of TAG). Their main objective is to improve ICP methods although some, given their academic background, would also hope to have their work published in the peer-reviewed literature.

There are some other uses of ICP data such as:

1. Understanding cost comparisons across countries in areas such as public sector wages, health and education to support policy analysis. International comparisons in these areas require data on national average prices with more detail (e.g. wages for occupation groups) than at the basic heading level.
2. Research into the price behaviour of tradeable goods and services to better understand the influence of tax barriers etc. requires average price data at the product level.

From the point of view of data users, including many of the donor organisations, the 2005 policy provided insufficient access to the more detailed ICP data they require to undertake their work. National average price data at the product level, preferably with some disaggregation to better understand the prices paid by the poor, would satisfy most of their needs. For some of the investigations into price index aggregation methods, access to individual price data may be desirable. Almost all of the researchers, involved

in studies on poverty on methodology, are working closely with the World Bank or the regional coordinating agencies or are staff members or work with partner institutions.

The more sophisticated data users accept that the ICP was not designed to provide high quality data at this detailed level. However, it was still useable if there was a quality indicator (e.g. coefficient of variation) and sufficient metadata to enable them to interpret and understand the detailed data. Among other things, this would allow users to adjust for extreme observations.

## ***2.2 World Bank and Regional Coordinator Needs***

The World Bank and the Regional Coordinators require national average price data for the regional and Global Core items to enable them to compile robust global and regional PPPs and Price Level Indexes.

In particular, they are required for the Quaranta and Dikhanov tables, which are used for data validation across countries for both the regional and Global Core items. Where editing indicates potential problems in national average price data or the identification of comparable products, there may need to be access to individual price observations to understand the nature of the problem and whether any adjustments are necessary. The alternative is to work very closely with the country concerned. The precise arrangements should be established through Memoranda of Understanding with the respective regions.

## ***2.3 Issues from a Country Perspective***

The previous two sections highlight that the most important requirement from a data user perspective, and also from the perspective of the World Bank and the Regional Coordinating authorities, is for access to national average price data at the product level together with a quality indicator and appropriate metadata. Access to individual price observations would also be useful for some purposes.

From a country perspective, what are the issues with providing national average price data at the product level? The stated concern from national statistical offices about providing access to more detailed data is often around confidentiality. However, the real concerns are often around data quality as evidenced by the survey of countries. There are two aspects to this. First, there are concerns that the quality of data at the detailed level is not good enough to support research uses. Second, there are concerns that the countries and Regional Coordinating Agencies might be publicly criticised because of these shortcomings. There might also be concerns that the ICP itself might be criticised because of the greater knowledge of quality weaknesses. It is this second concern that has to be recognised in the ICP policy on data access.

There are also concerns about statistical confidentiality particularly for those products where there is a single distributor of the product.

From a country perspective, there are also issues about providing individual price observations in many countries. Even though prices may be largely observable (although

not in all cases), the statistical laws of some countries prevent the release of these data. Other countries can only release microdata under very specific conditions which may restrict ICP uses. Other countries may be able to release unidentifiable data on price observations but may have the quality concerns mentioned above.

#### *2.4 Analysis of the Key Issues*

As outlined in the previous Section, quality is a major issue that has to be managed in providing broader access to national average price data. Criticisms of quality are something that many official statistical agencies, that provide access to detailed data including micro data, need to currently manage. It is not a new problem. With respect to the ICP, the most important messages to be provided to the users of detailed data are:

1. The ICP is of good quality for the purposes for which it was primarily designed.
2. It was not designed to provide data at a very detailed level.
3. However, it is recognised that data at that level might be useful for some specialist users particularly if supported by quality indicators and appropriate metadata.

These explanations are much easier to make if there is a close working relationship with the data users – more of the way of a partnership.

Establishing realistic expectations on the quality of the data is also very important. It follows that access should be restricted to those data users who understand and accept the limitations and whose proposed uses of the data are not dependent on data of a higher quality.

With respect to the confidentiality concerns with products with a single supplier, a solution may be to seek the consent of the manufacturer to be able to release the data to approved data users on the understanding that they would not be able to identify branded products in any of their published work

With respect to individual price observations, given the variety of legal arrangements for releasing microdata, the only practical solution is for the data users to liaise directly with those countries for which they would like data access.

