

## Validation Tables

The main price data analysis at the regional and global levels of validation is carried out using two validation tables: the Quaranta table, named after Vincenzo Quaranta, who first proposed the table for use in the European purchasing power parities (PPP) program in 1990, and the Dikhanov table, named after Yuri Dikhanov, who first proposed the table for use during the 2005 round of the International Comparison Program (ICP). The purpose of both tables is to screen the national average prices for possible errors by comparing the average prices of the same items across countries. Both tables provide similar measures of price variation for basic headings (BHs), countries, and items.

The main difference between the two tables is that the Quaranta table is employed to edit prices within basic headings, whereas the Dikhanov table can also be used to edit prices within aggregates. The Dikhanov table can be set as well to show only key indexes while hiding other details. When presented in this compact form, the Dikhanov table is better suited for editing prices across the basic headings and the items composing an aggregate.

The Dikhanov table is specific to the country product dummy (CPD) or country product representative dummy (CPRD) method of calculating PPPs, whereas the Quaranta table has a broader application that includes the Éltető-Köves-Szulc (EKS) and EKS\* methods, as well as the CPD, CPRD, and weighted CPD (CPD-W) methods.

This chapter describes the characteristics of the Quaranta and Dikhanov tables. Chapter 15 on validation of the household consumption survey discusses the use of these tables during the validation process. And chapter 23, dealing with elementary aggregation using the CPD method, describes the PPP calculation methods employed in the Quaranta and Dikhanov tables.

### QUARANTA TABLE

The Quaranta table consists of a set of tables for basic headings—one for each basic heading as a whole and one for each item within the basic heading. For comparison of average prices, the Quaranta table provides three main measures: exchange rate (XR)-ratio, purchasing power parity (PPP)-ratio, and price level index (PLI). For analysis of price variation, the Quaranta table provides four measures, ranging from variation of individual price observations to variation of items within a basic heading. Each of these indexes is discussed in detail in the rest of this section.

### Average Price Measures

Once converted to a common currency, the average prices of different countries for *the same item* can be compared and extreme values

can be identified. However, prices cannot be compared *across items* directly, even when expressed in the same currency, but the price ratios of countries pricing an item can be compared with the equivalent price ratios for other items once they have been "standardized."

The standardized price ratios for an item are the ratios of the individual average prices in countries pricing the item to the geometric mean of the average prices in all countries pricing the item, when the average prices are expressed in a common currency. Both exchange rates and PPPs are used in validation to convert the average prices to a common currency, and both the exchange rate-converted average prices and the PPP-converted average prices are used to derive the standardized price ratios.

The standardized price ratios<sup>1</sup> based on exchange rate-converted prices are called XR-ratios, and the standardized price ratios based on PPP-converted prices are called PPP-ratios.<sup>2</sup>

The third measure used to compare prices is the PLI. This index is defined as the ratio of the basic heading PPP to the exchange rate and is expressed as a percentage. A PLI greater (less) than 100 indicates that when the national average prices are converted at exchange rates, the resulting prices within the basic heading tend to be higher (lower) on average than the prices in the base country of the group.

### Measures of Price Variation

The Quaranta table provides four coefficients of variation:

- *Overall average coefficient of variation*—measures dispersion among all the PPP-ratios for a basic heading. In doing so, it measures the homogeneity of the price structures of the countries covered by the basic heading and the reliability of the PPPs calculated for the basic heading. Naturally, the higher the value of the coefficient, the less homogeneous will be the price structures, resulting in less reliable PPPs.
- *Country coefficient of variation*—measures dispersion among a country's PPP-ratios for a basic heading. In other words, it measures the variation in a country's price levels

among the items in a basic heading and the reliability of its PPP for the basic heading. The higher the coefficient's value, the less uniform will be the country's price levels, leading to less reliable PPPs.

- *Item coefficient of variation*—measures dispersion among the PPP-ratios for an item. It is an indicator of comparability and accuracy that addresses the question of whether comparable products have been priced for an item. The higher the coefficient's value, the less uniform will be the item price levels. Low uniformity would raise questions about the comparability and accuracy of the item's pricing across countries.
- *Price observation coefficient of variation*—measures variation in the price observations on which the average price reported for an item by a country is based. It is taken straight from the average price table and is used to identify extreme values among average prices during the country-level validation.

Besides serving as editing tools, the coefficients provide a means of monitoring progress during the validation stage and, at its conclusion, of assessing the effectiveness of the entire process of editing and verification in reducing the incidence of nonsampling errors among the price data. Coefficients should be significantly smaller at the end of validation than they were at the beginning.

### Description of the Quaranta Table

Table 14.1 is an example of a Quaranta table. The numbers in italics have been added for ease of reference. Explanatory notes follow the table. The table has four sections:

1. *Data selection criteria* provides general details about the table such as the run date [3], averaging method [4], and imputation method [5].
2. *Summary information* gives information that relates to the basic heading as a whole such as the number of items [6] and countries [8] included in the analysis, as well as the average weight of the basic heading in the total expenditure [7] and the average coefficient of variation [9].

**Table 14.1** Example of a Quaranta Table, ICP 2011

QUARANTA TABLE DIAGNOSTICS—Rice								
Data selection criteria								
<b>[1] Basic heading code</b>	110111.1	<b>[2] Time period</b>	Q1, 2011	<b>[3] Run date</b>	3/29/2012			
<b>[4] Averaging method</b>	Arithmetic mean	<b>[5] Imputation method</b>	CPD					
Summary information								
<b>[6] No. of items included in the analysis</b>	11 out of 11	<b>[7] Average weight of basic heading in total expenditure</b>	0.0					
<b>[8] No. of countries included in the analysis</b>	5 out of 5	<b>[9] Average coefficient of variation</b>	30.9					
<b>[10] Base country</b>	Country							
Country-level details								
* Shares are multiplied by 10,000.								
<b>[11]</b>	<b>[12]</b>	<b>[13]</b>	<b>[14]</b>	<b>[15]</b>	<b>[16]</b>	<b>[17]</b>		
<b>Country</b>	<b>XR</b>	<b>PPP</b>	<b>PLI (%)</b>	<b>Weight*</b>	<b>Items</b>	<b>Var. co.</b>		
Country 1	3,104.03	2,167.5700	6983.10%	0.0	8;*5	28.6		
Country 2	15.38	9.9332	64.60%	0.0	6;*2	41.4		
Country 3	1.00	1.0000	100.00%	0.0	7;*7	19.5		
Country 4	9.27	10.6217	114.58%	0.0	11;*5	31.9		
Country 5	9.49	8.5859	90.45%	0.0	8;*3	33.1		
Item-level details								
<b>[18]</b>	<b>[19]</b>				<b>[20]</b>	<b>[21]</b>		
110111.101	Long grain rice, parboiled				Var. co.:	34.2	1 kilogram	
<b>[22]</b>	<b>[23]</b>	<b>[24]</b>	<b>[25]</b>	<b>[26]</b>	<b>[27]</b>	<b>[28]</b>	<b>[29]</b>	<b>[30]</b>
<b>Country</b>	<b>NC-price</b>	<b>Quotations</b>	<b>Var. co.</b>	<b>XR-pr</b>	<b>XR-ratio</b>	<b>PPP price</b>	<b>PPP-ratio</b>	<b>Pref. UoM</b>
Country 1	—	—	—	—	—	—	—	1 kilogram
Country 2	18.000	*14	6.0	1.17	108.19	1.81	151.49	1 kilogram
Country 3	1.309	*17	17.5	1.31	120.94	1.31	109.40	1 kilogram
Country 4	8.605	*19	26.2	0.93	85.79	0.81	67.73	1 kilogram
Country 5	9.150	5	1.5	0.96	89.08	1.07	89.09	1 kilogram
<b>Geometric mean</b>			<b>[31]</b>	<b>[32]</b>				
			1.08	1.20				

