



GEORGIA

TECHNICAL ASSISTANCE REPORT—RESIDENTIAL PROPERTY PRICE INDICES

May 2020

This Technical Assistance Report paper on Georgia was prepared by a staff team of the International Monetary Fund. It is based on the information available at the time it was completed on March 16, 2020.

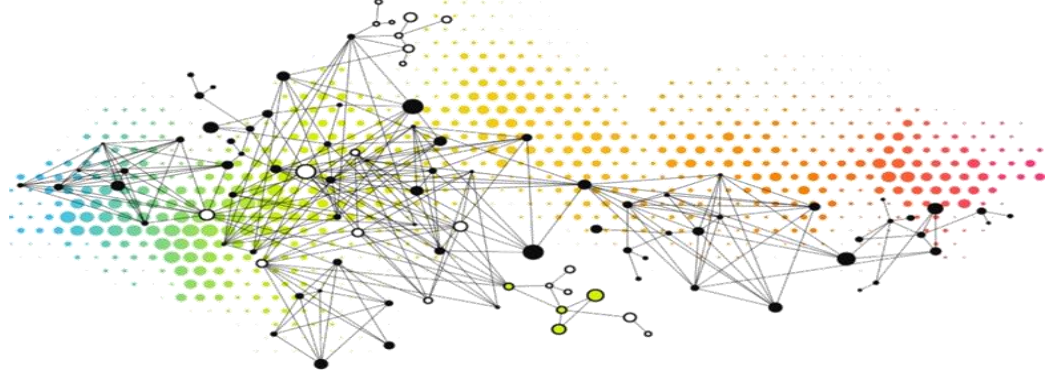
Disclaimer:

This document was prepared before COVID-19 became a global pandemic and resulted in unprecedented strains in global trade, commodity, and financial markets. It, therefore, does not reflect the implications of these developments and related policy priorities. We direct you to the [IMF Covid-19 page](#) that includes staff recommendations with regard to the COVID-19 global outbreak.

Copies of this report are available to the public from

International Monetary Fund • Publication Services
PO Box 92780 • Washington, D.C. 20090
Telephone: (202) 623-7430 • Fax: (202) 623-7201
E-mail: publications@imf.org Web: <http://www.imf.org>
Price: \$18.00 per printed copy

**International Monetary Fund
Washington, D.C.**



GEORGIA

DECEMBER
2019

REPORT ON RESIDENTIAL PROPERTY PRICE INDICES MISSION (SEPTEMBER 23–OCTOBER 4, 2019)

Prepared by Vanda Guerreiro

The contents of this report constitute technical advice provided by the staff of the International Monetary Fund (IMF) to the authorities of Georgia (the "TA recipient") in response to their request for technical assistance. This report (in whole or in part) or summaries thereof may be disclosed by the IMF to IMF Executive Directors and members of their staff, as well as to other agencies or instrumentalities of the TA recipient, and upon their request, to World Bank staff, and other technical assistance providers and donors with legitimate interest, unless the TA recipient specifically objects to such disclosure (see [Operational Guidelines for the Dissemination of Technical Assistance Information](#)). Publication or Disclosure of this report (in whole or in part) or summaries thereof to parties outside the IMF other than agencies or instrumentalities of the TA recipient, World Bank staff, other technical assistance providers and donors with legitimate interest, shall require the explicit consent of the TA recipient and the IMF's Statistics Department.

CONTENTS

Glossary	3
SUMMARY OF MISSION OUTCOMES AND PRIORITY RECOMMENDATIONS	4
DETAILED TECHNICAL ASSESSMENT AND RECOMMENDATIONS	6
A. Governance	6
B. RPPI for Flats	8
C. RPPI for Detached Houses	10
D. Scanner Data	12
E. Officials Met During the Mission	14
FIGURES	
1. Overall RPPI by Method	5
2. Breakdown of the RPPI Compiled with Online Data	6
3. Histograms for Flats Before and After Data Treatment—4Q 2018: Prestigious and Non-Prestigious	9
4. RPPI for Flats and Sub-Indices	10
5. Histograms for Detached Before and After Data Treatment - 4Q 2018: Prestigious and Non-Prestigious	11
6. RPPI for Detached and Sub-Indices	12
TABLES	
1. Priority Recommendations	5
2. Number of Observations and Value by Strata for Flats	6
3. Parameters for Model Selection	8
4. Number of Observations and Number of Variables for Flats	8
5. Thresholds and Categories for Flats	9
6. Number of Observations and Number of Variables for Detached	10
7. Thresholds and Categories for Detached	11
APPENDIX	
1. Files Management	15