### **3. PROPOSED 2011 ICP DATA ACCESS AND ARCHIVING POLICY**

#### *3.1 Policy Directives*

The ICP Data Access and Archiving Policy is set out below. The key underlying assumptions are that except for products where there is only a single supplier, National Average Price Data are not statistically confidential in the sense that individual businesses can be identified, but Individual Price Observations may be statistically confidential and prevented from further release by the statistical laws of most countries.

Data quality is a further consideration in determining whether ICP data should be released to specific users.

The policy directives are:

Directive 1: PPPs, Price Level Indexes and Expenditure Data will be published at the analytical category level on the ICP web site. They will be available to the public.

Directive 2: PPPs, Price Level Indexes and Expenditure Data for the unpublished categories including basic headings will be available to approved data users only through an online database that is only accessible through a user name/password combination to reduce the effort involved in the servicing of requests. The confidentiality requirements will apply to these users.

Directive 3: National Average Price data at the product level, with appropriate disaggregation and measures of quality, will be available to approved data users only.

Directive 4: Individual Price Observations (microdata) will not be available through the ICP Global Office. Where permitted by the laws of individual countries, the ICP Global Office could provide contact details to those researchers who are interested in negotiating with individual countries.

Directive 5: For ICP compilation and archiving purposes, National Average Price Data will be provided to the ICP Global Office and the relevant Regional Coordinating agencies for both the regional and Global Core items.

Directive 6: There will be supporting metadata for all ICP data releases consistent with good statistical practice.

These policies take account of country views expressed in the recent survey. The results from the survey are found in Annex 4.

### ***3.2 Conditions for Approval of Requests***

It is proposed that the conditions of approval for Directive 3 be more stringent than Directive 2 primarily because the data are of lower quality and will only be useful for specific research purposes by sophisticated researchers. Thus, a user may be provided access under Directive 2 but not Directive 3.

For a data access request to be approved, the following conditions are to be satisfied:

1. There must be a written application seeking approval which, among other things, outlines the uses to which the ICP data will be put.
2. The data users must sign an undertaking agreeing to the conditions of access.

3. The conditions of access include:
  - (a) users confirm that they understand the quality limitations and that the data is satisfactory for their use;
  - (b) users will not publish data at a more detailed level than the analytical categories which will be the most detailed level at which ICP data is published;
  - (c) users will provide feedback to the World Bank Development Data Group (e.g. a copy of a research report) on the results of the use of the ICP data;
  - (d) users will not provide the ICP data to third parties; and
  - (e) for approved users under Directive 3, the data should only be used for research purposes.
4. The user is able to demonstrate that the value of the proposed use of the ICP data warrants the effort involved in supporting the data request.

With respect to (a), the understanding of the quality of the data and how it can be managed in research should be greater for approved users under Directive 3 than Directive 2.

With regards to the entities responsible for approving and providing data access:

1. On Directive 1, after collaboration with OECD/Eurostat and the Regional Coordinating Agencies, the ICP Global Office will make recommendations to the Executive Board for the analytical categories to be published in the public domain.
2. On Directives 2 and 3, the Director of the World Bank Development Data Group, based on advice from the ICP Global Office, will make decisions on whether to approve requests after collaboration with the relevant Regional Coordinating Agencies if the request is in respect of a small number of countries. To ensure the ICP is responsive to data users, Regional Coordinating Agencies should only be given a specified time (e.g. 2 weeks) to respond. The Director of the Development Data Group may choose to seek the advice of experts to help with these decisions.
3. On Directive 4, the participating country will decide whether to release or not but may choose to seek advice from the ICP Global Office or the Regional Coordinating Agency.

Should the ICP Global Office cease to exist following the end of the current ICP 2011 round or any official decision by the World Bank authorities, the ICP Global Office's responsibilities related to the release of data to users will be entrusted by the World Bank Development Data Group to a World Bank staff or unit or team.

### 3.3 Analysis of Policy

The main improvements from the 2005 ICP Data Access and Archiving Policy are illustrated in the table below.

