International Comparison Program

Validation Tables

Draft version

Operational Guide

Contents

Validation Tables[1]

1. Introduction

The main price data analysis within regional and global level validation is carried out by using two validation tables. These are Quaranta Table (QT), named after Vincenzo Quaranta, who first proposed them for use in the European PPP program in 1990, and Dikhanov Table (DT), named after Yuri Dikhanov, who first proposed the table for use during the course of ICP 2005. The purpose of both tables is to screen the national average prices for possible errors by comparing average prices of same items across countries. Respectively, both tables provide similar measures of price variation for basic headings, countries, and items.

The main difference between the two types of table is that the QT is employed to edit prices within basic headings (BHs), whereas the DT can also be used to edit prices within aggregates. In addition, the DT can be set to show only key indices while hiding other available details. When presented in this compact form, the DT is better suited for editing prices across the basic headings and items comprising an aggregate.

The DT is specific to the CPD or CPRD method of calculating PPPs, while the QT has a broader application that includes the EKS and EKS* methods as well as the CPD and CPRD methods.

This chapter of the ICP Operational Guide explains the characteristics of the Quaranta and the Dikhanov Tables. Next chapter "*Validation of Household Consumption Survey* focuses" will discuss the usage of these tables during the validation process. Finally, the chapter on "*Computation and Linking"* explains the PPP calculation methods used in the Quaranta and the Dikhanov Tables.

2. Quaranta Table

The Quaranta Table consists of a set of tables for basic headings, one for the BH as a whole and one for each item within the BH. For the comparison of average prices, the QT provides three main measures: XR-Ratios, PPP-Ratios and a Price Level Index (PLI). For the analysis of price variation, the QT provides four measurements ranging from variation of individual price observations to variation of items within a BH. Each of these indices is discussed in detail below.

3.1 Average price measures

Once converted to a common currency, the average prices of different countries for *the same item* can be compared with each other and extreme values can be identified. However, prices cannot be compared *across items* directly, even when expressed in the same currency. On the other hand, the price ratios of countries pricing an item can be compared with the equivalent price ratios for other items once they have first been "standardized".

[1] This chapter is prepared by Marko Rissanen, with input from Biokou Mathieu Djayeola, based on material prepared by David Roberts, Yuri Dikahnov, and Nada Hamadeh.

Standardized price ratios for an item are the ratios between the individual average prices of the countries pricing the item and the geometric mean of the average prices of all the 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 standardized price ratios.

The standardized price ratios[2] based on exchange rate converted prices are called "XR-Ratios" and the standardized price ratios based on PPP converted prices are called "PPP-Ratios" [3].

In addition to XR- and PPP-Ratios, the third measure to compare prices is the Price Level Index (PLI). This is defined as the ratio of the basic heading PPP to the XR and is expressed as a percentage. A PLI that is 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 prices in the base country of the group.

3.2 Measures of price variation

The Quaranta Table provides four variation coefficients as follows:

- *Overall average variation coefficient*: Measures dispersion among all the PPP-Ratios for a basic heading. In doing so, it measures price structures' homogeneity 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 the price structures will be resulting in less reliable PPPs.

- *Country variation coefficient*: 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 coefficients' value, the less uniform the country's price levels will be leading to less reliable PPPs.

- *Item variation coefficient*: Measures dispersion among the PPP-Ratios for an item. It is an indicator of comparability and accuracy that addresses the question whether comparable products have been priced for the item in question. The higher the coefficient's value, the less uniform the item price levels will be. This would raise questions regarding the comparability and accuracy of the item's pricing across countries.

- *Price observation variation coefficient*: 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 country level validation.

[2] A standardized price ratio equals ($CC\text{-}Price_{1A}$ / [$CC\text{-}Price_{1A} * CC\text{-}Price_{1B} * \ldots\ldots CC\text{-}Price_{1N}$]$^{1/N}$) 100 where $CC\text{-}Price_{1A}$ is the average price for product 1 in country A in the common currency. $CC\text{-}Price_{1A}$ is itself equal to $NC\text{-}Price_{1A}$ / CC_{1A} where $NC\text{-}Price_{1A}$ is the average price for product 1 in country A in national currency and CC_{1A} is the currency conversion rate between the national currency of A and the common currency. The currency conversion rate is either the exchange rate or the PPP: $CC_{1A} = XR_{1A}$ or PPP_{1A}.

