6th Meeting of the ICP Inter-Agency Coordination Group

September 24-26, 2018

World Bank, 1818 H St. NW, Washington, DC

MC 10-100

INTERNATIONAL COMPARISON PROGRAM

DATA ACCESS AND ARCHIVING POLICY

DISCUSSION DOCUMENT

August 2017

CONTENTS

I. Introduction 3

II. Data Access Objectives 4

III. Guiding Principles 5

IV. Procedures for Data Archiving 7

V. Procedures for Data Access 7

Annex 1. ICP Classification of Final Expenditure on GDP 9

Annex 2. Proposed ICP Publication Level 10

Annex 3. Proposed ICP Data Archiving and Access Process 12

Annex 4. ICP Datasets 13

Annex 5. ICP Metadata 15

Annex 6. Application for Data Access and Statement of Confidentiality 17

I. INTRODUCTION

The International Comparison Program (ICP) is a worldwide statistical initiative to estimate purchasing power parities (PPPs) of the world's economies, and to provide comparable price and volume measures of gross domestic product (GDP) and its expenditure aggregates among countries within and across regions. The ICP entails large price data collection and detailed national accounts expenditure data compilation for participating countries.

In addition to the general uses of PPPs and related measures, the collected data on prices and expenditures can support various analytical uses of policy makers and researchers from international, regional and national agencies, as well as academic and research institutions. Hence, the underlying datasets generated by the ICP have the potential to be an important contributor to the overall value of the program.

In the 2005 ICP round, the value of the detailed price and expenditure data was left mostly untapped as the Data Archive and Access Policy strongly limited access to unpublished data. By the end of the 2005 round, it was clear that the user community was pressing for greater access to more detailed data in subsequent ICP rounds.

In response to mounting user demands for more detailed data, the 2011 ICP Executive Board endorsed the 2011 ICP Data Access and Archiving Policy, which improved access to more detailed data while respecting confidentiality constraints and data quality limitations. To increase the quality and utility of the data collected, the 2011 ICP round also focused on collecting, archiving, and providing access to metadata. Moreover, throughout the course of the 2011 ICP implementation, calls for greater government transparency and accountability through the open data movement gained traction. These open data initiatives aimed at improving the availability and reuse of data.

Recognizing that inadequate access to underlying data inhibits research, policy development, and restricts feedback about quality, the United Nations Friends of the Chair (FOC) evaluation of the 2011 ICP noted the need "to reach out and demonstrate the value of ICP data to policymakers and other important users and donors, in particular showcasing to donors that ICP generally responds to user needs and strives for further openness with regard to access to data and metadata". The evaluation concluded by recommending that "the Governing Board establish policies that improve openness with regard to access to ICP data and metadata".

This document responds to the FOC recommendation by outlining an enhanced ICP Data Access and Archiving Policy. Elements of this enhanced policy include the objectives, guiding principles, and procedures for data archiving and access. The annexes provide additional information on the ICP classification, proposed publication level, summary of the proposed data access and archiving process, ICP datasets and accompanying metadata, as well as the application for data access.

II. Data Access Objectives

The long-standing objective of the ICP data access and archiving policy is as follows: *Data derived from the ICP should be utilized to the maximum extent possible for statistical, research, and analytical purposes*. To enhance data access per the FOC recommendation, the current overall objectives are:

> Objective 1: To provide users with more detailed results beyond what was published through the 2005 and 2011 ICP, by increasing the number of published headings.
>
> Objective 2: To expand user access to unpublished data at a finer level of detail by providing, for the first time, access to sub-national average prices at the item level.

The current ICP Classification of Final Expenditure on GDP is a hierarchy of 6 main aggregates, 28 categories, 63 groups, 126 classes and 155 basic headings, as shown in Annex 1. The publication level, for which the 2005 and 2011 ICP results were published, covered only main aggregates and categories, as well as additional aggregates, such as actual individual consumption. In total, 24 headings were published through the 2005 ICP, while 26 headings were published through the 2011 ICP.

Currently, the Statistical Office of the European Union (Eurostat) and the Organisation for Economic Co-operation and Development (OECD) are publishing their respective results at a finer level of detail, covering 61 and 50 headings respectively. The Commonwealth of Independent States (CIS) have also published their 2014 results at a similar level of detail as the OECD. To make progress on the enhanced access policy's first objective, the ICP should follow suit and extend the publication of the ICP results to include finer levels of detail, such as groups and classes, where applicable. Thus, going forward, it is proposed to extend the publication level from 26 headings in the 2011 ICP to 48 headings, as shown in Annex 2.