	2005 Policy	Proposed 2011 Policy
<b>A. Global Data</b>		
a. Individual price observations at the product level (microdata) for the products on the Global Core list with identifiers for type of outlet, and urban/rural designation.	No access	No access - Requires country consent
b. National annual average prices for the products on the Global Core list. These include the prices for consumer goods and services from the Global Core list plus the price or quantity data for the global product lists for housing, health, education, government, equipment, and construction.	Ring average prices - Access to World Bank consultants only	Access to approved researchers
c. Disaggregated national annual average prices (according to agreed levels of disaggregation such as urban/rural or capital city vs. other geographical regions).	Not available - Ring survey framework design (Urban/ capital city only)	Access to approved researchers
d. PPPs, Price Level Indexes and expenditure data for the unpublished categories and basic headings for all countries	Access to approved researchers	Access to approved researchers
e. PPPs, Price Level Indexes and expenditure data at the agreed analytical level for all countries.	Public access	Public access
f. Population data and exchange rates for all countries (they are the same as the population data and exchange rates under Regional data below).	Public access	Public access
g. Metadata including estimates for the coefficients of variation.	No access	Access to approved researchers
<b>B. Regional Data</b>		
a. Individual price observations at the product level (microdata) with identifiers for type of outlet, and urban/rural designation.	No access	No access - Requires country consent
b. National annual average prices used to compute basic heading purchasing power parities (PPPs) in each of the regional programs. These include the prices for consumer goods and services from the regional product lists (which includes items selected from the Global Core list) plus the price or quantity data for the global product lists for housing, health, education, government, equipment, and construction.	No access	Access to approved researchers
c. Disaggregated national annual average prices (according to agreed levels of disaggregation such as urban/rural or capital city vs. other geographical regions).	Not mandatory in the survey framework design	Access to approved researchers
d. PPPs, Price Level Indexes and expenditure data for the unpublished categories and basic headings.	Access to approved researchers	Access to approved researchers
e. PPPs, Price Level Indexes and expenditure data at the agreed analytical level.	Public access	Public access
f. Population data and exchange rates.	Public access	Public access
g. Metadata including estimates for the coefficients of variation.	No access	Access to approved researchers

The analytical categories to be published cannot be prescribed at this stage. They will depend on other issues particularly data quality, both with respect to the PPPs and the National Accounts data. Also, there should be some commonality between the basic headings that are combined so that the analytical category can be interpreted in a sensible way. The analytical categories published by OECD/Eurostat should provide some guidance and may provide a useful starting point.

Data for the unpublished categories including basic headings will only be provided to approved data users. This is similar to the 2005 policy except that it is proposed that this data be provided through an online database. However, by using user name/password protection, it would only be accessible by approved users. Releasing through the web should reduce the effort involved in these types of requests.

National Average Price data at the product level is the core data set for satisfying those important data users working on poverty analysis. There will need to be some adjustments to provide a special treatment for confidential items (e.g. products with a single supplier) or items that cannot be released at this detail for other reasons. There may also be a need to exclude some countries because of quality or other considerations. For these reasons, an 'externally releasable' version of national average prices should be created to save effort and ensure consistency in releases. This will provide the data base from which approved requests might be serviced. However, the ICP Global Office and the Regional Coordinating Agencies should receive national average price data without these amendments.

Even though the ICP Global Office will have the national average price data available for all countries, it does not follow that the full database will be made available to service requests. It may be that requests are only for data in respect of a single country or a small number of countries. In these circumstances, it is only necessary to release data for those countries.

There may be special interest in the Global Core product list because it will facilitate analysis across the regions. If so, a version of national average prices for Global Core list products only should be created.

To support poverty analysis, it is necessary to identify the prices paid by the poor. This will be assisted greatly if national average prices can be produced separately for certain areas (e.g. urban/rural) and for selected outlet types. This would not be necessary for countries with low levels of poverty or where it can be demonstrated that prices paid by the poor do not differ greatly from those paid by others. The World Bank and the Regional Coordinating Agencies should agree on the countries for which the dissection is important. The precise arrangements may need to be determined on a one-on-one basis.

A quality measure, such as the coefficient of variation (available from the validation tables), should also be provided to assist interpretation and allow users to make some adjustment for extreme observations. Data on the number of observations would be useful but may have to be adjusted so that those national average prices with one or two observations cannot be identified (e.g. it could be said that three or fewer observations are involved).