QUARANTA TABLE DIAGNOSTICS		
<b>[1]</b>	Basic heading code	Code for basic heading covered by table.
<b>[2]</b>	Time period	Period during which prices for products covered by table were collected.
<b>[3]</b>	Run date	Date table was computed.
<b>[4]</b>	Averaging method	Method used to calculate average values in table.
<b>[5]</b>	Imputation method	Method used to calculate basic heading PPPs in column <b>[13]</b> . Currently, it is the country product dummy (CPD), but it could also be the country product representative dummy (CPRD), weighted CPD (CPD-W), Éltető-Köves-Szulc (EKS), or EKS*.

table continues next page

**Table 14.1** (Continued)

Summary information		
[6]	No. of items	Number of items specified and included in basic heading analysis.
[7]	Average weight	Average expenditure weight for group of countries covered by basic heading. Unweighted arithmetic mean of the national weights in column [15]. Like the national weights, it is scaled to 10,000.
[8]	No. of countries	Total number of countries included in basic heading analysis.
[9]	Average coefficient of variation (var. co. 1)	Overall average coefficient of variation or, more precisely, average item coefficient of variation for items priced for basic heading. It is calculated as unweighted arithmetic mean of product coefficients of variation at [20]. It measures average variation of PPP-ratios in column [29] of all products priced for basic heading.
[10]	Base country	Country and respective currency selected as numéraire. Any of the countries and respective currencies included in the analysis could be chosen as numéraire.
Country-level details		
[11]	Country	Names of countries covered by table.
[12]	XR	Market exchange rates (XRs) of countries expressed as number of units of national currency (NC) per unit of numéraire currency specified in [10].
[13]	PPP	Purchasing power parities (PPPs) for basic heading calculated as specified in [5] and expressed as number of units of national currency per unit of selected numéraire currency specified in [10]. Prices used to calculate PPPs are average prices in national currencies that countries report for products they priced for basic heading—that is, the NC-prices in column [23].
[14]	PLI	Price level index (PLI). PPPs in column [13] expressed as percentage of the corresponding exchange rate in column [12].
[15]	Weight	National expenditure weights multiplied by 10,000. That part of a country's GDP that is spent on basic heading when both expenditures are expressed in national currency and valued at national price levels.
[16]	No. of items	Number of items priced by each country for basic heading. Number with asterisk (*) is number of important items.
[17]	Var. co. 2	Country coefficient of variation. Standard deviation of country's PPP-ratios in column [29] for all products priced by country for basic heading, expressed as percentage of arithmetic mean of country's PPP-ratios in column [29] for all products priced by country for basic heading.
Item-level details		
[18], [19]	Item code and name	Code, name, and summary definition of product covered in subsequent product section.
[20]	Var. co. 3	Item coefficient of variation. Standard deviation of item's PPP-ratios in column [29] expressed as percentage of arithmetic mean of product's PPP-ratios in column [29].
[21]	Unit and quantity of measurement	Unit and quantity to which price observations are converted.
[22]	Country	Names of countries covered by table.
[23]	NC-price	Average price for product in national currency.
[24]	Quotations	Number of price observations on which average prices in national currency in column [23] are based. Asterisk (*) indicates whether the item is important for given country.
[25]	Var. co. 4	Price observation coefficient of variation. Standard deviation of price observations underlying product's average price in column [23] expressed as percentage of arithmetic mean of price observations underlying product's average price in column [23].
[26]	XR-price	Average prices in national currency in column [23] converted to numéraire currency with exchange rates in column [12].
[27]	XR-ratio	Standardized price ratios based on exchange rate—converted prices in column [26]. XR-prices are expressed as percentage of their geometric mean at [31].
[28]	PPP price	Average prices in national currency in column [23] converted to numéraire currency with PPPs in column [13].

**Table 14.1** (Continued)

Item-level details		
[29]	PPP-ratio	Standardized price ratios based on PPP-converted prices in column [28]. PPP prices are expressed as percentage of their geometric mean at [32].
[30]	Pref. UoM	Preferred unit of measurement (UoM) for item.
[31]	Geometric mean	Geometric mean of exchange rate-converted prices in column [26]. Use of geometric mean here and in [32] ensures invariance with respect to choice of numéraire.
[32]	Geometric mean	Geometric mean of PPP-converted prices in column [28].

Source: ICP, <http://icp.worldbank.org/>.

Note: — = not available.

3. *Country-level details* gives key indexes that relate to individual countries at the basic heading level. Included are PPPs [13], PLIs [14], number of priced items [16], and country coefficients of variation [17].

4. *Item-level details* covers the items priced for the basic heading. Each item has its own table that shows the item coefficient of variation [20]; the average prices reported by countries in national currencies [23]; the average prices converted to a common currency with their exchange rates [26], their geometric mean [31], and their XR-ratios [27]; and the average prices converted to a common currency with the PPPs for the basic heading—that is, PPP prices [28], their geometric mean [32], and their PPP-ratios [29].

### Use of XR-Ratios and PPP-Ratios

The XR-ratios and PPP price ratios provide valuable information for screening the national average prices. Each of these ratios refers to a particular item in a particular country. A high (low) XR-ratio or PPP-ratio means that the national average price for the item in question is high (low) compared with the prices of the same product in other countries when converted into a common numéraire currency using the exchange rate or basic heading PPP. When the same logic is used to screen individual price observations within a country, there comes a point at which the XR or PPP price is high (low) enough to raise the question of whether the price may be erroneous. Experience suggests that the appropriate threshold for the individual XR-ratio is 50 and PPP-ratio is 150. To ease the process of detecting high (low) PPP-ratios, the color scheme presented in table 14.2 is used.

**Table 14.2** Color Scheme for PPP-Ratios, ICP 2011

PPP-ratios with values	Color code
Between 78 and 128	None
Between 47 and 78 or 128 and 212	Yellow
Between 14 and 47 or 212 and 739	Red
Less than 14 or greater than 739	Black

Source: ICP, <http://icp.worldbank.org/>.