In addition to published results, users have been able to request access to unpublished underlying data from the 2005 and 2011 ICP rounds. In the 2005 ICP, users were able to access basic heading level data and a limited set of national average prices at item level for the 18 countries that participated in the global linking exercise[1]. The access was expanded in the 2011 ICP to include national average prices at item level for all participating countries. Going forward, to make progress on the enhanced policy's second objective, it is proposed to expand user access beyond the 2011 ICP to include sub-national average price data at the item level.

[1] Referred to as the "Ring" exercise.

The detailed ICP data access and archiving policy objectives, reflecting the proposed enhanced publication and access, are:

1. PPPs, PLIs and expenditure data for all countries are disseminated in ICP publications and through an online database at the agreed publication level of 48 headings, with supporting metadata.

2. PPPs, PLIs and expenditure data at various levels of detail below the publication level for all countries are available to users, with supporting metadata, through a formal application.

3. National and sub-national[2] average price data at the item level for items on the Global Core Lists for all countries, with supporting metadata and measures of quality, are available to users through a formal application, except when the confidentiality of respondents is jeopardized.

4. National and sub-national average price data at the item level for regional items not on the Global Core Lists for all countries, with supporting metadata and measures of quality, are available to users through a formal application, except when the confidentiality of respondents is jeopardized.

5. Individual price observations with supporting metadata are available to users where permitted by the laws of individual countries, as long as the confidentiality of respondents is protected[3].

III. Guiding Principles

The following principles guide the management of ICP data and support the objective of enhancing data access:

1. **Appropriate use**: ICP data should be made available for analytical, research, and statistical purposes. Users should not misuse the data by attempting to deduce underlying confidential data.

2. **Equality of access**: ICP data are a public good and, thus, should be made available on an equal basis to anyone who wants to use them, in the same way that most national statistical offices make data available to users.

[2] Capital city or first-level administrative divisions within a country (e.g. states, provinces, etc.).
[3] For example, through data anonymization techniques.

3. **Preservation of respondent confidentiality**: provision of ICP data should be consistent with legal and other necessary arrangements that ensure the confidentiality of respondents.

4. **Transparency**: the principles and procedures for access to ICP data, as well as the uses of these data, should be transparent and publicly available.

5. **Consistency**: the principles and procedures for data access should strive to be consistent across all regions and countries, to promote equality in the treatment of all countries.

6. **Reciprocity:** reciprocity between participating countries should be established to the maximum extent possible. All participating countries are automatically considered as approved users of ICP data. Non-participating countries are not considered as approved users of ICP data, but may apply for access to these data following the procedures stipulated in Section V.

7. **Reliability**: releases of ICP data should be accompanied by appropriate metadata, including metadata that describes the quality limitations of the data.

8. **Quality limitations**: users of ICP data should be informed of the quality limitations and they should agree that the data are still useful for their purposes.

9. **Serviceability**: the ICP data should be archived to ensure it can be used to service future approved requests for access to data, it is available for possible use in future ICP comparisons, and it is available as a back-up in case these data are lost through disaster or other reasons by a region or a country.

10. **Disclosure limitations**: users accessing unpublished ICP data should not disseminate these data.

11. **Promotion of uses**: to promote the use of ICP data, users are required to share their research findings with all stakeholders, consistent with the disclosure limitations in this policy.

12. **Limitations on users' findings:** indicators computed by users based on ICP data are not considered part of the official results of the ICP.

13. **Ease of access**: data access procedures should ensure a simple and expedited process for access to ICP data.

IV. Procedures for Data Archiving

This section describes the procedures for archiving the ICP data and metadata listed in Annexes 4 and 5.

1. PPPs, PLIs and expenditure data for all countries are disseminated at the agreed publication level of 48 headings, with supporting metadata. They are archived by the World Bank Global ICP Unit and by the relevant Regional Implementing Agencies.

2. PPPs, PLIs and expenditure data at various levels of detail below the publication level for all countries, with supporting metadata, are archived by the World Bank Global ICP Unit and by the relevant Regional Implementing Agencies.