Although some data users would like access to individual price observations, this is prevented by law in many countries or only available under prescribed conditions. It is not possible to develop a general policy that allows access. In reality, data users only require individual price observations for specific studies in specific countries. The number of justifiable requests is likely to be quite small. It is most sensible if the data user negotiates directly with the country concerned. The ICP Global Office or the Regional Coordinating Agency might assist by providing appropriate contacts.

As with all uses of statistics, metadata requirements are important. Metadata are vital for understanding the data and its quality. For example, it is important that data users understand where reference PPPs have been used rather than the PPPs being compiled on the basis of actual data. In order to keep service costs to a minimum, a single metadata document should be produced for all users. An international metadata standard for statistics should be used. The ICP Quality Assurance Framework should be the overarching framework for providing the metadata required to support these data users. The proposed content of the metadata is shown in Annex 2.

## **4. IMPLEMENTATION OF THE POLICY**

### ***4.1 Modalities***

A written application for access to data must be submitted and signed by the senior member of the data user group. An outline of the information required in this application is shown at Annex 5. It should be sent to the Director of the Development Data Group at the World Bank.

The request may be for data for the unpublished categories including basic headings and/or National Average Prices. If the more detailed National Average Price data is being requested, the request should contain the justification for accessing this more detailed data.

In accordance with the above policy, the Director of the Development Data Group, based on advice from the ICP Global Office, will make decisions on whether to approve requests but after collaboration with the relevant Regional Coordinating Agencies if the request is in respect of a small number of countries.

The Director may seek other expert advice before making a final decision on applications.

### ***4.2 Criteria for Approving Applications***

The following criteria for deciding whether to approve data access requests are proposed:

1. Does the proposal have merit? Does it really require data at the detailed level?
2. Are the data quality limitations understood? Is the data quality satisfactory for the proposed use? Do the data users agree with this? How will they manage the data quality limitations?
3. Does the data user have the capability to use the detailed data in a realistic way? For example, do they have a prior track record in a related area of research?
4. Are the data users prepared to publish the results of the research or make it available on request (and a copy to the ICP Global Office)?

5. Do they accept the conditions for access as outlined in this policy?

The answers to these questions should be positive before a request is approved. The distinction between requests for access to data for the unpublished categories including basic headings (Directive 2) from those for national average prices (Directive 3) revolves around criteria 1 and 2. Requests for national average price data should pass a higher bar than those for data for the unpublished categories including basic headings before being approved.

Technical advice may be sought in assessing criteria 1, 2 and 3.

#### ***4.3 Feedback to Countries***

The following quote, made in the survey of countries, captures the feeling of many countries.

*“A mechanism should be in place to keep National Statistical Offices (NSOs) informed of details of all data access requests and .....inform NSO once the results are published. In particular, the following information regarding the approved access should be provided: identity of the researcher and the institutions he/she belongs to; research proposal outlining the objectives of the research, how the ICP data will be used, research methodology, expected results, intended use of results, what data will be published and dissemination plan.”*

Thus, there is a need to efficiently provide regular feedback to countries that are impacted by ICP data access requests.

#### ***4.4 Treatment of Breaches by Data Users***

Breaches of the conditions of access should be treated seriously to discourage inappropriate practice. The consequences of breaches should be spelt out in both the Request for Data Access and the approval letter in response to requests.

If breaches do occur, the following measures should be taken:

1. The data user should be asked to return the data and forbidden from future access to ICP data.
2. The institution(s), to which the data user belongs, should be advised of the breach.
3. For some data users, it may be possible to take legal action if the breach is sufficiently serious in nature.
4. For internal World Bank users, this may entail applying the Bank's official disciplinary mechanisms.

#### ***4.5 Archival Requirements***

There needs to be agreement on the archival arrangements. It is most important that the relevant ICP data be archived so:

1. they can be used to service future approved requests for access to data;
2. they are available for possible use in future ICP rounds; and
3. they are available as back-up in case the data are lost for one reason or another by a region or country.

The World Bank Development Data Group and Regional Coordinating Agencies are best placed to provide this service.