[3] Alternative term is CUP-Ratio, where CUP stands for Conventional Unit for expressing Parities.

Besides serving as editing tools, the coefficients provide the means to monitor progress during the validation stage and, at its conclusion, to assess the effectiveness of the entire process of editing and verification in reducing the incidence of non-sampling error among the price data. Coefficients should be significantly smaller at the end of validation than they were at the beginning.

3.3 Description of the Quaranta Table

Table 1 below illustrates 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:

- The first section, **Data Selection Criteria**, provides general details about the table such as the averaging method *[4]*, imputation method *[5]* and the run date *[3]*.

- The second section, **Summary Information**, gives information that relates to the basic heading as a whole such as number of items *[6]* and countries *[8]* included in the analysis as well as average weight of the basic heading in total expenditure *[7]* and Average Coefficient Variation *[9]*.

- The third section, **Country Level Details**, gives key indices that relate to individual countries at BH level. Included are PPPs *[13]*, PLIs *[14]*, number of priced items *[16]* and country variation coefficients *[17]*.

- The fourth and the final section, **Item Level Details**, covers the items priced for the basic heading. Each items has its own table that shows the item variation coefficient *[20]*; the average prices reported by countries in national currencies *[23]* the average prices converted to a common currency with 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 i.e. PPP-Price *[28]*, their geometric mean *[32]* and their PPP-Ratios *[29]*.

Table 1: An example of a Quaranta Table

QUARANTA TABLE DIAGNOSTICS - Rice

Data Selection Criteria

[1] **Basic Heading Code**	1101111	*[2]* **Time Period**	Q1, 2011	*[3]* **Run Date**	3/29/2012
[4] **Averaging Method**	Arithmetic Mean	*[5]* **Imputation**	CPD		

Summary Information

[6] **No of Items included in the Analysis**	11 out of	*[7]* **Average weight of Basic Heading in Total Expenditure**	0.0
[8] **No of Countries included in the Analysis**	5 out of 5	*[9]* **Average Coefficient Variation**	30.9
[10] **Base Country**	Country		

Country Level Details

 # Shares are multiplied by 10000.

[11] Country	*[12]* XR	*[13]* PPP	*[14]* PLI(%)	*[15]* Weight #	*[16]* Items	*[17]* Var.Co.
Country 1	3104.03	2167.5700	69.83%	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

[18] 110111101 — [19] Long grain rice - Parboiled — *[20]* **Var.Co.:** 34.2 — *[21]* 1-Kilograms

[22] Country	*[23]* NC-Price	*[24]* Quotations	*[25]* Var.Co.	*[26]* XR-pr	*[27]* XR-ratio	*[28]* PPP-price	*[29]* PPP-ratio	*[30]* Pref. UoM
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
	Geo Mean			*[31]* 1.08		*[32]* 1.20		

QUARANTA TABLE DIAGNOSTICS		
[1]	Basic Heading Code	Code for the basic heading covered by the table
[2]	Time period	Period during which the prices for the products covered by the Table were collected.
[3]	Run Date	Date the Table was computed.
[4]	Averaging Method	Method used to calculate average values in the table.
[5]	Imputation	Method used to calculate the basic heading PPPs in column *[13].* Currently the CPD but it could also be the CPRD, the EKS or the EKS*.

SUMMARY INFORMATION		
[6]	No. of Items	Number of items specified and included in the basic heading analysis.
[7]	Av. Weight	Average expenditure weight for the group of countries covered by the basic heading. The unweighted arithmetic mean of the national weights in column *[15]*. Like the national weights it is scaled to 100,000.
[8]	No. of Countries	Total number of countries and number of countries included in the basic heading analysis.
[9]	Average Coefficient Variation (Var.Co.1)	Overall average variation coefficient or, more precisely, the average item variation coefficient for the items priced for the basic heading. It is calculated as the unweighted arithmetic mean of the product variation coefficients at [20]. It measures the average variation of the PPP-Ratios in column [29] of all products priced for the basic heading.
[10]	Base Country	Country and respective currency selected as numéraire. Any of the countries and respective currencies included to the analysis can be chosen as numéraire.