3. National and sub-national average price data at the item level for items on the Global Core Lists for all countries, with supporting metadata and measures of quality, are archived by the World Bank Global ICP Unit and by the relevant Regional Implementing Agencies.

4. National and sub-national average price data at the item level for regional items not on the Global Core Lists for all countries, with supporting metadata and measures of quality, are archived by the World Bank Global ICP Unit and by the relevant Regional Implementing Agencies.

5. Participating countries are responsible for archiving their individual price observations with supporting metadata. However, some countries may request the World Bank Global ICP Unit and/or relevant Regional Implementing Agencies to archive the individual price observations because they do not have their own facilities to archive these data. If a country asks the World Bank Global ICP Unit and/or relevant Regional Implementing Agencies to archive these data, the country in question can chose to have the data encrypted and hold the encryption key.

The unpublished data archived by the World Bank Global ICP Unit will be treated with confidentiality. The data will be archived in a secure database with limited access rights and administered by a designated data custodian. Access to the data (or any portions of the data) will be subject to the procedures specified in Section V.

V. Procedures for Data Access

This section describes the procedures for data access for the ICP data and metadata listed in Annexes 4 and 5.

1. PPPs, PLIs and expenditure data for all countries at the agreed publication level of 48 headings, with supporting metadata, will be disseminated in ICP publications and through an online database.

2. PPPs, PLIs and expenditure data at various levels of detail below the publication level for all countries, with supporting metadata, can be accessed by users through a formal application addressed to the World Bank Global ICP Unit.

3. National and sub-national average price data at the item level for items on the Global Core Lists for all countries, with supporting metadata and measures of quality, can be accessed by users through a formal application addressed to the World Bank Global ICP Unit, consistent with the confidentiality laws and processes of participating countries. Countries should inform the relevant Regional Implementing Agencies, who in turn will inform the World Bank Global ICP Unit, which information are considered confidential, and hence cannot be shared. Item names and descriptions will be anonymized before sharing with users.

4. National and sub-national average price data at the item level for regional items not on the Global Core Lists for all countries, with supporting metadata and measures of quality, can be accessed by users through a formal application addressed to the World Bank Global ICP Unit, consistent with the confidentiality laws and processes of participating countries. Countries should inform the relevant Regional Implementing Agencies, who in turn will inform the World Bank Global ICP Unit, which information are considered confidential, and hence cannot be shared. Item names and descriptions will be anonymized before sharing with users.