The data that should be archived includes:

1. Data for the unpublished categories including basic headings (PPPs, PLIs and expenditure data)
2. National Average Prices at the product level, disaggregated where appropriate
3. Quality measures derived from the data validation tables (Quaranta and Dikhanov tables, as well as tables related to the validation of expenditure data)
4. Supporting metadata

Some countries may like the World Bank and/or Regional Coordinating Agencies to archive microdata because they do not have their own facilities to archive this data. This will require effort by the World Bank and/or Regional Coordinating Agencies and a decision needs to be made on whether this is justified or not.

#### ***4.6 Procedures for Reviewing Policy***

There may be good reason for reviewing this policy in light of experience or other developments. Proposals for change should come from the ICP Global Office or through the ICP Global Office. The final decision should be made by the ICP Executive Board and may be through email exchange. However, the specific comments of the Regional Coordinating Agencies should be sought before any submission is made to the Board.

## 5. ANNEXES

### ANNEX 1: ICP DATA TYPES

#### A. Global Data

- a. Individual price observations at the product level (microdata) for the products on the Global Core list with identifiers for type of outlet, and urban/rural designation.
- b. National annual average prices for the products on the Global Core list. These include the prices for consumer goods and services from the Global Core list plus the price or quantity data for the global product lists for housing, health, education, government, equipment, and construction.
- c. Disaggregated national annual average prices (according to agreed levels of disaggregation such as urban/rural or capital city vs. other geographical regions).
- d. PPPs, Price Level Indexes and expenditure data for the unpublished categories and basic headings for all countries
- e. PPPs, Price Level Indexes and expenditure data at the agreed analytical level for all countries.
- f. Population data and exchange rates for all countries (they are the same as the population data and exchange rates under Regional data below).
- g. Metadata including estimates for the coefficients of variation.

#### B. Regional Data

- a. Individual price observations at the product level (microdata) with identifiers for type of outlet, and urban/rural designation.
- b. National annual average prices used to compute basic heading purchasing power parities (PPPs) in each of the regional programs. These include the prices for consumer goods and services from the regional product lists (which includes items selected from the Global Core list) plus the price or quantity data for the global product lists for housing, health, education, government, equipment, and construction.
- c. Disaggregated national annual average prices (according to agreed levels of disaggregation such as urban/rural or capital city vs. other geographical regions).
- d. PPPs, Price Level Indexes and expenditure data for the unpublished categories and basic headings.
- e. PPPs, Price Level Indexes and expenditure data at the agreed analytical level.
- f. Population data and exchange rates.
- g. Metadata including estimates for the coefficients of variation.

## **ANNEX 2: ICP METADATA**

### **A. At the global level**

1. Broad description of the ICP methodology, including description of methodologies for comparison resistant areas
2. Global product lists, with supporting descriptions, grouped by basic headings
3. Table showing which countries collected price data for particular global products
4. Matrix of importance of Global Core items for all countries
5. National Accounts reporting forms and related instruction notes
6. Table showing which countries collected direct prices for machinery and equipment and/or implemented the alternative approach of the Price Factor Method
7. PPPs which were imputed and the source of the reference PPPs
8. The relative weight of reference PPPs basic headings in the total GDP for each country

### **B. At the regional level**

1. Regional product lists, with supporting descriptions, grouped by basic headings
2. Table showing which countries collected price data for particular regional products
3. Table showing which countries provided other basic data for basic headings for which such data were required
4. Table showing which countries implemented the proposed three methods for housing
5. Table showing which countries provided expenditure data for the different basic headings
6. PPPs which were imputed and the source of the reference PPPs
7. The relative weight of reference PPPs basic headings in the total GDP for each country

### **C. At the national level**

1. Table showing products for which price data were collected

2. Sample framework for ICP
3. Quarters for which prices were collected
4. Estimation method for national average prices (especially when price survey was not spatially representative)
5. Price indices used to adjust prices collected to the entire year 2011, when the price survey did not exactly cover the year 2011
6. Imputations used in lieu of data collection
7. Methodology for deriving expenditure data
8. Methodologies used for housing
9. Methodology used for machinery and equipment

### **ANNEX 3: ICP DATA ACCESS PRINCIPLES**

The ICP Executive Board agreed that there should be a set of principles that underpin the data access policy. There are three basic types of data – (A) those relevant to data for the published categories, unpublished categories, and basic headings such as PPPs, Price Level Indexes and National Accounts data, (B) those relevant to the national average prices at the product level, and (C) those relevant to individual price observations. It may be convenient to have two sets of principles. The first set deals with aggregated data i.e. data types A and B, whereas the second set of principles deals with microdata.