COUNTRY LEVEL DETAILS		
[11]	Country	Names of countries covered by the Table.
[12]	XR	Market exchange rates of the countries expressed as the number of units of national currency per unit of the numéraire currency specified in *[10]*.
[13]	PPP	Purchasing power parities for the basic heading calculated as specified in *[5]* and expressed as the number of units of national currency per unit of the selected numéraire currency specified in *[10]*. The prices used to calculate the PPPs are the average prices in national currencies that countries report for the products they priced for the basic heading – that is, the NC-Prices in column *[23]*.
[14]	PLI	Price Level Indices. The PPPs in column *[13]* expressed as a percentage of the corresponding exchange rate in column *[12]*.
[15]	Weight	National expenditure weights multiplied by 10000. That part of a country's GDP that is spent on the basic heading when both expenditures are expressed in national currency and valued at national price levels.

[16]	No. of Items	Number of items that are priced by each country for the basic heading The number with an asterisk (*) represent the number of important items.
[17]	Var.Co.2	Country variation coefficient. The standard deviation of the country's PPP-Ratios in column [29] for all products priced by the country for the basic heading expressed as a percentage of the arithmetic mean of the country's PPP-Ratios in column [29] for all products priced by the country for the basic heading.

ITEM LEVEL DETAILS		
[18], [19]	Item code and name	Code, name, and summary definition of the product covered in the subsequent product section.
[20]	Var.Co.3	Item variation coefficient. The standard deviation of the item's PPP-Ratios in column *[29]* expressed as a percentage of the arithmetic mean of the 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 the Table.
[23]	NC-Price	Average price for the product in national currency (NC).
[24]	Quotations	Number of price observations on which the average prices in national currency in column *[23]* are based. An asterisk (*) indicates whether the item is important for the given country.
[25]	Var.Co.4	Price observation variation coefficient. The standard deviation of the price observations underlying the product's average price in column *[23]* expressed as a percentage of the arithmetic mean of the price observations underlying the product's average price in column *[23]*.
[26]	XR-Price	The average prices in national currency in column *[23]* converted to the numéraire currency with the exchange rates in column *[12]*.
[27]	XR-Ratio	Standardized price ratios based on the exchange rate converted prices in column *[26]*. The XR-Prices expressed as a percentage of their geometric mean at *[31]*.
[28]	PPP-Price	The average prices in national currency in column *[23]* converted to the numéraire currency with the PPPs in column *[13]*.
[29]	PPP-Ratio	Standardized price ratios based on the PPP converted prices in column *[28]*. The PPP-Prices expressed as a percentage of their geometric mean at *[32]*.
[30]	Pref. UoM	Preferred unit of measurement for the item.
[31]	Geo Mean	Geometric mean of the exchange rate converted prices in column [26]. The use of a geometric mean here and in [32] insures invariance with respect to choice of numéraire.
[32]	Geo Mean	Geometric mean of the PPP converted prices in column [28].

3.4 Use of XR- and PPP-Ratios

The XR 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- or PPP-ratio means that the national average price for the item in question is high (low) compared to the prices of the same product in other countries when converted into a common numéraire currency using the XR or basic heading PPP. When using the same logic 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 whether the price(s) may be erroneous. Experience suggests that the appropriate thresholds for the individual XR or PPP-Ratios are 50 and 150. To ease detecting high (low) PPP-Ratios, a color scheme presented below in Table 2 is used.

Table 2: Color scheme for the PPP-Ratios

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

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 the fact 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 to be 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 another. This could happen in real world scenario, but on the other hand, it could also 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 variation coefficients for the PPP prices or the PPP price ratios[4].

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

3.5 Country Diagnostic Report

The standard Quaranta Table gives tables for each basic heading and 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 basic heading level analysis. This grouping is called Country Diagnostic Report. In essence the report compiles basic headings and item rows for a single country, and presents these as an individual report. All information is gathered and calculated as with the standard Quaranta Table; the difference is only in the presentation. Table 3 below gives an example of a Country Diagnostic Report.

The Country Diagnostic Report has four sections, similar to the standard Quaranta Table. The first two sections, Data Selection Criteria and Summary Information, are identical with the standard table. However the following two last sections, Country Level Details and Item Level Details, contain information only for the given country.

This way of presentation has two main benefits:

- It can serve as an effective validation tool especially for the countries, by quickly pinpointing any potential problems with data; and

- It allows focusing on internal price structure and potential systematic problems in a country's data, which may go unnoticed when information is presented for all countries included in the analysis.

However, it should be noted that the 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 full item context, by studying the respective information and indices 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.