An XR-ratio that lies outside these limits may signal a questionable observation. However, it must be remembered that the principal reason for calculating PPPs is that when the prices of a given product are converted into a common currency unit using exchange rates, they are not in fact equal in all countries. The general level of prices tends to be systematically higher or lower in some countries than in others. Thus a high or low XR price for an individual item in one country may be largely due to the fact that the general price level for that country is high or low when exchange rates are used. It may not signal any abnormality in that particular price. For this reason, XR price ratios are less useful than PPP price ratios for validation purposes.

On the other hand, PPPs are the rates of currency conversion that are designed to equalize price levels for the products covered. The PPPs for a basic heading such as fresh or chilled vegetables other than potatoes are the rates of currency conversion that should enable a given amount of currency to purchase the same basket of vegetables in all countries. Thus if the patterns of relative prices for the different items within the basic heading were similar in different countries, the PPP prices for the same item in different countries would tend to be bunched together and the PPP price ratios (i.e., the ratios of the individual PPP prices to the geometric

mean of the PPP prices for all the countries) would cluster around 100. There would be little dispersion between countries in either the PPP prices or the PPP price ratios derived from them.

Conversely, a high level of dispersion in the PPP prices of the PPP price ratios for the same item across different countries implies that the relative price of the item tends to vary a lot from country to country. This could happen in the real world, but it also could signal that one or more PPP prices are wrong. Thus the dispersion in the PPP prices or price ratios for the same item in different countries becomes a key indicator for purposes of regional level validation. It can be measured by calculating the coefficients of variation for the PPP prices or the PPP price ratios.<sup>3</sup>

In conclusion, if the coefficient of variation for the PPP prices or price ratios for the same item in different countries exceeds some predetermined threshold, the national average prices for that item become questionable and require further investigation.

### Quaranta Summary Tables

Although the basic heading Quaranta tables continue to be the main tool for in-depth price validation, a useful starting point is an overview of the price data provided by the countries. This overview should identify the most important cases for validation activities. To this end, two summary tables are added to the standard Quaranta table: (1) the PLI summary and (2) the coefficient of variation (CV) summary.

#### *Price Level Index Summary*

The PLI expresses the price level of one country relative to another by dividing the PPPs by the current nominal exchange rate. If the PLI of a country is higher than 100, the country is relatively expensive compared with the other in the comparison. Conversely, if the PLI is lower than 100, the country is relatively cheap compared with the other country. PLIs can be calculated for individual basic headings or for aggregates.

PLIs are useful in the validation process because the price levels for a certain country are expected to be relatively stable across the basic

headings. For example, if an overall (aggregate) PLI for a country is 100, a basic heading with a PLI of 300 would have to be checked for potential problems with the underlying price data. However, a high or low PLI is not an error by definition; it can be the result of a certain economic structure. Nevertheless, these cases should be carefully verified. It should also be noted that valid PLIs tend to fit within certain thresholds. Thus extremely high or low basic heading PLIs—1,000 or 0.1, for example—are clear indicators of problems with the price data for the basic heading.

The PLI summary table brings together basic heading PLIs for each country covered by the Quaranta table, and it adds certain indicators to assist in identifying problematic cases (see table 14.3 for an example of a PLI summary table).

The bottom part of the table is a matrix of basic headings and countries included in the analysis. The table presents five food basic headings and four countries. The PLIs appear as percentages for each country.

The upper part of the table gives the following information for each country:<sup>4</sup>

- (1) Rank of the country based on the geometric mean (4)
- (2) Minimum PLI
- (3) Maximum PLI
- (4) Unweighted geometric mean of all BH PLIs
- (5) Set upper limit for BH PLIs (here 2 times the geometric mean)
- (6) Set lower limit for BH PLIs (here 0.5 times the geometric mean)
- (7) Number of missing BH PLIs (thus missing PPPs)<sup>2</sup>
- (8) Number of BHs below set threshold in (5)
- (9) Number of BHs above set threshold in (6)
- (10) Total number of problematic BHs—sum of (7), (8), and (9).

In table 14.3, the geometric mean of all BH PLIs for country 1 is 149 percent, and the set upper limit for PLI variation is 297 percent. For country 1, most of the BH PLIs are close to the geometric mean, except the one for bread, for which the PLI is 300 percent. This basic heading should be flagged for further analysis to determine whether it is plausible that the price level

**Table 14.3** Example of PLI Summary Table, ICP 2011

PLI QUARANTA TABLE						
(1)	Rank	1	2	3	4	...
(2)	Minimum	115%	43%	65%	40%	...
(3)	Maximum	300%	124%	210%	107%	...
(4)	Geometric mean (GM)	149%	96%	92%	81%	...
(5)	Upper limit	297%	186%	184%	162%	...
(6)	Lower limit	74%	46%	46%	40%	...
(7)	No. of missing basic headings (BHs)	0	1	1	0	...
(8)	No. of BHs <0.5 GM	0	1	0	1	...
(9)	No. of BHs >2 GM	1	0	1	0	...
(10)	No. of problematic BHs	1	2	2	1	...
BH code	BH	Country 1	Country 2	Country 3	Country 4	...
110111.1	Rice	130%	124%	70%	107%	...
110111.2	Other cereals, flour, and other cereal products	115%	—	210%	94%	...
110111.3	Bread	300%	120%	75%	84%	...
110111.4	Other bakery products	120%	116%	—	40%	...
110111.5	Pasta products	135%	43%	65%	102%	...

Source: ICP, <http://icp.worldbank.org/>.

Note: — = not available; BH = basic heading.

for bread products in country 1 is over 2 times higher than those for the other products such as rice and other bakery products.

In this example, the upper and lower limits were set to be 2.0 and 0.5 times the geometric mean, respectively. However, these limits can be either wider or narrower; the selection depends on the quality of the price data. During the initial data validation, the focus is normally on correcting the extreme errors that make any finer data validation impossible. In this case, wide limits are normally selected. After a series of validation rounds, the limits can be set to be narrower. Gradually, after a number of rounds of verification with prices of all countries participating in the comparison, a convergence will occur, and the return on further rounds of verification will be deemed marginal and therefore not worth pursuing.

### **Coefficient of Variation Summary**

The Quaranta tables use two measures for variation for each basic heading:

- *Country coefficient of variation*—measures dispersion among a country's PPP-ratios for a

basic heading. In other words, it measures the variation in a country's price levels among the items for the basic heading and the reliability of its PPP for the basic heading. The higher the coefficient's value, the less uniform will be the country's price levels, leading to less reliable PPPs.

- *Basic heading coefficient of variation*—measures dispersion among all the PPP-ratios for a basic heading. In doing so, it measures homogeneity of the price structures of the countries covered by the basic heading and the reliability of the PPPs calculated for the basic heading. Naturally, the higher the coefficient's value, the less homogeneous will be the price structures, resulting in less reliable PPPs.

The coefficient of variation summary places these measures in matrix form and adds the average expenditure share for the group of countries covered by the basic heading. Table 14.4 is an example of a CV summary.