Furthermore, there are some general principles that should apply to all forms of data.

The first and second sets of principles are the most important as the majority of access requests will be for aggregated data. Microdata access requests can be expected to be the exception rather than the rule.

#### **A. General Principles**

The following four Principles are suggested.

1. Given the ICP has involved considerable effort and cost, the data derived from the ICP should be utilized to the maximum extent possible, given the quality limitations of the data, after taking confidentiality constraints into consideration.
2. The agreed policy and procedures for data access should be acceptable to the vast majority of countries. Given that has been achieved; all participating countries should comply with the agreed policy and procedures.
3. The procedures for researcher access to detailed ICP data, as well as the uses of this data, should be transparent and publically available.
4. Releases of aggregated and microdata should be accompanied by appropriate metadata, including metadata that describes the quality limitations of the data.

#### **B. Principles with Aggregated Data**

The following five Principles are suggested.

1. Aggregated data are generally not confidential (in the sense that individual businesses can be identified) and this should not be cited as the reason for not providing access unless genuine confidentiality problems exist.
2. Some proposed uses of ICP may require data quality that goes beyond the design intentions of ICP and therefore access to this data should be limited to researchers who are informed of the quality limitations and agree that the data is still useful for

their purposes and are prepared to make an undertaking with the World Bank that outlines conditions on their use of their data.

3. The integrity and past record of the institution of the researchers should be taken into account when making decisions on access.
4. As there are costs associated with deriving and providing the required data, which can be significant in some cases, access should only be provided where the value of work justifies the marginal cost of providing the data in a form that is suitable to the researcher.
5. Expert advice will be sought to inform decisions on access that are not straightforward.
6. The results of work utilizing ICP data may be put in the public domain but aggregated data should not be published at a more detailed level than is currently the case for the ICP.

### **C. Principles with Microdata**

The 2007 Session of the UN Statistical commission agreed on a set of Principles for access to microdata. These seem like a reasonable starting point for Principles for access to ICP microdata.

The agreed UN Statistical Commission Principles were:

Principle 1: It is appropriate for microdata collected for official statistical purposes to be used for statistical analysis to support research as long as confidentiality is protected.

Principle 2: Microdata should only be made available for statistical purposes.

Principle 3: Provision of microdata should be consistent with legal and other necessary arrangements that ensure that confidentiality of the released microdata is protected.

Principle 4: The procedures for researcher access to microdata, as well as the uses and users of microdata should be transparent, and publicly available.

Putting this into the context of the ICP, these Principles might read as follows.

1. It is appropriate for microdata collected for the ICP to be used for statistical analysis to support research as long as confidentiality is protected.
2. The microdata should only be made available for statistical purposes.
3. Provision of microdata should be consistent with the confidentiality laws of the country.

An equivalent to Principle 4 is not required as it is covered by General Principle 3.

It should be noted that Principle 1 does not imply that National Statistical Offices should release microdata. There may be other considerations such as quality that need to be taken into account before deciding whether to release or not.