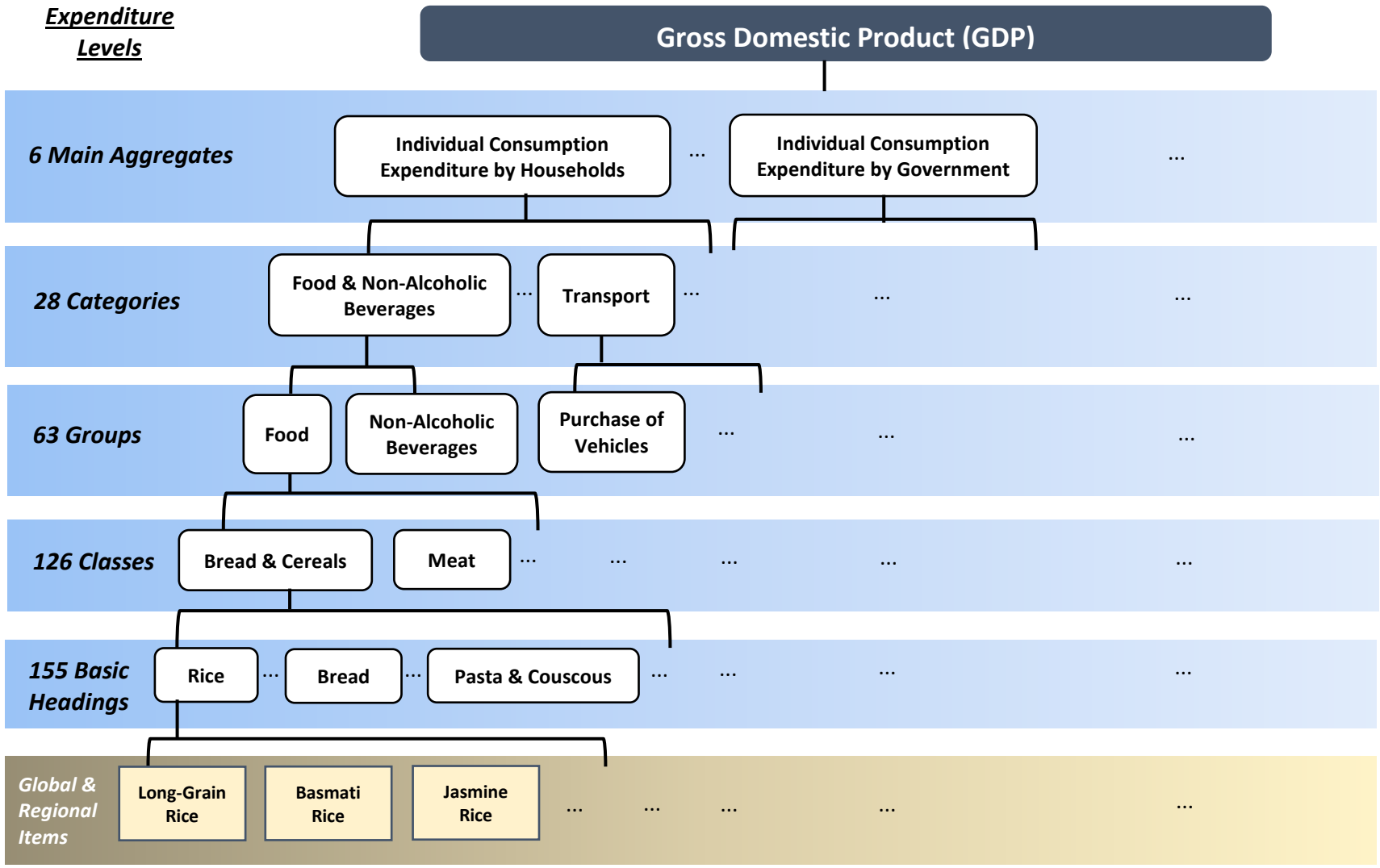
5. Access to individual price observations with supporting metadata is restricted in general. However, users may submit formal application(s) of access to these data directly to the countries in question.

Applying for access to ICP datasets listed in 2, 3 and 4 above is initiated by the user(s) with a formal application addressed to the World Bank Global ICP Unit. An outline of the information required in this application is shown in Annex 6. The World Bank Global ICP Unit makes decisions on whether to approve requests, in line with the access policy approved by the ICP Governing Board and outlined in this document.

Alternatively, users may approach relevant Regional Implementing Agencies for access to regional data (regional PPPs, PLIs, expenditure data, and average prices). In this case, Regional Implementing Agencies follow the access policy agreed by the ICP Governing Board. This access to regional data does not require clearance from the World Bank Global ICP Unit.

Annex 3 provides a summary of the data access and archiving process for ICP data and metadata.

ANNEX 1. ICP CLASSIFICATION OF FINAL EXPENDITURE ON GDP[4]



[4] The full ICP Classification of Final Expenditure on GDP can be found at: http://www.worldbank.org/en/programs/icp/brief/methodology-national-accounts

ANNEX 2. PROPOSED ICP PUBLICATION LEVEL

Proposed Publication Level		ICP 2011 Publication Level
Heading Name	Heading Level	
GROSS DOMESTIC PRODUCT	**GDP**	Yes
DOMESTIC ABSORPTION	**Additional Aggregate**	Yes
FINAL CONSUMPTION EXPENDITURE	**Additional Aggregate**	
ACTUAL INDIVIDUAL CONSUMPTION	**Main Aggregate (Actual)**	Yes
FOOD AND NON-ALCOHOLIC BEVERAGES	Category	Yes
FOOD	Group	
Bread and cereals	Class	
Meat	Class	
Fish and seafood	Class	
Milk, cheese and eggs	Class	
Oils and fats	Class	
Fruits, vegetables, potatoes	Class (Merged)	
Other food	Class (Merged)	
NON-ALCOHOLIC BEVERAGES	Group	
ALCOHOLIC BEVERAGES, TOBACCO AND NARCOTICS	Category	Yes
ALCOHOLIC BEVERAGES	Group	
TOBACCO	Group	
CLOTHING AND FOOTWEAR	Category	Yes
HOUSING, WATER, ELECTRICITY, GAS AND OTHER FUELS	Category (Actual)	Yes
FURNISHINGS, HOUSEHOLD EQUIPMENT AND ROUTINE HOUSEHOLD MAINTENANCE	Category	Yes
HEALTH	Category (Actual)	Yes
TRANSPORT	Category	Yes
PURCHASE OF VEHICLES	Group	
TRANSPORT SERVICES	Group	
COMMUNICATION	Category	Yes