For a *region as whole*, the CV summary helps to identify basic headings that are most problematic based on the basic heading CV and that

**Table 14.4** Example of Coefficient of Variation Summary, ICP 2011

Coefficient of variation (CV) of PPP-ratios							
Basic heading (BH) code	Basic heading	Country 1	Country 2	Country 3	Country 4	BH CV	Expenditure share
110111.1	Rice	41.9	245.2	151.5	42.0	120.1	10.76
110453.1	Other fuels	31.5	21.1	42.8	50.1	36.4	6.69
110111.2	Other cereals, flour, and other cereal products	141.9	51.7	88.2	36.2	79.5	5.99
110117.1	Fresh or chilled vegetables other than potatoes	311.8	301.1	155.0	367.9	284.0	5.49
110115.3	Other edible oils and fats	19.1	30.1	13.7	27.2	22.5	4.92

Source: ICP, <http://icp.worldbank.org/>.

Note: Cells with CVs higher than 50 are highlighted.

also have high-expenditure shares. The basic headings are ranked row-wise based on the expenditure share. These basic headings are of utmost importance in the validation process because they have a greater impact when the BH PPPs are aggregated. Again, all problematic cases should be analyzed, but the starting point should be problematic basic headings with high expenditure shares.

For *individual countries*, the CV summary identifies the most problematic cases based on the basic heading CV. Countries are ranked column-wise from the most problematic to the least problematic, based on the geometric mean of the basic heading CVs.

In this example, CVs higher than 50 are highlighted, but, as explained earlier, the set limit depends on the phase of the validation.

### Country Diagnostic Report

The standard Quaranta table includes tables for each basic heading and each item under the basic heading. The item tables present information for all countries included in the analysis. An alternative way to present this information is to group countries, rather than items, for BH-level analysis. This grouping is called a country diagnostic report. In essence, the report compiles basic headings and item rows for a single country and presents them as an individual report. All information is gathered and calculated for the standard Quaranta table; the only difference between the standard Quaranta table and the country diagnostic report is the presentation (see table 14.5 for an example of a country diagnostic report).

The country diagnostic report has four sections, much like the standard Quaranta table. The first two sections, data selection criteria and summary information, are identical to those in the standard table. However, the two last sections, country-level details and item-level details, contain information only for the given country.

This kind of presentation has two main benefits: (1) it can serve as an effective validation tool, especially for the countries, by quickly pinpointing any potential problems with data; and (2) it allows focusing on the internal price structure and the potential systematic problems in a country's data that may go unnoticed when information is presented for all countries in the analysis.

However, identification of potentially problematic cases in the country diagnostic report is only the first step; the next step should be verification of each case in the full item context by studying the respective information and indexes for all countries pricing the item. Without this wider context, it is not possible to see the price relations of countries pricing the item and possible problems arising from it.

### DIKHANOV TABLE

The second validation table used in the context of the ICP is the Dikhanov table. As explained at the outset of this chapter, both the Dikhanov and Quaranta tables use a similar approach and concepts: studying item price deviations for each country in a two-dimensional space—that of items and countries. The Quaranta table is

**Table 14.5** Example of Country Diagnostic Report, ICP 2011

COUNTRY DIAGNOSTIC REPORT—Country								
<b>Basic heading code</b>	1101111			<b>Time period</b>	Q1, 2011	<b>Run date</b>	3/29/2012	
<b>Averaging method</b>	Arithmetic mean			<b>Imputation method</b>	CPD			
Summary information								
<b>No. of items included in the analysis</b>	11 out of 11			<b>Average weight of basic heading in total expenditure</b>	0.0			
<b>Base country</b>	Country			<b>Average coefficient of variation</b>	30.9			
Country-level details								
	Country	XR	PPP	PLI (%)	Weight*	Items	Var. co.	
	Country	1.00	1.0000	100.000	0.0	7;*7	19.5	
<b>* Shares are multiplied by 10,000.</b>								
Item-level details								
Product code	Product name	Pref. UoM	NC-price	Quotations	Var. co.	XR-ratio	PPP-ratio	
110111.101	Long grain rice—parboiled	1 kilogram	1.309	*17	17.5	120.94	109.40	
110111.102	Long grain rice—non-parboiled	1 kilogram	1.347	*14	2.8	117.09	107.99	
110111.103	Long grain rice—family pack	5 kilograms	3.031	*12	7.0	88.52	82.17	
110111.104	Jasmine rice	10 kilograms	6.639	*7	3.5	95.41	88.00	
110111.105	Basmati rice	1 kilogram	1.104	*4	53.7	104.10	89.42	
110111.106	White rice, 25% broken	1 kilogram	—	*—	—	—	—	
110111.107	White rice, medium grain	1 kilogram	—	—	—	—	—	
110111.108	Brown rice—family pack	5 kilograms	—	*—	—	—	—	
110111.109	Short grain rice	1 kilogram	—	—	—	—	—	
110111.110	Uncle Ben's rice	5 kilograms	6.313	*10	5.9	103.47	93.60	
110111.111	Thailand rice	1 kilogram	0.811	*2	4.9	150.66	139.87	

Source: ICP, <http://icp.worldbank.org/>.

Note: — = not available; XR = exchange rate; PLI = price level index; var. co. = coefficient of variation; UoM = unit of measurement; NC = national currency.

intended to serve as a diagnostic tool for prices at the basic heading level, whereas the Dikhanov table can be processed at any level, from the total gross domestic product (GDP)<sup>6</sup> down to the basic heading. The Dikhanov table can also be processed for intermediate aggregates such as goods and services. The Quaranta table shows additional information about item prices within a basic heading such as the number of quotations, the price variance and average prices, as well as the exchange rate ratios, whereas the Dikhanov table emphasizes the between-basic heading validation, adding features to detect anomalies across both countries and basic headings.

In the Quaranta table, PPPs are computed using one of the four methods: (1) EKS, (2) EKS\*,

(3) CPD or CPRD, or (4) weighted CPD. In the Dikhanov table, the CPD and CPRD methods are used in computations because the EKS method does not generate the average product price, an important measure that enters into various computations in the Dikhanov table.

### Average Price Measures

Similar to the Quaranta table, the Dikhanov table provides XR-ratios at the item level and PLIs at the aggregate or basic heading level. However, for the PPP-ratios the approach of the Dikhanov table differs from that of the Quaranta table. As explained in the previous section, the PPP-ratio is the double-normalized product price. The first normalization is to convert the

price of the product into the numéraire currency by dividing it by the basic heading PPP (this is the so-called PPP price). The second normalization is to divide the PPP price by the geometric mean of the PPP prices across all the countries.

The Dikhanov table uses CPD residuals instead of PPP-ratios as the double-normalized item price. When the CPD is used, the CPD residuals in the Dikhanov table are equal to the logarithms of the PPP-ratios in the Quaranta table. However, this identity holds only if

- The PPPs in both tables are calculated with the CPD (it will not hold if the CPD is used for the Dikhanov table but EKS for the Quaranta table).
- The items are included in the calculation as a single group—that is, either as a basic heading or as an aggregate (it will not hold if the CPD residuals for items in the basic heading in the Dikhanov table are based on PPPs for an aggregate and the PPP-ratios in the Quaranta table are derived from PPPs for the basic heading).