Glossary

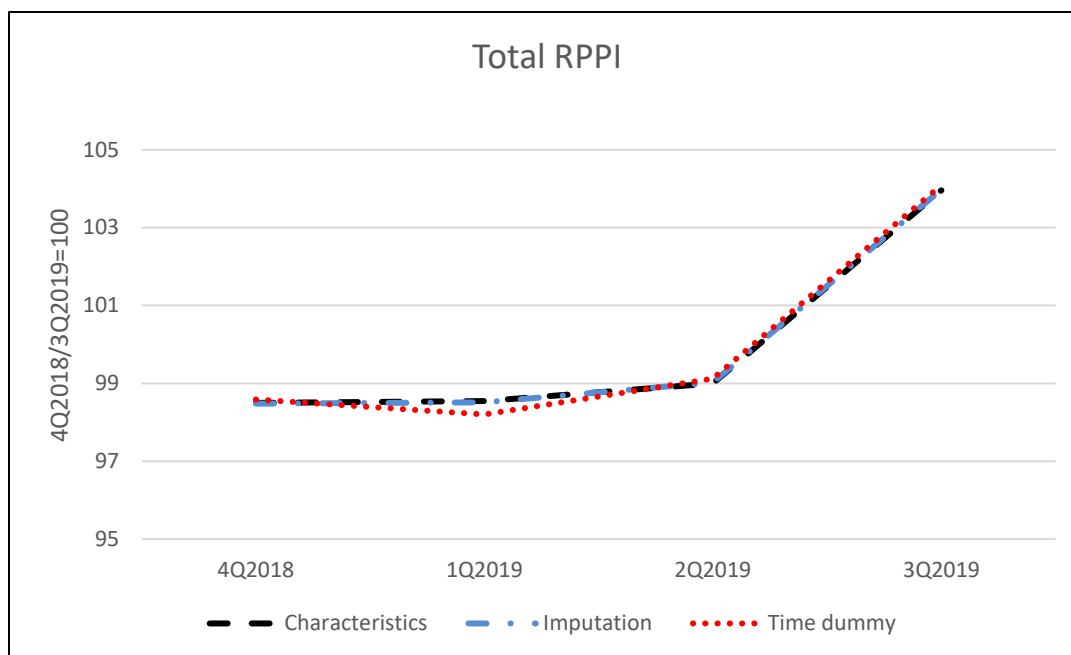
AIC	Akaike Criterion
ANOVA	Analysis of Variance
COICOP	Classification of Individual Consumption According to Purpose
Geostat	National Statistics Office of Georgia
NAPR	National Agency of Public Registry of Ministry of Justice
OLS	Ordinary Least Squares
RPPI	Residential Property Price Index
RRSE	Robust Residual Standard Errors
SD	Scanner Data

SUMMARY OF MISSION OUTCOMES AND PRIORITY RECOMMENDATIONS

- 1. The purpose of the mission was to assist the National Statistics Office (Geostat) in progressing on the compilation of a residential property price index (RPPI).** This was the third price statistics mission to Georgia conducted under the auspices of the three-year *Project to Improve National Accounts and Price Statistics in Eastern and Southeastern Europe*. This project is funded by the Government of The Netherlands. The aim of the first mission, held during April 25 to May 4, 2018 was to assist Geostat in obtaining data for RPPI compilation. The second mission held during in December 2018 focused on the assessment of the data obtained and training Geostat staff on the RPPI compilation. The aim of the current mission was to assist Geostat in compiling an experimental RPPI.
- 2. The Second Phase of the G-20 Data Gaps Initiative and guidance on Financial Soundness Indicators identify RPPI as a critical ingredient of financial stability policy analysis and macroprudential measures.** In addition, an RPPI is on its own a macro-economic indicator of growth and a key indicator for understanding financial market conditions.
- 3. Geostat is aiming at compiling a quarterly RPPI covering new flats and new detached houses for the capital city, Tbilisi.** The mission recommended to obtain all data including the secondary market and rest of the country to be stored for future expansion of the index coverage.
- 4. Following the recommendations from the previous mission, Geostat obtained data on dwellings transactions from a second web site with advertised properties for sale.** In addition, there was an improvement on the web scraping of data from the first web site. Data sets cover all types of new dwellings in Tbilisi.
- 5. Data from the National Agency of Public Registry of Ministry of Justice (NAPR) began being automatically transmitted from October 2019.** The NAPR data cover only Tbilisi, including flats and land plots with unspecified buildings. The data are very incomplete with regards of the dwelling's characteristics and Geostat should continue negotiations with NAPR to improve the survey.
- 6. The mission implemented successfully the programs developed in R based on the IMF's draft for the *RPPI Practical Compilation Guide* with the available data.** The three hedonic methods compiled during the mission should continue being tested during 2019. Geostat staff is capable to perform changes and adaptations of the R scripts if needed.
- 7. Data from the first website are available for four quarters since the fourth quarter of 2018.** After the second quarter of 2019 the data have more observations and variables due to the improvement on the web scraping technique used. The experimental indices were compiled

using the initial data structure, the 4Q2018 as the base period and, referencing the indices for the available four quarters. The results can be seen in Figure 1.