[4] In this particular case, it does not matter which is used as the PPP prices and PPP ratios for the same item in different countries differ only by a scalar. The PPP price ratios are derived by dividing the PPP prices by a constant, their geometric mean. As the variation coefficient (i.e., the standard deviation divided by the arithmetic mean) is essentially a measure of relative dispersion is the same for vectors of observations that differ only by a scalar.

Table 3: An example of a Country Diagnostic Report

COUNTRY DIAGNOSTICS REPORT - Country

Basic Heading Code	1101111	**Time Period**	Q1, 2011	**Run Date**	3/29/2012
Averaging Method	Arithmetic Mean	**Imputation**	CPD		

Summary Information

No of Items included in the Analysis	11 out of 11	**Average weight of Basic Heading in Total Expenditure**	0.0
Base Country	Country	**Average Coefficient Variation**	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

 # Shares are multiplied by 10000.

Item Level Details

Product Code	Product Name	Pref. UoM	NC-Price	Quotations	Var.Co.	XR-Ratio	PPP-Ratio
110111101	Long grain rice - Parboiled	1 - Kilogram	1.309	*17	17.5	120.94	109.40
110111102	Long grain rice - Non-Parboiled	1 - Kilogram	1.347	*14	2.8	117.09	107.99
110111103	Long grain rice - Family Pack	5 - Kilogram	3.031	*12	7.0	88.52	82.17
110111104	Jasmine Rice	10 - Kilogram	6.639	*7	3.5	95.41	88.00
110111105	Basmati Rice	1 - Kilogram	1.104	*4	53.7	104.10	89.42
110111106	White rice, 25% broken	1 - Kilogram	-	*-	-	-	-
110111107	White rice, Medium Grain	1 - Kilogram	-	-	-	-	-
110111108	Brown rice - Family Pack	5 - Kilogram	-	*-	-	-	-
110111109	Short-grained rice	1 - Kilogram	-	-	-	-	-
110111110	Uncle Ben's Rice	5 - Kilogram	6.313	*10	5.9	103.47	93.60
110111111	Thailand Rice	1 - Kilogram	0.811	*2	4.9	150.66	139.87

3. Dikhanov Table

The second validation table used in the context of the International Comparison Program is the Dikhanov Table (DT). As explained in the introduction, both DT and QT tables use a similar approach and concepts: studying item price deviations for each country in a two-dimensional space: that of items and countries. While the Quaranta Table (QT) is intended to serve as a diagnostic tool for prices at the basic heading level, the Dikhanov Table can be processed at any level from total GDP[5] down to the basic heading. The DT can also be processed for intermediate aggregates, such as goods, services etc. The QT 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 DT places an emphasis on the between basic heading validation, adding features to detect anomalies across both countries and basic headings.

In the QT, PPPs are computed using one of the four methods: EKS, EKS*, CPD or CPRD. In the DT, the CPD and CPRD are used in computations as EKS does not generate the average product price, an important measure that enters in various computations in the DT.

3.1 Average price measures

Similarly to the QT, DT provides XR-Ratios at the item level and Price Level Indices at the aggregate or BH level. However, for the PPP-ratios, the approach of the DT differs from that of the QT. 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 DT uses CPD residuals instead of PPP-ratios as the double-normalized item price. When the CPD is used, the CPD residuals in the DT are equal to the logarithms of the PPP-Ratios in the QT. However, this identity only holds if:

- The PPPs in both tables are calculated with CPD. It will not hold if the CPD is used for the DT but, EKS for the QT; and

- 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 DT are based on PPPs for an aggregate and the PPP-Ratios in the QT are derived from PPPs for the basic heading.

Table 4 below summarizes the connection of the CPD residuals and PPP-Ratios, given that the above mentioned conditions hold.

[5] 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 ball park estimate of the final PPP for the GDP.

Sections *3.3 Description of the Dikhanov Table* and *3.4 Use of CPD Residuals* below give details on the CPD residuals.

Table 4: CPD residuals and respective PPP-Ratios

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

3.2 Measures of price variation

The Dikhanov Table uses standard deviations instead of variation coefficients to measure overall country and item variation. The standard deviations for the CPD residuals in the DT are variation coefficients as the mean of the residuals is 1. The two sets of coefficients are not the same due to differences in computation. The overall variation coefficient in the Quaranta table is an average of the variation coefficients of the items priced for the basic heading, whereas the overall variation coefficient in the DT is computed with all the CPD residuals in the tables' item section, thereby ensuring consistency between the overall variation coefficient, the item variation coefficients and the country variation coefficients.