Table 14.6 summarizes the connection of the CPD residuals and PPP-ratios when these conditions hold.

Details on the CPD residuals are presented later in this section.

### Measures of Price Variation

The Dikhanov table uses standard deviations (STDs) instead of coefficients of variation to measure overall country and item variation. The STDs for the CPD residuals in the Dikhanov table are CVs because the mean of the residuals is 1. The two sets of coefficients are not the

same because of differences in computation. The overall CV in the Quaranta table is an average of the CVs of the items priced for the basic heading, whereas the overall CV in the Dikhanov table is computed with all the CPD residuals in the table's item section, thereby ensuring consistency among the overall CV, the item CVs, and the country CVs.

In addition, the item CVs in the Quaranta table should in theory be calculated using logarithms because the PPP-ratios are based on the geometric mean of the PPP prices, but for practical reasons they are calculated using the arithmetic mean and standard deviation of the PPP-ratios. By contrast, the item CVs in the Dikhanov table are based on CPD residuals, which are logarithms of the PPP-ratios. Despite the computational differences, these two sets of coefficients of variation are of similar orders of magnitude and reliability in terms of identifying extreme values.

### Description of the Dikhanov Table

The Dikhanov table can be presented in two versions: extended<sup>2</sup> and collapsed. Table 14.7A is an example of an extended table, and table 14.7B illustrates a collapsed table for the household final consumption expenditure (HFCE). Both tables are calculated at the basic heading level, and neither table is complete. CPD residuals are shown for only 3 (table 14.7A) or 6 (table 14.7B) of the 809 products priced and only for 5 of the 18 countries included in the comparison. The PPPs for the aggregate, HFCE, are not weighted. They have been calculated by a CPD that uses the whole set of products and their prices without taking basic heading expenditures into account.

**Table 14.6** CPD Residuals and Respective PPP-Ratios, ICP 2011

CPD residuals with values	PPP-ratio equivalence
Between -0.25 and 0.25	Between 78 and 128
Between -0.75 and -0.25 or 0.25 and 0.75	Between 47 and 78 or 128 and 212
Between -2.0 and -0.75 or 0.75 and 2.0	Between 14 and 47 or 212 and 739
Less than -2.0 or greater than 2.0	Less than 14 or greater than 739

Source: ICP, <http://icp.worldbank.org/>.

**Table 14.7A** Example of Extended Dikhanov Table (Basic Heading Level of Analysis), ICP 2011

Dikhanov temporal analysis		[1] Country 1	[1] Country 2	[1] Country 3	[1] Country 4	[1] Country 5	[2] STD	[3] Count
		[4] Yearly, 2005						
	[5] PPP	658.129	4.040	590.222	7.873	96.795		
	[6] STD	0.256	0.29	0.251	0.279	0.248	0.259	
	[7] No. of priced items	513	572	605	420	481		803
	[8] ER (LCU/US\$)	527.470	5.780	527.470	6.360	100.500		
	[9] Rebased_XR	959.036	10.509	959.036	11.564	182.727		
	[10] PLI	0.686	0.384	0.615	0.681	0.530		
Item-level details		[1] Country 1	[1] Country 2	[1] Country 3	[1] Country 4	[1] Country 5	[11] STD	[12] Count
Item code	Item name	[4] Yearly, 2005						
[13] 110111.1	[14] Rice							
	[15] PPP	718.297	4.849	831.093	6.634	39.381		
	[16] STD	0.073	0.274	0.262	0.517	0.195		
	[17] PLI	0.749	0.461	0.867	0.574	0.216		
	[18] No. of priced items	5	6	5	4	5		
[19] 110111.101	[20] Long grain rice, prepacked	[21]—	[21] 0.267	[21] -0.068	[21] 0.223	[21] -0.162	0.251	5
	[22] Average price	—	5.51	675.00	7.21	29.14		
	[23] No. of observations	—	10	16	4	17		
	[24] Coefficient of variation	—	3.00	26.00	7.84	7.55		
	[25] XR-ratio	—	108.78	146.03	129.36	33.09		
[19] 110111.102	[20] Long grain rice, sold loose	[21] -0.01514	[21] -0.02193	[21] -0.40873	[21] —	[21] -0.16078	0.225	10
	[22] Average price	517.57	3.47	404	—	24.53		
	[23] No. of observations	7	10	14	—	19		
	[24] Coefficient of variation	23.36	7.89	19.69	—	10.16		
	[25] XR-ratio	134.11	82.05	104.68	—	33.36		
[19] 110111.103	[20] Basmati rice	[21] -0.02923	[21] 0.28665	[21] 0.00527	[21] 0.44355	[21] 0.28587	0.282	13
	[22] Average price	1,371.85	12.70	1,643.00	20.33	103.07		
	[23] No. of observations	19	9	11	8	20		
	[24] Coefficient of variation	12.75	8.48	50.58	15.75	10.07		
	[25] XR-ratio	120.05	101.42	143.78	147.55	47.34		

Explanatory notes keyed to the italicized numbers follow table 14.7B. The table is organized in three sections:

- *General section at the top*—gives general details and key indexes that relate to the aggregate or basic heading as a whole, such as number of items [3] included in the analysis, as well as the PPP [5], STD [6], and PLI [10] for each country.
- *Basic heading section in the middle*—gives information that relates to the basic heading as a whole, such as the number of items [12] included in the analysis, as well as the PPP [15], STD [16], and PLI [17] for individual countries.
- *Item section at the bottom*—covers the items priced for the basic heading. Shown for the extended version of the table are the

**Table 14.7B** Example of Collapsed Dikhanov Table (Basic Heading Level of Analysis), ICP 2011

Dikhanov temporal analysis	[1] Country 1	[1] Country 2	[1] Country 3	[1] Country 4	[1] Country 5	[2] STD	[3] Count
	[4] Yearly, 2005						
[5] PPP	658.129	4.040	590.222	7.873	96.795		
[6] STD	0.256	0.292	0.251	0.279	0.248	0.259	
[7] No. of priced items	513	572	605	420	481		803
[8] ER (LCU/US\$)	527.470	5.780	527.470	6.360	100.500		
[9] Rebased_XR	959.036	10.509	959.036	11.564	182.727		
[10] PLI	0.686	0.384	0.615	0.681	0.530		

