

A faint, dotted world map serves as the background for the slide.

Estimation of annual average prices

Missing Data Imputation

June 27-29, 2017
ECLAC, Santiago de Chile



Missing Data Imputation

- Calculation of PPP relies in annual average price of products (arithmetic mean of the four quarters)
- How to impute missing data for products?
 - An association between National Costumer Price Index and the ICP list is used to impute missing values for products / quarters

Why is imputation important?

- Besides usual difficulties to collect data for all products, imputation will be essential in the 2017 Cycle, since price collection will start only in the second semester of the reference year
- Proposed survey calendar:

Even if the country collects prices for all the products of the list, imputation will be necessary to estimate values to complete the reference year, 2017

2017	2018
-	30-50% of products
-	30-50% of products
100% of products	<i>Rolling benchmark to be decided...</i>
100% of products	

Same criteria for the Rolling Benchmark, from 2018 onwards, that DOES NOT collect prices for all the products at each period.

Required Data

- Table with the greatest possible disaggregation of the monthly National CPI.

Code	Name of product	jan.11	feb.11	mar.11	average Q1 11	apr.11	may.11	jun.11	average Q2 11	jul.11	aug.11	sep.11	average Q3 11	Oct.11	Nov.11	Dec.11	average Q4-11
0	All Products	104,36	104,57	105,67	104,87	105,94	106,38	107,6	106,64	107,59	107,80	107,6	107,67	107,57	107,64	107,76	107,66
1	GROSS DOMESTIC PRODUCT	104,36	104,57	105,67	104,87	105,94	106,38	107,6	106,64	107,59	107,80	107,6	107,67	107,57	107,64	107,76	107,66
11	INDIVIDUAL CONSUMPTION EXPENDITURE BY HOUSEHOLD	104,36	104,57	105,67	104,87	105,94	106,38	107,6	106,64	107,59	107,80	107,6	107,67	107,57	107,64	107,76	107,66
11,01	FOOD AND NON-ALCOHOLIC BEVERAGES	107,13	107,25	108,73	107,72	109,59	110,07	110,5	110,08	110,99	111,59	111,5	111,37	112,21	112,40	112,70	112,44
11.01.1	FOOD	107,39	107,48	109,04	107,97	110,06	110,53	111,1	110,58	111,60	112,25	112,1	112,01	112,92	113,12	113,45	113,17
11.01.11	Bread and cereals	101,59	101,60	101,71	101,64	104,40	105,16	105,2	104,94	104,93	105,33	105,5	105,26	106,13	106,13	106,27	106,18
11.01.11.1	Rice	105,80	105,56	107,43	106,28	107,40	109,79	106,9	108,05	106,89	107,31	107,2	107,16	107,87	107,85	106,87	107,53
11.01.11.2	Other cereals, flour and other products	100,57	100,27	98,83	99,91	104,95	105,81	106,4	105,74	106,06	106,97	106,8	106,63	107,42	107,11	107,30	107,28
11.01.11.3	Bread	100,87	100,92	101,30	101,03	101,34	101,55	102,6	101,83	101,92	101,92	101,8	101,90	101,87	102,22	102,31	102,13
11.01.11.4	Other bakery products	101,87	101,55	104,59	102,67	104,38	104,54	105,5	104,83	105,78	105,56	106,8	106,08	108,75	109,50	110,11	109,45
11.01.11.5	Pasta products	99,89	102,36	99,60	100,62	103,80	102,58	102,2	102,87	102,06	102,08	103,3	102,50	104,62	103,87	106,28	104,92