Additionally, the item variation coefficients in the QT 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. This is not the case with the item variation coefficients in the DT. These are based on CPD residuals which are logarithms of the PPP-Ratios. Despite the computational differences, these two sets of variation coefficients are of similar orders of magnitude and reliability in terms of identifying extreme values.

3.3 Description of the Dikhanov Table

The Dikhanov tables can be presented in two versions: extended and collapsed. Table 6A below provides an example of an extended table and Table 6B of a collapsed table for household final consumption expenditure (HFCE). Both tables are calculated at BH level and neither table is complete. CPD residuals are shown for only three (6A) or six (6B) of the 809 products priced and only for five 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.

The numbers in *italics* have been added to the tables for ease of reference. Explanatory notes follow the table. The table is organized in three sections:

- **General section** at the top of the table gives general details and key indices that relate to the aggregate or basic heading as a whole such as number of items *[3]* included to the analysis as well as PPP *[5]*, STD *[6]* and PLI *[10]* for each country.

- **BH section** at the middle of the table gives information that relates to the basic heading as a whole such as number of items *[12]* included to the analysis as well as PPP *[15]*, STD *[16]* and PLI *[17]* for the individual countries.

- **Item section** at the bottom covers the items priced for the basic heading. For the *extended* version of the table shown are:

    - The item variation coefficient *[11],* Number of countries pricing the item *[12*], CPD residuals *[21]*, average price in national currency *[22]*, no. of observations *[23]*, coefficient of variation *[24]* and the XR Ratio *[25]*.

  For the *collapsed* version of the table shown are only:

    - The item variation coefficient *[11],* Number of Countries pricing the item *[12]* and CPD residuals *[21].*

In addition, the cells in the report with CPD residuals are color-coded to facilitate visual diagnostics. The color codes are shown below in Table 5.

Table 5: Color scheme for the CPD residuals

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

Table 6A: An example of an extended Dikhanov Table (BH level analysis)

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	*[4]* Yearly - 2005	*[4]* Yearly - 2005	*[4]* Yearly - 2005	*[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	[4] Yearly - 2005	[4] Yearly - 2005	[4] Yearly - 2005	[4] Yearly - 2005		
***[13]* 99.11.01.11.1**	***[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] 99.11.01.11.1.01	*[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] 99.11.01.11.1.02	*[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] 99.11.01.11.1.03	*[20]* Basmati Rice	*[21]* -0.02923	*[21]* 0.28665	*[21]* 0.00527	*[21]* 0.44355	*[21]* 0.28587	0.282	13
	[22] Average Price	1371.85	12.70	1643.00	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		

Table 6B: An example of a collapsed Dikhanov Table (BH level analysis)

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	*[4]* Yearly - 2005	*[4]* Yearly - 2005	*[4]* Yearly - 2005	*[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	*[4]* Yearly - 2005	*[4]* Yearly - 2005	*[4]* Yearly - 2005	*[4]* Yearly - 2005		
***[13]* 99.11.01.11.1**	***[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] 99.11.01.11.1.01	*[20]* Long grain rice, prepacked	*[21]* -	*[21]* 0.26746	*[21]* -0.06845	*[21]* 0.22277	*[21]* -0.16158	0.251	13
[19] 99.11.01.11.1.02	*[20]* Long grain rice, sold loose	*[21]* -0.01514	*[21]* -0.02193	*[21]* -0.40873	*[21]* -	*[21]* -0.16078	0.225	10
[19] 99.11.01.11.1.03	*[20]* Basmati Rice	*[21]* -0.02923	*[21]* 0.28665	*[21]* 0.00527	*[21]* 0.44355	*[21]* 0.28587	0.282	13
[19] 99.11.01.11.1.04	*[20]* Medium grain rice	*[21]* 0.11338	*[21]* -0.38882	*[21]* 0.40975	*[21]* -	*[21]* -0.14866	0.223	12
[19] 99.11.01.11.1.05	*[20]* Short grain rice	*[21]* 0.03658	*[21]* -0.33214	[21] 0.06217	*[21]* -0.88125	*[21]* 0.18515	0.342	12
[19] 99.11.01.11.1.06	*[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 the Table.
[2]	STD 1	Standard deviation of the CPD residuals of all products priced for the basic heading or aggregate. It can be converted to an overall variation coefficient for products by multiplying it by 100. The mean of all product residuals is 1.
[3]	Count	Number of products specified for the basic heading or aggregate.
[4]	Frequency - year	Period during which the prices for the products covered by the Table were collected.
[5]	PPP	Purchasing power parities for the basic heading or aggregate covered by the Table. They are expressed as the number of local currency units per unit of the selected numéraire currency. The prices used to calculate the PPPs are the average prices in local currencies that countries report for the products they priced for the basic heading or aggregate - that is the average prices in row [22].
[6]	STD 2	Standard deviation of each country's CPD residuals for the basic heading or aggregate. It can be converted to a country variation coefficient by multiplying by 100. The mean of each country's residuals is 1.
[7]	No.of Priced Items	Number of products that are priced by each country for the BH or aggregate.
[8]	ER (LCU/US$)	Market exchange rates of countries expressed as the number of local currency units per US dollar.
[9]	Rebased_XR	Exchange rates [8] rebased to the numéraire currency. Number of local currency units per unit of numéraire currency.
[10]	PLI	Price level indices. The PPPs in row [5] expressed as a ratio of the corresponding rebased exchange rates in row [9] for the basic heading or aggregate.