Proposed Publication Level		ICP 2011 Publication Level
Heading Name	Heading Level	
RECREATION AND CULTURE	Category (Actual)	Yes
EDUCATION	Category (Actual)	Yes
RESTAURANTS AND HOTELS	Category	Yes
MISCELLANEOUS GOODS AND SERVICES	Category (Actual)	Yes
NET PURCHASES ABROAD	Category	Yes
INDIVIDUAL CONSUMPTION EXPENDITURE BY HOUSEHOLDS	**Main Aggregate**	Yes
INDIVIDUAL CONSUMPTION EXPENDITURE BY HOUSEHOLDS WITHOUT HOUSING	**Additional Aggregate**	Yes
GOVERNMENT FINAL CONSUMPTION EXPENDITURE	**Additional Aggregate**	
INDIVIDUAL CONSUMPTION EXPENDITURE BY GOVERNMENT	**Main Aggregate**	Yes
COLLECTIVE CONSUMPTION EXPENDITURE BY GOVERNMENT	**Main Aggregate**	Yes
GROSS CAPITAL FORMATION	**Main Aggregate**	
GROSS FIXED CAPITAL FORMATION	Category	Yes
MACHINERY AND EQUIPMENT	Group	Yes
CONSTRUCTION	Group	Yes
OTHER PRODUCTS	Group	Yes
CHANGES IN INVENTORIES AND VALUABLES	Category (Merged)	Yes
BALANCE OF EXPORTS AND IMPORTS	**Main Aggregate**	Yes
TOTAL GOODS	**Additional Aggregate**	
CONSUMER GOODS	**Additional Aggregate**	
CAPITAL GOODS	**Additional Aggregate**	
TOTAL SERVICES	**Additional Aggregate**	
CONSUMER SERVICES	**Additional Aggregate**	
GOVERNMENT SERVICES	**Additional Aggregate**	
TOTAL PUBLISHED HEADINGS	**48**	**26**

Annex 3. Proposed ICP Data Archiving and Access Process

	1	2	3	4	5
Datasets	PPPs, PLIs and expenditure data at the publication level for all countries, with supporting metadata	PPPs, PLIs and expenditure data at various levels of detail below the publication level for all countries, with supporting metadata	National and sub-national average price data for items on the Global Core Lists for all countries, with supporting metadata and measures of quality	National and sub-national average price data for regional items not on the Global Core Lists for all countries, with supporting metadata and measures of quality	Individual price observations, with supporting metadata
Access	Data will be released in ICP publications and though an online database	Data can be accessed through a formal application addressed to the World Bank Global ICP Unit	Data can be accessed through a formal application addressed to the World Bank Global ICP Unit	Data can be accessed through a formal application addressed to the World Bank Global ICP Unit or Regional Implementing Agencies	Restricted in general; however, users can address applications of access directly to the countries in question
Archiving	Archived by the World Bank Global ICP Unit and Regional Implementing Agencies	Archived by the World Bank Global ICP Unit and Regional Implementing Agencies	Archived by the World Bank Global ICP Unit and Regional Implementing Agencies	Archived by the World Bank Global ICP Unit and Regional Implementing Agencies	Archived by the countries in question

Degree of Detail

ANNEX 4. ICP DATASETS

A. Global Data

1. Global PPPs, PLIs and expenditure data disseminated at the publication level for all countries.

2. Global PPPs, PLIs and expenditure data at various levels of detail below the publication level for all countries.

3. National annual average prices for items on the Global Core Lists[5] for all countries.

4. Disaggregated sub-national annual average prices for items on the Global Core Lists for all countries.

5. Population and exchange rate data for all countries.

B. Regional Data

1. Regional PPPs, PLIs and expenditure data disseminated at the publication level for countries in each region.

2. Regional PPPs, PLIs and expenditure data at various levels of detail below the publication level for countries in each region.