11.01.12	Meat	110,45	112,95	113,04	112,15	113,99	113,52	114,3	113,95	116,05	116,05	115,1	115,74	116,18	116,21	116,63	116,34
11.01.12.1	Beef and veal	116,41	119,62	122,64	119,56	124,16	127,29	125,4	125,62	123,52	123,74	124,2	123,84	123,71	124,19	125,68	124,52
11.01.12.2	Pork	113,89	115,59	115,61	115,03	117,11	118,93	116,3	117,47	119,31	118,03	117,2	118,19	117,82	115,80	118,71	117,44
11.01.12.3	Lamb, mutton and goat	115,84	127,90	126,81	123,52	126,81	119,52	119,5	121,95	121,37	125,84	125,8	124,35	126,13	124,46	131,75	127,44
11.01.12.4	Poultry	109,01	111,48	111,31	110,60	112,14	110,99	112,5	112,5	112,5	112,5	112,5	112,5	114,05	114,42	113,78	114,08
11.01.12.6	Delicatessen and other meat preparations	110,22	109,96	110,65	110,28	111,94	114,13	114,8	114,8	114,8	114,8	114,8	114,8	117,91	118,23	118,88	118,34
11.01.13	Fish and seafood	105,05	104,24	108,32	105,87	108,68	109,31	109,1	109,03	110,17	110,88	110,7	110,61	110,95	111,81	112,55	111,77
11.01.13.1	Fresh, chilled or frozen fish and seafood	101,34	100,77	104,44	102,18	103,79	104,08	104,0	103,98	105,55	105,89	105,8	105,77	105,89	105,89	106,39	106,05
11.01.13.2	Preserved or processed fish and seafood	114,03	112,61	114,40	113,68	116,42	117,68	117,1	117,08	117,52	118,84	118,6	118,32	119,00	121,32	122,49	120,94
11.01.14	Milk, cheese and eggs	106,80	106,19	111,00	108,00	112,57	111,60	111,3	111,85	111,65	112,88	113,4	112,67	114,76	116,30	116,63	115,90
11.01.14.2	Preserved milk and other milk products	103,34	102,67	106,10	104,04	107,40	106,08	106,0	106,51	105,98	107,46	108,2	107,24	109,68	110,89	111,20	110,59
11.01.14.3	Cheese	132,75	132,60	135,85	133,73	140,07	140,12	139,4	139,88	141,60	141,54	141,5	141,58	143,39	146,59	147,26	145,75
11.01.14.4	Eggs and egg-based products	110,43	109,79	109,43	109,90	109,15	109,70	108,5	109,12	108,91	109,84	109,3	109,38	108,72	110,47	110,44	109,88
11.01.15	Oils and fats	107,47	108,26	110,42	108,72	113,03	114,05	115,8	114,32	115,45	116,27	116,4	116,04	117,16	118,77	121,10	119,01
11.01.15.1	Butter	120,29	120,41	121,55	120,75	122,21	123,89	124,2	123,44	124,10	124,40	124,2	124,25	124,12	123,95	128,13	125,40
11.01.15.2	Margarine	118,10	117,85	120,03	118,68	125,31	126,52	128,3	126,73	128,68	128,09	128,8	128,53	128,95	129,11	131,52	129,86
11.01.15.3	Other edible oils and fats	102,28	103,39	104,87	103,51	107,39	108,08	110,3	108,60	109,63	110,96	111,0	110,54	112,16	114,61	116,40	114,39
11.01.16	Fruit	110,14	110,41	110,37	110,31	109,41	109,61	112,5	110,52	106,42	107,32	106,3	106,68	107,69	107,75	109,28	108,24
11.01.16.1	Fresh or chilled fruit	111,06	111,31	110,81	111,06	111,03	110,61	114,5	112,06	107,10	107,91	106,4	107,16	108,10	107,88	109,57	108,52
11.01.16.2	Frozen, preserved or processed fruit and fruit-based proc	100,23	100,67	107,70	102,90	100,25	103,79	101,2	101,78	102,46	103,87	105,3	103,88	105,31	107,00	107,52	106,61

Quarterly averages

Required Data

- Make full correspondence between each ICP Product and the most disaggregated National CPI Product