ITEM LEVEL DETAILS		
[13]	BH Code	Code of the Basic Heading covered by the table.
[14]	BH Name	Name of the Basic Heading covered by the table.
[15]	PPP	Purchasing power parities for each country for the Basic Heading covered.
[16]	STD 3	Standard deviation of each country's CPD or CPRD residuals for the Basic Heading

ITEM LEVEL DETAILS		
[17]	PLI	Price Level Indices for each country for the Basic Heading covered.
[18]	No. of Priced Items	Number of products that are priced by each country for the Basic Heading.
[19]	Item Code	Codes of the items covered by the table.
[20]	Item Name	Names of the items covered by the table.
[21]	CPD Residuals	CPD residuals by product and country. See section 4.2 *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 the average prices at [22] are based.
[24]	Coefficient of Variation	Price observation variation coefficient for each country.
[25]	XR Ratio	Standardized price ratios based on the exchange rate converted prices. The converted prices expressed as a percentage of their geometric mean.

3.4 Use of CPD Residuals

CPD residuals are used throughout the Dikhanov table. As shown in Annex 1, the residuals from CPD regressions are presented as follows:

$$\varepsilon_{cp} = \ln p_{cp} - x_{cp}\beta = \ln p_{cp} - Dc_c - Dp_p$$

where Dc_c and Dp_p are the country and product dummies.

During the validation, CPD residuals and STD of CPD residuals should be analyzed in a similar manner to the way PPP-Ratios and variation coefficients are analyzed in the Quaranta Tables. The process of data validation with the DT 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 Venezuela is very low compared to the price in other Latin American countries, therefore the large deviation (CPD residual) for the product as shown in the DT at the GDP level is not a mistake in data. In general, the overall STD by country should reflect the quality of 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 their beverages in gallons instead of liters, its basic heading level data alone could be very consistent (as everything is priced in gallons), but

inconsistent with all other countries (that priced in liters). Tables 7A and 7B give examples of the Dikhanov tables that are calculated at BH Level (7A) and at GDP level (7B). Table 7A has the same problem for that basic heading as the QT. However, Table 7B explicitly shows that inconsistency.

It is indeed recommended to run the DT at different levels of aggregation[6], for example: at the basic heading level; for a higher aggregate level (such as Food),;and at the GDP level. Processing the DT 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 STD of residuals (the upper right corner of the table): Table 7A predictably shows a smaller value than Table 7B (0.71 vs. 0.80)[7]. However, the difference is not large (and it is even smaller for cases with data of poorer quality in which case the within basic heading deviations would dominate the between basic heading ones). The overall STD of residuals for larger regions with more variety tends to be larger than that for smaller and more uniform ones.

It is informative to study STDs of residuals both by country and by item. The same overall picture can be observed here as he STD values in Table 7B are greater than those in Table 7A.

As with the QT, it may be concluded that the goal of the validation is to reduce the overall STD 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.

[6] The difference between the residuals run at different levels of processing such as the basic heading and GDP can be expressed as:

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

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

$$\varepsilon_{cp}(BH) - \varepsilon_{cp}(GDP) = \{\ln(PPP_c(GDP)) - \ln(PPP_c(BH))\} + \{Dp_p(BH) - Dp_p(BH)\}$$

The first component is the difference between the basic heading PPP and the GDP PPP (the relative price level of the basic heading vis-à-vis the overall price level at the GDP level) while 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.