Item-level details		[1] Country 1	[1] Country 2	[1] Country 3	[1] Country 4	[1] Country 5	[11] STD	[12] Count
Item code	Item name	[4] Yearly, 2005						
[13] 110111.1	[14] Rice							
	[15] PPP	718.297	4.849	831.093	6.634	39.381		
	[16] STD	0.073	0.274	0.262	0.517	0.195		
	[17] PLI	0.749	0.461	0.867	0.574	0.216		
	[18] No. of priced items	5	6	5	4	5		
[19] 110111.101	[20] Long grain rice, prepacked	[21]—	[21]0.26746	[21]–0.06845	[21]0.22277	[21]–0.16158	0.251	13
[19] 110111.102	[20] Long grain rice, sold loose	[21]–0.01514	[21]–0.02193	[21]–0.40873	[21]—	[21]–0.16078	0.225	10
[19] 110111.103	[20] Basmati rice	[21]–0.02923	[21]0.28665	[21]0.00527	[21]0.44355	[21]0.28587	0.282	13
[19] 110111.104	[20] Medium grain rice	[21]0.11338	[21]–0.38882	[21]0.40975	[21]—	[21]–0.14866	0.223	12
[19] 110111.105	[20] Short grain rice	[21]0.03658	[21]–0.33214	[21]0.06217	[21]–0.88125	[21]0.18515	0.342	12
[19] 110111.106	[20] Brown rice	[21]–0.10559	[21]0.18878	[21]—	[21]0.21493	[21]—	0.287	11

Summary Information		
[1]	Country 1, 2, . . . , n	Names of countries covered by table.
[2]	STD 1	Standard deviation (STD) of country product dummy (CPD) residuals of all products priced for basic heading or aggregate. It can be converted to an overall coefficient of variation for products by multiplying it by 100. The mean of all product residuals is 1.
[3]	Count	Number of products specified for basic heading or aggregate.
[4]	Frequency, year	Period during which prices for products covered by table were collected.
[5]	PPP	Purchasing power parity (PPP) for basic heading or aggregate covered by table. It is expressed as number of local currency units per unit of selected numéraire currency. Prices used to calculate PPPs are average prices in local currencies that countries report for products they priced for basic heading or aggregate—that is, the average prices in row [22].
[6]	STD 2	Standard deviation of each country's CPD residuals for basic heading or aggregate. It can be converted to a country coefficient of variation by multiplying by 100. Mean of each country's residuals is 1.
[7]	No. of priced items	Number of products priced by each country for basic heading or aggregate.
[8]	ER (LCU/US\$)	Market exchange rate (ER) of countries expressed as number of local currency units (LCUs) per U.S. dollar.
[9]	Rebased_XR	Exchange rate (XR) [8] rebased to numéraire currency. Number of local currency units per unit of numéraire currency.
[10]	PLI	Price level index (PLI). PPPs in row [5] expressed as ratio of corresponding rebased exchange rates in row [9] for basic heading or aggregate.

Item-Level Details		
[13]	BH code	Code of basic heading covered by table.
[14]	BH name	Name of basic heading covered by table.
[15]	PPP	Purchasing power parity for each country for basic heading covered.
[16]	STD 3	Standard deviation of each country's CPD or country product representative dummy (CPRD) residual for basic heading.
[17]	PLI	Price level index for each country for basic heading covered.
[18]	No. of priced items	Number of products priced by each country for basic heading.
[19]	Item code	Code of item covered by table.
[20]	Item name	Name of item covered by table.
[21]	CPD residual	CPD residual by product and country. See section on use of CPD residuals for additional information.
[22]	Average price	Average item price in local currency units.
[23]	No. of observations	Number of price observations on which average prices at [22] are based.
[24]	Coefficient of variation	Price observation coefficient of variation for each country.
[25]	XR-ratio	Standardized price ratio based on exchange rate—converted price. Converted price is expressed as a percentage of its geometric mean.

Source: ICP, <http://icp.worldbank.org/>.

Note: — = not available. Cells with CPD residuals are color-coded to facilitate visual diagnostics. See table 14.8 for color codes.

**Table 14.8** Color Scheme for CPD Residuals, ICP 2011

CPD residuals with values	Color code
Between -0.25 and 0.25	None
Between -0.75 and -0.25 or 0.25 and 0.75	Yellow
Between -2.0 and -0.75 or 0.75 and 2.0	Red
Less than -2.0 or greater than 2.0	Black

Source: ICP, <http://icp.worldbank.org/>.

item coefficient of variation [11], number of countries pricing the item [12], CPD residuals [21], average price in national currency [22], number of observations [23], coefficient of variation [24], and XR-ratio [25]. Only the item coefficient of variation [11], number of countries pricing the item [12], and CPD residuals [21] are shown for the *collapsed version* of the table.

### Use of CPD Residuals

CPD residuals are used throughout the Dikhanov table (see annex). As shown in the annex, the residuals from CPD regressions are presented as

$$\varepsilon_{cp} = \ln p_{cp} - x_{cp}\beta = \ln p_{cp} - Dc_c - Dp_p, \quad [14.1]$$

where  $Dc_c$  and  $Dp_p$  are the country and product dummies.

During the validation, CPD residuals and the standard deviation of CPD residuals should be analyzed in a manner similar to that in which PPP-ratios and coefficients of variation are analyzed in the Quaranta tables. The process of data validation with the Dikhanov table should thus start with checking the entries with the largest negative or positive residuals, trying to investigate and resolve these issues. Some of the deviations, even very large ones, can be legitimate. For example, the price of gasoline in República Bolivariana de Venezuela is very low compared with the prices in other Latin American countries, and therefore the large deviation (CPD residual) for the product as shown in the Dikhanov table at the GDP level is not a mistake in data. In general, the overall STD by country should reflect the quality of the price data.

It should also be noted that not every problem can be observed when the CPD is run only at the basic heading level. For example, if a country erroneously priced all its beverages in gallons instead of liters, its BH-level data alone could be very consistent (because everything is priced in gallons), but inconsistent with those of all other countries (priced in liters).

It is recommended that the Dikhanov table be run at different levels of aggregation<sup>8</sup>—for example, at the basic heading level, at a higher