Figure 1. Overall RPPI by Method



8. The mission provided some guidance on the use of scanner data (SD) on the CPI compilation. As per request of the Geostat Director, the mission also addressed the use of SD. The introduction of SD should be made on a step-wise approach to avoid huge impacts on the CPI and to make it more manageable, reliable, and safe.

Table 1. Priority Recommendations

Target Date	Priority Recommendation	Responsible Institutions
December/2020	<i>Compile the RPPI with data from the two websites and assess the results.</i>	Geostat
December/2020	<i>Continue negotiations with NAPR to improve the survey.</i>	Geostat
October/2019	<i>Obtain data for the secondary market.</i>	Geostat

Further details on the priority recommendations and the related actions/milestones can be found in the action plan under *Detailed Technical Assessment and Recommendations*.

DETAILED TECHNICAL ASSESSMENT AND RECOMMENDATIONS

A. Governance

9. **Geostat is aiming at compiling a quarterly RPPI covering new flats and new detached houses for Tbilisi.** The mission recommended to obtain all data including secondary market and rest for the country and store it for future expansion of the index coverage. The index structure should be as follows:

Figure 2. Breakdown of the RPPI Compiled with Online Data

RPPI	Apartments	Prestigious
		Non-Prestigious
Detached houses	Prestigious	
	Non-Prestigious	

10. **The stratification by location was made by clustering the district areas according to being more or less prestigious.** The allocation of a district to each stratum was made in cooperation with the sampling experts from Geostat. Dwellings were classified by 44 district areas and a cut-off approach was implemented to split between the two strata. The cut-off was based on the median price per square meter. An outlier detection was made at this stage using the coefficient of variation. The resulting number of observations by strata is exemplified in Table 2 for Flats.

Table 2. Number of Observations and Value by Strata for Flats

Periods	Prestigious		Value Share	Non-Prestigious		Value Share
	No. Obs	Total value		No. Obs	Total value	
4Q2018	2091	661,362,304	0.72	1851	251,898,956	0.28
1Q2019	2926	2,460,546,608	0.89	2326	311,423,711	0.11
2Q2019	8061	3,458,924,804	0.80	6363	843,508,634	0.20
3Q2019	10694	3,458,924,804	0.76	7241	1,104,995,366	0.24
Total annual		10,039,758,520	0.80		2,511,826,667	0.20

- 11. A new staff member was added to the prices team to work on the RPPI.** Four statisticians are working on the RPPI: the head of the prices unit and three staff members who are also working on the CPI and PPI.
- 12. The mission developed a framework for the file management during the experimental period.** Files should be kept safely and following a comprehensive schema to be in line with the R scripts and to be easily tracked. Particularly during the experimental period this system is rather complex since the RPPI is being compiled with different methods and different data sources generating multiple versions of each file. Backups of all the files generated during the process are also highly recommended. The file management framework proposed by the mission can be found in Annex 1.
- 13. Following the recommendations from the previous mission, Geostat obtained data on dwellings transactions from a second web site.** Data cover all types of dwellings in Tbilisi for the primary and secondary market.
- 14. Data from the two web sites should be merged.** The data structure of the two web sites is similar for flats thus can be easily merged. In contrast, the data structure for detached houses is rather different. The second web site has many more variables. The significance of the extra variables should be investigated. If significant, a different model for each can be compiled, with the same method. The resulting sub-indexes of the web sites (per strata) are two separate indexes can be compiled and then aggregated with a weighted average. The weights should be obtained from the total value of the dwellings. When merging the data, the duplicates should be deleted.
- 15. Data from the National Agency of Public Registry of Ministry of Justice (NAPR) is being automatically transmitted from October 2019.** The NAPR data cover only Tbilisi, including flats and land plots with unspecified buildings. It is not possible to distinguish between primary and secondary market in the new data. The data structure has changed and the variable which identifies primary and secondary market is no longer available. Geostat will contact the NAPR in near future to clarify the content of the data to be transmitted. In addition, data are very limited with regards to the dwelling characteristics and Geostat should continue negotiations with NAPR to improve the survey.
- 16. The model and methodology should be kept fixed for at least one year.** It is recommended to have stability on the model and the method for five years. While for the weights, the average characteristics or, the base price or, the dwellings sample—depending on the methodology chosen—should be kept fixed for at least one year (ideally five) and be updated every first quarter.
- 17. The experimental RPPI was compiled with the data from the first web site.** The second web site began being web scraped in September therefore data were still not available.