[7] The overall STD of residuals is run on the whole country-product tableau of the CPD residuals.

Table 7A: An example of a collapsed Dikhanov Table, calculated at BH level

Dikhanov Temporal Analysis	Country 1	Country 2	Country 3	Country 4	Country 5	Country 6	STD	Count
	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005		
PPP	1.000	4.066	0.049	3.442	8.548	4.727		
STD	0.502	0.728	0.750	0.711	0.800	0.717	0.705	
No.of Priced Items	706	747	562	766	661	691		844
ER (LCU/US$)	88.600	527.468	5.110	527.468	1138.000	393.383		
Rebased_XR	1.000	5.953	0.058	5.953	12.844	4.440		
PLI	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	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005		
1101111	**Rice**								
	PPP	1.000	5.504	0.065	1.480	8.052	1.579		
	STD	0.596	0.640	0.441	0.634	0.614	0.653		
	PLI	1.000	0.924	1.120	0.249	0.627	0.356		
	No.of Priced Items	7	7	3	7	5	7		
110111101	Long-grained rice	-0.176	0.582	-	-0.402	0.406	-0.410	0.416	5
110111102	Long-grained rice	-0.665	0.698	0.502	-0.231	-	-0.305	0.515	5
110111103	Long-grained rice	-0.288	-0.400	-0.571	0.886	-0.541	0.913	0.643	6
110111104	Medium-grained rice	0.202	-1.157	-	0.071	0.802	0.082	0.638	5
110111105	Short-grained rice	-0.667	-0.464	-	0.976	-0.863	1.018	0.824	5
110111106	Basmati Rice	1.094	0.205	0.068	-0.722	-	-0.645	0.660	5
110111107	Broken rice	0.501	0.536	-	-0.579	0.195	-0.654	0.518	5
1101112	**Other cereals, flour and other products**								
	PPP	1.000	3.531	0.052	3.274	4.962	6.010		
	STD	0.496	0.774	0.806	0.726	1.062	0.704		
	PLI	1.000	0.593	0.905	0.550	0.386	1.354		
	No.of Priced Items	14	20	13	21	11	19		
110111201	Wheat Flour	0.221	0.285	-0.371	0.357	-0.256	-0.236	0.293	6
110111202	Wheat Flour	0.419	0.122	-0.328	0.313	-0.087	-0.439	0.315	6
110111203	Couscous	-	0.311	-0.288	0.291	-	-0.314	0.302	4
110111204	Couscous (millet)	-	-1.450	-	0.986	-	0.465	1.047	3

Table 7B: An example of a collapsed Dikhanov Table, calculated at GDP level

Dikhanov Temporal Analysis	Country 1	Country 2	Country 3	Country 4	Country 5	Country 6	STD	Count
	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005		
PPP	1.000	4.066	0.049	3.442	8.548	4.727		
STD	0.623	0.812	0.839	0.785	0.893	0.817	0.797	
No.of Priced Items	706	747	562	766	661	691		844
ER (LCU/US$)	88.600	527.468	5.110	527.468	1138.000	393.383		
Rebased_XR	1.000	5.953	0.058	5.953	12.844	4.440		
PLI	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	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005	Yearly - 2005		
110111101	Long-grained rice	0.163	1.224	-	-0.906	0.686	-1.167	0.914	5
110111102	Long-grained rice	-0.392	1.273	1.050	-0.802	-	-1.128	0.979	5
110111103	Long-grained rice	-0.051	0.140	-0.058	0.279	-0.364	0.054	0.200	6
110111104	Medium-grained rice	0.541	-0.515	-	-0.433	1.082	-0.675	0.689	5
110111105	Short-grained rice	-0.328	0.179	-	0.472	-0.583	0.261	0.392	5
110111106	Basmati Rice	1.366	0.780	0.616	-1.293	-	-1.469	1.156	5
110111107	Broken rice	0.841	1.178	-	-1.084	0.475	-1.410	1.047	5
110111201	Wheat Flour	0.293	0.216	-0.237	0.379	-0.728	0.077	0.380	6
110111202	Wheat Flour	0.491	0.052	-0.194	0.335	-0.558	-0.127	0.347	6
110111203	Couscous	-	0.142	-0.254	0.214	-	-0.102	0.188	4
110111204	Couscous (millet)	-	-1.608	-	0.919	-	0.688	1.141	3

3.5 Modified Dikhanov Table

Another variation of the Dikhanov Table is called modified Dikhanov Table. This table keeps the original presentation of the DT, either collapsed or extended, but instead of using the DT specific indices, it uses the indices used in the Quaranta Table. Table 8 below summarizes the differences between the original and modified DT. Further, Table 9 gives an example of a Modified Dikhanov Table.