**Table 14.9A** Example of Collapsed Dikhanov Table (Calculated at Basic Heading Level), ICP 2011

Dikhanov temporal analysis		Country 1	Country 2	Country 3	Country 4	Country 5	Country 6	STD	Count
		Yearly, 2005							
<b>PPP</b>		1.000	4.066	0.049	3.442	8.548	4.727		
<b>STD</b>		0.502	0.728	0.750	0.711	0.800	0.717	0.705	
<b>No. of priced items</b>		706	747	562	766	661	691		844
<b>ER (LCU/US\$)</b>		88.600	527.468	5.110	527.468	1138.000	393.383		
<b>Rebased_XR</b>		1.000	5.953	0.058	5.953	12.844	4.440		
<b>PLI</b>		1.000	0.683	0.851	0.578	0.666	1.065		
<b>Item-level details</b>		<b>Country 1</b>	<b>Country 2</b>	<b>Country 3</b>	<b>Country 4</b>	<b>Country 5</b>	<b>Country 6</b>	<b>STD</b>	<b>Count</b>
<b>Item code</b>	<b>Item name</b>	Yearly, 2005							
<b>110111.1</b>	<b>Rice</b>								
	<b>PPP</b>	1.000	5.504	0.065	1.480	8.052	1.579		
	<b>STD</b>	0.596	0.640	0.441	0.634	0.614	0.653		
	<b>PLI</b>	1.000	0.924	1.120	0.249	0.627	0.356		
	<b>No. of priced items</b>	7	7	3	7	5	7		
110111.101	Long grain rice	-0.176	0.582	—	-0.402	0.406	-0.410	0.416	5
110111.102	Long grain rice	-0.665	0.698	0.502	-0.231	—	-0.305	0.515	5
110111.103	Long grain rice	-0.288	-0.400	-0.571	0.886	-0.541	0.913	0.643	6
110111.104	Medium grain rice	0.202	-1.157	—	0.071	0.802	0.082	0.638	5
110111.105	Short grain rice	-0.667	-0.464	—	0.976	-0.863	1.018	0.824	5
110111.106	Basmati rice	1.094	0.205	0.068	-0.722	—	-0.645	0.660	5
110111.107	Broken rice	0.501	0.536	—	-0.579	0.195	-0.654	0.518	5
<b>110111.2</b>	<b>Other cereals, flour, and other cereal products</b>								
	<b>PPP</b>	1.000	3.531	0.052	3.274	4.962	6.010		
	<b>STD</b>	0.496	0.774	0.806	0.726	1.062	0.704		
	<b>PLI</b>	1.000	0.593	0.905	0.550	0.386	1.354		
	<b>No. of priced items</b>	14	20	13	21	11	19		
110111.201	Wheat flour	0.221	0.285	-0.371	0.357	-0.256	-0.236	0.293	6
110111.202	Wheat flour	0.419	0.122	-0.328	0.313	-0.087	-0.439	0.315	6
110111.203	Couscous	—	0.311	0.288	0.291	—	-0.314	0.302	4
110111.204	Couscous (millet)	—	-1.450	—	0.986	—	0.465	1.047	3

aggregate level (such as food), and at the GDP level. Processing the Dikhanov table at a level higher than the basic heading can help analyze price points for "bad" basic headings with partially erroneous price entries. These "bad" basic headings would be distorted, which would make the processing impossible at the basic heading level. However, processing at a higher level would help identify prices that are consistent within a broader set of products.

It is important to study the overall standard deviation of residuals (upper right-hand corner of the table). Table 14.9A, an example of a collapsed Dikhanov table calculated at the basic heading level, predictably shows a smaller value than table 14.9B (0.71 versus 0.80), an example of a collapsed Dikhanov table calculated at the GDP level.<sup>2</sup> However, the difference is not large (and even smaller for countries with data of poorer quality, in which case the within-basic heading

**Table 14.9B** Example of Collapsed Dikhanov Table (Calculated at GDP Level), ICP 2011

Dikhanov temporal analysis	Country 1	Country 2	Country 3	Country 4	Country 5	Country 6	STD	Count
	Yearly, 2005							
<b>PPP</b>	1.000	4.066	0.049	3.442	8.548	4.727		
<b>STD</b>	0.623	0.812	0.839	0.785	0.893	0.817	0.797	
<b>No. of priced items</b>	706	747	562	766	661	691		844
<b>ER (LCU/US\$)</b>	88.600	527.468	5.110	527.468	1138.000	393.383		
<b>Rebased_XR</b>	1.000	5.953	0.058	5.953	12.844	4.440		
<b>PLI</b>	1.000	0.683	0.851	0.578	0.666	1.065		

Item-level details		Country 1	Country 2	Country 3	Country 4	Country 5	Country 6	STD	Count
Item code	Item name	Yearly, 2005							
110111.101	Long grain rice	0.163	1.224	—	-0.906	0.686	-1.167	0.914	5
110111.102	Long grain rice	-0.392	1.273	1.050	-0.802	—	-1.128	0.979	5
110111.103	Long grain rice	-0.051	0.140	-0.058	0.279	-0.364	0.054	0.200	6
110111.104	Medium grain rice	0.541	-0.515	—	-0.433	1.082	-0.675	0.689	5
110111.105	Short grain rice	-0.328	0.179	—	0.472	-0.583	0.261	0.392	5
110111.106	Basmati rice	1.366	0.780	0.616	-1.293	—	-1.469	1.156	5
110111.107	Broken rice	0.841	1.178	—	-1.084	0.475	-1.410	1.047	5
110111.201	Wheat flour	0.293	0.216	-0.237	0.379	-0.728	0.077	0.380	6
110111.202	Wheat flour	0.491	0.052	-0.194	0.335	-0.558	-0.127	0.347	6
110111.203	Couscous	—	0.142	-0.254	0.214	—	-0.102	0.188	4
110111.204	Couscous (millet)	—	-1.608	—	0.919	—	0.688	1.141	3

Source: ICP, <http://icp.worldbank.org/>.

Note: — = not available. See tables 14.7A and 14.7B for explanation of abbreviations. Cells with CPD residuals are color-coded to facilitate visual diagnostics. See table 14.8 for color codes.

deviations would dominate the between–basic heading ones). The overall standard deviation of residuals for larger regions with more variety tends to be larger than those for smaller and more uniform regions.

It is informative to study the standard deviations of residuals both by country and by item. The same overall picture can be observed: the STD values in table 14.9B are greater than those in table 14.9A.

As for the Quaranta table, it may be concluded that the goal of the validation is to reduce the overall standard deviation with the understanding that there are limits to its reduction, and that many large CPD residuals may be quite legitimate. However, each large residual needs to be investigated.

### Modified Dikhanov Table

Another variation of the Dikhanov table is the modified Dikhanov table. This table keeps the original presentation of the Dikhanov table, either collapsed or extended, but instead of using the Dikhanov table–specific indexes, it uses the indexes in the Quaranta table. Table 14.10 summarizes the differences between the original and modified Dikhanov tables. Table 14.11 is an example of a modified Dikhanov table.

The main advantage of the modified Dikhanov table is that users familiar with Quaranta indexes can benefit from the layout of the Dikhanov table, as well as from the option to calculate the tables and the respective PPPs at different levels of aggregation.

**Table 14.10** Indexes Used in the Original and Modified Dikhanov Tables, ICP 2011

Dikhanov table	Modified Dikhanov table
CPD residuals	PPP-ratios
Standard deviation	Coefficient of variation
Price level index (absolute)	Price level index (percentage)
No. of priced items	No. of priced items and no. of important items

Source: ICP, <http://icp.worldbank.org/>.