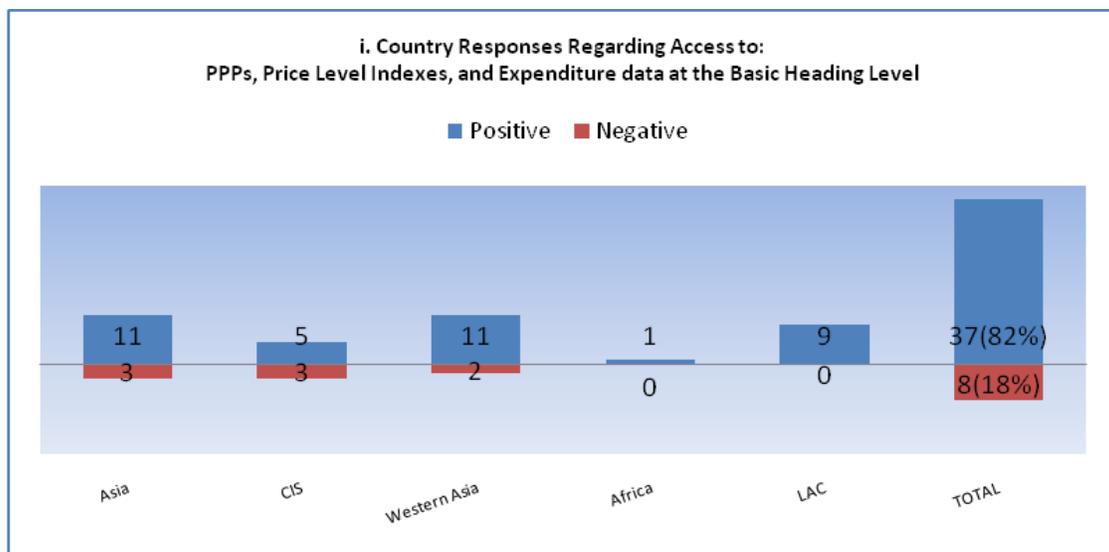
Statistical analysis should be interpreted broadly. It includes the compilation of statistical aggregations of various forms, fitting of statistical models, research into statistical methods and analyses of statistical differences between sub-populations.

#### ANNEX 4: RESULTS FROM COUNTRY RESPONSES TO THE DATA ACCESS QUESTIONNAIRE

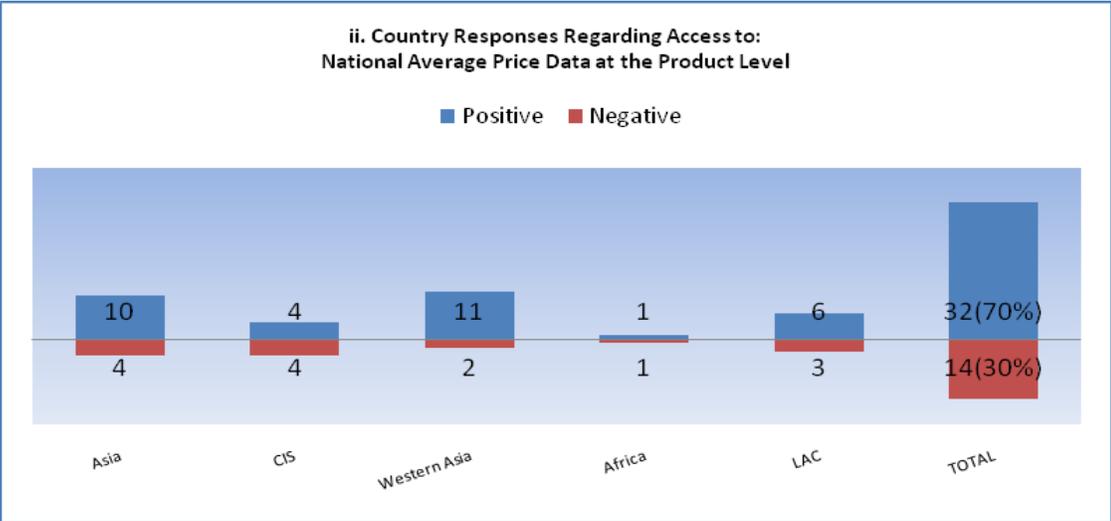
The Global Office, with the help of Regional Coordinating Agencies surveyed country positions with regards to data access. This was done through a questionnaire that was translated into various languages. The questionnaire was sent to countries in Africa, Asia, Latin America, CIS, and Western Asia.