Table 8: Indices used in the original and modified DT

Dikhanov Table	Modified Dikhanov Table
CPD Residuals	PPP-Ratios
Standard Deviation	Variation Coefficient
Price Level Index (absolute)	Price Level Index (percentage)
# of priced items	# prices items and # important items

The key advantage of the modified DT is that users familiar with QT indices can benefit from the lay-out of the DT as well as from the option to calculate the tables and the respective PPPs at different levels of aggregation.

Table 9: An example of a Modified Dikhanov Table, calculated at BH level

Dikhanov Temporal Analysis (Calc. Method: CPD)		Country 1	Country 2	Country 3	Country 4	Country 5	Var.Co.	Count
		Q12011	Q12011	Q12011	Q12011	Q12011		
	PPP	3942.845296	11.33127847	1	11.76329728	16.74579211		
	Variation Coefficient	53.8	85.8	72.2	58.5	533.5	129.8	
	No.of Priced Items	941;*712	604;*112	441;*429	925;*501	551;*472		1027
	ER (LCU/US$)	1193.5	5.9121	0.3845	3.5644	3.65		
	Rebased_XR	3104.031209	15.37607282	1	9.270221066	9.492847854		
	PLI(%)	127.023	73.694	100	126.893	176.404		
Item Level Details		**Country 1**	**Country 2**	**Country 3**	**Country 4**	**Country 5**	**Var.Co.**	**Count**
Item Code	**Item Name**	Q12011	Q12011	Q12011	Q12011	Q12011		
1101111	**Rice**							
	PPP	2167.57	9.93324	1	10.6217	8.58586		
	Variation Coefficient	28.6	41.4	19.5	31.9	33.1		
	PLI(%)	69.831	64.602	100	114.579	90.446		
	No.of Priced Items	8;*5	6;*2	7;*7	11;*5	8;*3		
110111101	**Long grain rice - Parboiled**	**-**	151.49	109.4	67.73	89.09	34.2	4
	Average Price	-	18	1.30859	8.60526	9.15		
	No.of Observations	-	*14	*17	*19	5		
	Variation Coefficient	-	6	17.5248	26.18	1.49651		
	XR Ratio	-	108.192	120.94	85.791	89.0825		
110111102	**Long grain rice - Non-Parboiled**	70.66	-	107.99	113.2	115.77	20.7	4
	Average Price	1910.53	-	1.34716	15	12.4		
	No.of Observations	-	-	*14	1	5		
	Variation Coefficient	-	-	2.8127	-	3.37363		
	XR Ratio	53.4946	-	117.085	140.631	113.529		
110111103	**Long grain rice - Family Pack**	159.34	-	82.17	76.37	-	43.7	3
	Average Price	12738.1	-	3.03063	29.9167	-		
	No.of Observations	-	-	*12	*24	-		
	Variation Coefficient	-	-	7.01726	12.3178	-		
	XR Ratio	119.858	-	88.5159	94.2567	-		

Annex 1: CPD Residuals in Dikhanov Table

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

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

$$x_{cp} = \left[Dc_2 \ldots Dc_{Nc} Dp_1 Dp_2 \ldots Dp_{Np} \right]$$
$$\beta = \left[\alpha_2 \ldots \alpha_{Nc} \gamma_1 \gamma_2 \ldots \gamma_{Np} \right]^T \quad (2)$$

where p_{cp} - price of product **p** in country **c**;

Dc_j and Dp_i - country and product dummies;

Np and *Nc* – number of products and countries, respectively;

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

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3)$$

Note that the first country dummy is dropped from the system because matrix **X** is of rank (*Np*+*Nc*-1). In fact, we can drop any variable 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

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \mathbf{y} \quad (4)$$

Using expression [3] we can present the error term as:

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

Those error terms [residuals] enter DT and are used in its statistics such as the standard deviations of residuals by country and product.

Note that 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, non-tradables, unprocessed food etc. In those cases the residuals will indicate the variability within those groups only.