**Table 14.11** Example of Modified Dikhanov Table (Calculated at Basic Heading Level), ICP 2011

Dikhanov temporal analysis (calculation method: CPD)		Country 1	Country 2	Country 3	Country 4	Country 5	Var. co.	Count
		Q1 2011						
<b>PPP</b>		1.566	1	224.374	1.804	1.384		
<b>Variation coefficient</b>		29.7	17.3	23.9	18.9	24.9	21.6	
<b>No. of priced items</b>		30;*30	30;*30	30;*30	30;*30	30;*30		30
<b>ER (LCU/US\$)</b>		1.646	1	178.100	1.785	0.8444		
<b>Rebased_XR</b>		1.646	1	178.100	1.785	0.844		
<b>PLI (%)</b>		95.144	100	125.982	101.066	163.892		

  

Item-level details		Country 1	Country 2	Country 3	Country 4	Country 5	Var. co.	Count
Item code	Item name	Q1 2011						
<b>110111.1</b>	<b>Rice</b>							
	<b>PPP</b>	1.50157	1	223.787	1.81769	1.18924		
	<b>Variation coefficient</b>	32.4	24.4	16.9	23.8	22.2		
	<b>PLI (%)</b>	91.214	100	125.652	101.847	140.838		
	<b>No. of priced items</b>	6;*6	6;*6	6;*6	6;*6	6;*6		
110111.101	Long grain rice, parboiled	76.22	143.6	104.14	120.56	72.77	28.9	5
	<i>No. of observations</i>	*692	*472	*51	*620	*139		
	<i>Coefficient of variation</i>	49.3029	24.1644	32.3469	27.9986	11.8166		
	<i>XR-ratio</i>	62.9456	130.007	118.465	111.165	92.7914		
110111.102	Long grain rice, non-parboiled	92.43	107.6	95.25	121.65	86.77	13.8	5
	<i>Average price</i>	52.2131	40.479	8,018.59	83.1814	38.8185		
	<i>No. of observations</i>	*685	*473	*46	*595	*139		
	<i>Coefficient of variation</i>	37.6515	38.8443	19.3825	25.3717	13.0202		
	<i>XR-ratio</i>	76.3336	97.4204	108.356	112.169	110.639		
110111.103	Long grain rice, family pack	170.06	66.26	77.65	126.56	90.31	39.8	5
	<i>Average price</i>	360.21	93.4659	24,510.9	324.504	151.494		
	<i>No. of observations</i>	*520	*457	*40	*549	*135		
	<i>Coefficient of variation</i>	36.3971	48.4808	47.7075	44.9962	16.5691		
	<i>XR-ratio</i>	140.44	59.9888	88.3309	116.699	115.149		

Source: ICP, <http://icp.worldbank.org/>.

Note: — = not available. See tables 14.7A and 14.7B for explanation of abbreviations. Cells with CPD residuals are color-coded to facilitate visual diagnostics. See table 14.8 for color codes. An asterisk (\*) indicates number of important items out of total number of priced items.

## Annex

### CPD Residuals in the Dikhanov Table

By definition, the CPD index for a set of countries and products can be presented as follows:

$$\ln p_{cp} = y_{cp} = x_{cp} \beta + \varepsilon_{cp} \quad (14A.1)$$

and

$$x_{cp} = [Dc_2 \dots Dc_{Nc} Dp_1 Dp_2 \dots Dp_{Np}] \quad (14A.2)$$

$$\beta = [\alpha_2 \dots \alpha_{Nc} \gamma_1 \gamma_2 \dots \gamma_{Np}]^T,$$

where  $P_{cp}$  is the price of product  $p$  in country  $c$ ;  $Dc_j$  and  $Dp_i$  are the country and product dummies, respectively; and  $Np$  and  $Nc$  are the number of products and countries, respectively.

In matrix notation, by stacking individual observations, this can be written as

$$y = X\beta + \varepsilon. \quad (14A.3)$$

Note that the first country dummy is dropped from the system because matrix  $X$  is of rank  $(Np + Nc - 1)$ . In fact, any variable can be dropped from the system; dropping the first country's dummy simply makes it the base country.

The solution is given (under the conditions of independently and identically distributed random disturbances) by

$$\hat{\beta} = (X^T X)^{-1} X^T y. \quad (14A.4)$$

With use of expression (14A.3), the error term can be presented as

$$\varepsilon_{cp} = \ln p_{cp} - x_{cp} \beta = \ln p_{cp} - Dc_c - Dp_p. \quad (14A.5)$$

Those error terms (residuals) enter the Dikhanov table and are used in its statistics such as the standard deviations of residuals by country and product.

The CPD residuals can be estimated at any level, starting from the basic heading and up to the GDP level. It is possible to compute the residuals at various other groupings as well—for example, for goods, services, nontradables, or unprocessed food. In those cases, the residuals will indicate the variability within those groups only.

## NOTES

1. A standardized price ratio is  $(\text{CC-price}_{1A} / [\text{CC-price}_{1A} * \text{CC-price}_{1B} * \dots * \text{CC-price}_{1N}]^{1/N}) * 100$  where  $\text{CC-price}_{1A}$  is the average price for product 1 in country A in the common currency.  $\text{CC-price}_{1A}$  is itself equal to  $\text{NC-price}_{1A} / \text{CC}_{1A}$  where  $\text{NC-price}_{1A}$  is the average price for product 1 in country A in the national currency, and  $\text{CC}_{1A}$  is the currency conversion rate between the national currency of country A and the common currency. The currency conversion rate is either the exchange rate or the PPP:  $\text{CC}_{1A} = \text{XR}_{1A}$  or  $\text{PPP}_{1A}$ .
2. An alternative term is the CUP-ratio, in which CUP stands for conventional unit for expressing parities.
3. In this particular case, it does not matter which is used because the PPP prices and PPP price ratios for the same item in different countries differ only by a scalar.
4. In the ICP Kit, the bottom part of the table is generated automatically, whereas the upper part must be completed manually.
5. If a base country approach is used in calculation of the PPPs, missing basic headings can be caused by missing price data for either the given country or the base country.
6. GDP PPP is estimated here as the CPD PPP utilizing the whole set of prices and products, and thus does not take into account basic heading expenditures. The advantage is that the CPD PPP at the aggregate level can be estimated before the actual basic heading weights are known, and it will still provide a ballpark estimate of the final PPP for the GDP.
7. Horizontal Quaranta tables, used in the European Comparison Program in 1993 and 1996, have a layout similar to that of the extended Dikhanov tables.
8. The difference between the residuals, run at different levels of processing such as the basic heading and GDP, can be expressed as

$$\varepsilon_{cp}(BH) - \varepsilon_{cp}(GDP) = -Dc_c(BH) - Dp_p(BH) + Dc_c(GDP) + Dp_p(GDP).$$

Because the country dummy is the log of the PPP, the difference between the CPD residuals in tables 14.9A and 14.9B can be broken down into two parts:

$$\begin{aligned} \varepsilon_{cp}(BH) - \varepsilon_{cp}(GDP) = & \left\{ \ln(PPP_c(GDP)) \right. \\ & \left. - \ln(PPP_c(BH)) \right\} \\ & + \left\{ Dp_p(GDP) - Dp_p(BH) \right\}. \end{aligned}$$

The first component is the difference between the BH PPP and the GDP PPP (the relative

price level of the basic heading vis-à-vis the overall price level at the GDP level), whereas the second component is the difference between the *logs* of the average prices of the product as computed with the CPD regression on all products and those as computed on the products within the basic heading only. Usually, the second component is insignificant.

9. The overall standard deviation of residuals is run on the whole country product tableau of the CPD residuals.