3. National annual average prices for items on the Regional Lists[6] for countries in each region.

4. Disaggregated sub-national annual average prices for items on the Regional Lists for countries in each region.

5. Population and exchange rate data for countries in each region.

C. Country Data

1. National annual average prices for items priced from the Regional Lists and Global Core Lists.

2. Disaggregated sub-national annual average prices for items priced from the Regional Lists and Global Core Lists.

---

[5] These include the Global Core Lists for consumer goods and services, housing, government, machinery and equipment, and construction.

[6] These include the Regional Lists for consumer goods and services, housing, government, machinery and equipment, and construction. These lists include items from the Global Core Lists selected for pricing in each region.

3. Individual price observations at the item level (microdata) for items priced from the Regional Lists and Global Core Lists with identifiers for the type of outlet, location, and urban/rural designation.

4. Population and exchange rate data.

ANNEX 5. ICP METADATA

A. At the global level

1. A broad description of ICP methodology, including a description of methodologies for consumer goods and services, housing, government, machinery and equipment, and construction, as well as the methodology for linking regional results and imputing PPPs for non-participating economies.

2. Global Core Lists, with supporting item descriptions.

3. The matrix of availability of price data for each item on the Global Core Lists for all countries.

4. The matrix of importance for consumer goods and services on the Global Core Lists for all countries.

5. Measures of quality including the number of individual price observations and coefficients of variation for Global Core List items for all countries.

6. A table showing the reference PPPs used to impute global PPPs.

7. National accounts reporting forms and related country practice questionnaires for all countries.

8. Survey framework questionnaires for all countries.

B. At the regional level

1. A broad description of ICP methodology implemented in each region, including the method implemented for housing.

2. Regional Lists, with supporting item descriptions.

3. The matrix of availability of price data for each item on the Regional Lists for all countries.

4. The matrix of importance for consumer goods and services on the Regional Lists for all countries.

5. Measures of quality including the number or individual price observations and coefficients of variation for items on the Regional Lists for countries in each region.

6. A table showing the reference PPPs used to impute regional PPPs.

7. National accounts reporting forms and related country practice questionnaires for countries in each region.

8. Survey framework questionnaires for countries in each region.

C. At the national level

1. Detailed national survey framework for the ICP.

2. A table showing items for which prices were collected and quarters in which prices were collected.

3. Estimation method for national average prices, especially when the price survey was not spatially representative.

4. Price indices used to adjust prices collected temporally, when the price survey did not cover the whole reference year.

5. Methodology for deriving expenditure data.

ANNEX 6. APPLICATION FOR DATA ACCESS AND STATEMENT OF CONFIDENTIALITY

Date:

From:

To: The World Bank Global ICP Unit

Subject: Request for access to unpublished data and metadata from the International Comparison Program

The following sections outline information which must accompany any request for access to **unpublished data and metadata** and define the limitations to their use.

A. Written proposal. Users must submit a written proposal to the World Bank Global ICP Unit that includes the following information:

1. Data, countries, and metadata for which access is requested.
2. Detailed project description including: the context of the research proposal; topic JEL classification code(s) of the planned research[7]; research plan; state of the literature on the topic; intended use; expected results; dissemination plans; and statement of expected benefits to the ICP.
3. Brief curriculum vitae of all the researchers involved in the research.

B. Declaration of use. The responsible senior manager, head of department or dean of the organization sponsoring the researcher must include a signed declaration stating that:

1. Users will not misuse the data by attempting to deduce underlying confidential data.
2. Users were informed of the quality limitations and they agree that the data are still useful for their purposes.
3. Users accessing unpublished ICP data should not disseminate these data.

[7] Users must pick and provide up to 4 three-character classification codes relevant to their planned research from the Journal of Economic Literature (JEL) classification system. The JEL classification system is a standard method of classifying literature in the field of economics. The classification, corresponding codes, and examples are available at: https://www.aeaweb.org/econlit/jelCodes.php.

4. To promote the use of ICP data, users will share their research findings with the World Bank Global ICP Unit.

5. The calculations and documentation underlying the research will be provided to the World Bank Global ICP Unit, if requested.

6. Users are aware that indicators computed by them based on ICP data are not considered part of the official results of the ICP.

7. Noncompliance with this declaration will exclude the requesting organization from any further access to unpublished ICP data.