Until date, forty six countries responded to the questionnaire. The results from the country responses are summarized below:

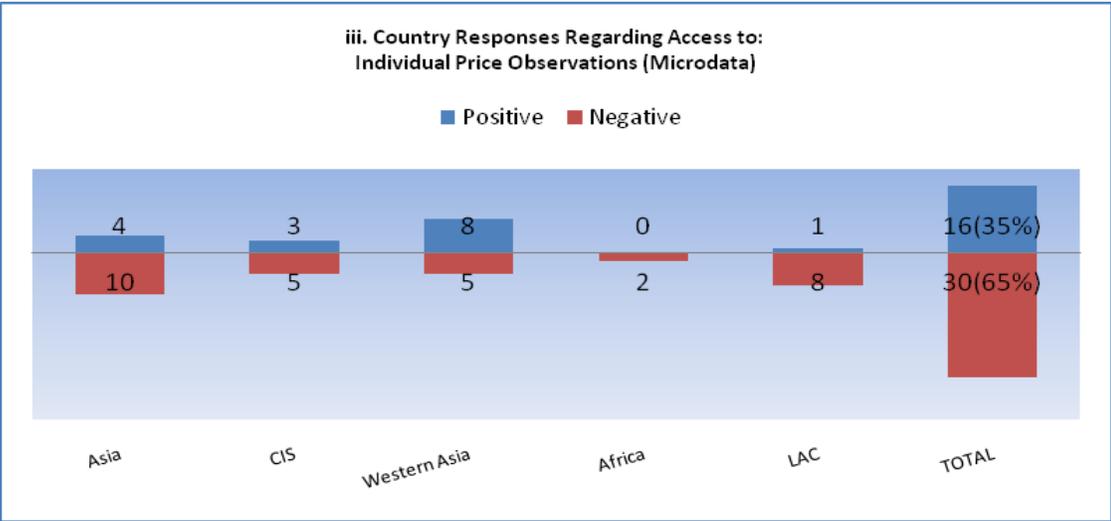
- PPPs, Price Level Indexes, and Expenditure data at the Basic Heading Level: The vast majority of countries (82%) support the release with Basic Heading level under reasonable conditions such as those outlined in the draft policy on data access.



- National Average Price Data at the Product Level: The majority of countries (70%) support the release of this data. Not surprisingly, this is lower than those supporting the release of data at the Basic Heading Level but it is still a sizeable majority.



- Individual Price Observations (microdata): Not surprisingly, the majority of countries (65%) cannot provide microdata for legal and other reasons.



**ANNEX 5: TEMPLATE FOR REQUEST FOR DATA ACCESS AND DECLARATION OF CONFIDENTIALITY**

Date:

From:

To: Director

Development Data Group

The World Bank

1818 H Street, NW

Washington, DC 20433

**Subject: Request for access to unpublished results (unpublished categories and basic heading PPPs and expenditures and/or national average prices) from the 2011 International Comparison Program.**

The following sections outline information, which must accompany any request for access to unpublished data and define the limitations to the use of those data and results extracted from the data.

**A. Written Proposal**

The researcher(s) should submit a written proposal to the ICP Global Office, World Bank that includes the following information:

- a. Data and countries or regions for which access is requested.
- b. Detailed project description including:
  - The context of the research proposal;
  - Whether data are required for the unpublished categories and basic headings or national average prices at the product level or both and the reasons if the national average prices are required;
  - The state of the literature on the topic;

- Statement of expected benefits (if any) to ICP global comparisons; and/or ICP regional comparisons,
  - Expected results, intended use, and dissemination plans.
- c. Research team:
- Brief curriculum vitae of the researchers involved in the research; and
  - Leader of the research team.
- d. Quality Limitations:
- A statement that they have discussed the quality limitations of the data with the ICP Global Office and agree that the data are of sufficient quality to satisfy the research proposal.

**B. Declaration of Confidentiality:**

The responsible senior management of the organization sponsoring the researcher(s) must include a declaration of confidentiality stating that:

- The ICP data will be used only for statistical analysis purposes, and the results of the research will either be published or made available upon request to the ICP Global Office.
- The ICP data will never be disseminated in the public domain in any form and will not be provided to third parties.
- The results of the research will not be published only at a more detailed level than is available to the public through the ICP web site unless permission has been granted to publish in greater detail. For that purpose, the version of the research to be published should be provided to the World Bank prior to publication to verify compliance.
- Copy of the research report and other feedback will be provided to the ICP Global Office.
- The calculations and documentation underlying the research will be provided to the World Bank if requested.
- Non-compliance to these rules will exclude the requesting researchers or their organization from any further access to detailed information. Furthermore, legal action may be taken to address the non-compliance.