ICP Code	ICP Product Name	Code	Product Name
110111101.LAC	Long grain rice - parboiled	11.01.11.1	Rice
110111102.LAC	Long grain rice - non-parboiled	11.01.11.1	Rice
110111107.LAC	Long grain rice - loose	11.01.11.1	Rice
110111108.LAC	Long grain rice - family pack	11.01.11.1	Rice
110111109.LAC	Jasmine rice	11.01.11.1	Rice
110111110.LAC	Basmati rice	11.01.11.1	Rice
110111112.LAC	White rice, medium grain - loose	11.01.11.1	Rice
110111113.LAC	Brown rice - family pack	11.01.11.1	Rice
110111201.LAC	Cornflakes (Kellogg s)	11.01.11.2	Other cereals, flour and other products
110111202.LAC	Wheat flour, not self-rising	11.01.11.2	Other cereals, flour and other products
110111204.LAC	Oats, rolled	11.01.11.2	Other cereals, flour and other products
110111206.LAC	Baby cereals	11.01.11.2	Other cereals, flour and other products
110111209.LAC	Wheat semolina (suji)	11.01.11.2	Other cereals, flour and other products
110111210.LAC	Maize flour (white)	11.01.11.2	Other cereals, flour and other products
110111211.LAC	Couscous	11.01.11.2	Other cereals, flour and other products
110111301.LAC	Baguette	11.01.11.3	Bread
110111302.LAC	Whole wheat bread	11.01.11.3	Bread
110111303.LAC	Roll	11.01.11.3	Bread
110111304.LAC	Sliced white bread	11.01.11.3	Bread
110111305.LAC	Pita bread	11.01.11.3	Bread
110111306.LAC	White bread	11.01.11.3	Bread

Best and most disaggregated match in National CPI

Guidelines

- Missing prices are imputed using the CPI variation of the corresponding product according to the formula:

$$P_m = P_d * \frac{CPI_m}{CPI_d}$$

P_m : missing quarter price


P_d : “donor” quarter price

CPI_m : mean CPI for missing quarter

CPI_d : mean CPI for “donor” quarter

Guidelines

- If options are available, which is going to be the `donor` quarter?
- General assumption:
 - Data that was collected latter is more reliable, since all the people involved in the ICP and data collection had more training and experience
 - Thus, to impute missing data, it is recommended to use as reference the prices collected after the missing quarter.
 - Example:



Product	II Quarter	III Quarter	IV Quarter
Beef Liver	15.42	?	15.50

The diagram illustrates the recommended method for imputing missing data. A blue arrow points from the IV Quarter (15.50) to the III Quarter (?), indicating that data from the later quarter is used as a reference. A red 'X' is placed over a blue arrow pointing from the II Quarter (15.42) to the III Quarter (?), indicating that data from the earlier quarter is not used for imputation.

Example

- Grenada's imputation for rice
- All varieties were associated to the same CPI Product (rice), thus imputed variation is equal

ICP Code	ICP Product Name	Code	Product Name	Average National Indexes								Imputed data and final averages				
				Q1-11	Q2-11	Q3-11	Q4-11	Q1-12	Q2-12	Q3-12	Q4-12	Q4-11	Q1-12	Q2-12	Q3-12	Q4-12
11011101.LAC	Long grain rice - parboiled	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77	4,29	4,27	4,28	4,13	4,85
11011102.LAC	Long grain rice - non-parboiled	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77	4,14	4,05	3,94	3,86	4,76
11011107.LAC	Long grain rice - loose	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77	3,68	3,37	4,14	4,14	3,90
11011108.LAC	Long grain rice - family pack	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77	19,01	18,91	18,91	19,20	18,97
11011109.LAC	Jasmine rice	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77	95,35	94,85	107,11	107,11	97,01
11011110.LAC	Basmati rice	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77	15,82	15,73	16,44	20,26	16,31
11011112.LAC	White rice, medium grain - loose	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77					
11011113.LAC	Brown rice - family pack	11.01.11.1	Rice	106,28	108,05	107,16	107,53	106,97	108,09	108,18	106,77					

$$P_m = P_{Q1-12} * \frac{107,53}{106,97}$$

Seasonal Products

- Some products might have no prices collected in a certain quarter because they did not exist in the market or they were scarce or not representative at the moment of price collection (seasonal products)
- For these products, no imputation is made, even if there is an available donor quarter
- Examples of seasonal products: seasonal fruits and vegetables, seasonal garments
- Countries must clearly inform missing values related to seasonality

Key Points

- Imputation will be a relevant part of the final database for ICP-PPP calculations
- A good correspondence between ICP and CPI is required
- The more disaggregated the CPI, the better
- The more integrated is CPI to IPC, the better



Thank You



International Comparison Program