18. Econometric procedures were made to obtain the best fitted model. An interactive procedure to estimate different models using robust residual standard errors (RRSE) deviation's transformation, to correct for heteroscedasticity, was made. The Akaike criterion (AIC) to evaluate the goodness of fit, then Analysis of Variance (ANOVA) were used for the comparison of the different models. The P-values and F-statistics confirm the best fit. In addition, different models were tried out providing very similar indices. The table below shows the results for the main parameters. The stratum Flats performs better since the number of observations is much higher.

Table 3. Parameters for Model Selection

Indicator	Detached	Flats
AIC	2007	3117
R ²	70	77
RRSE	0.54	0.28

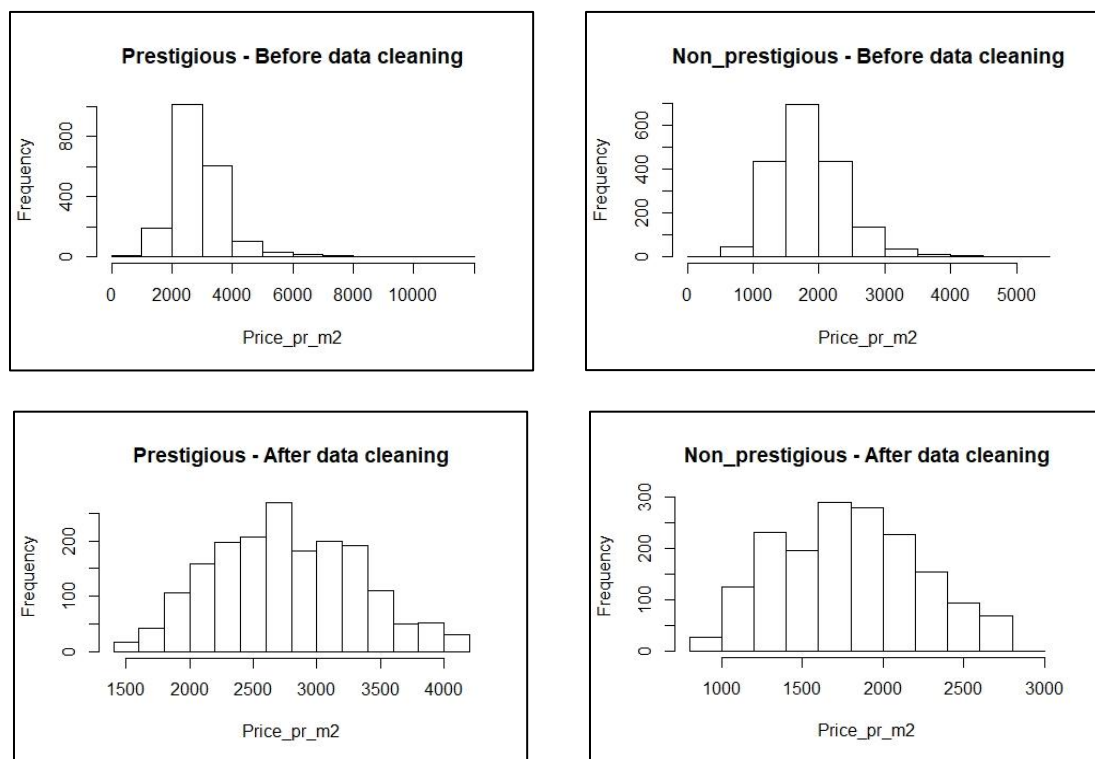
B. RPPI for Flats

19. The data were analyzed to treat the outliers, the missing values, and further consistency checks. The number of rooms was limited to between one and seven and the maximum number of floors was set to 40. The table and figures with the histograms below show the data analytics before and after data treatment. Around 28.6 percent of the observations were ruled out of the compilation which indicates good quality of the data. Table 4 shows a summary of the data and the histograms in Figure 3 are showing the results before and after the data have been cleaned for the two sub-strata.

Table 4. Number of Observations and Number of Variables for Flats

Periods	No. Obs before cleaning	No. Obs after cleaning	% of excluded Obs	Total no. of variables
4Q2018	3942	3345	15,1	11
1Q2019	5252	4778	9,0	11
2Q2019	14424	12344	14,4	20
3Q2019	17935	12809	28,6	20

Figure 3. Histograms for Flats Before and After Data Treatment—4Q 2018: Prestigious and Non-Prestigious



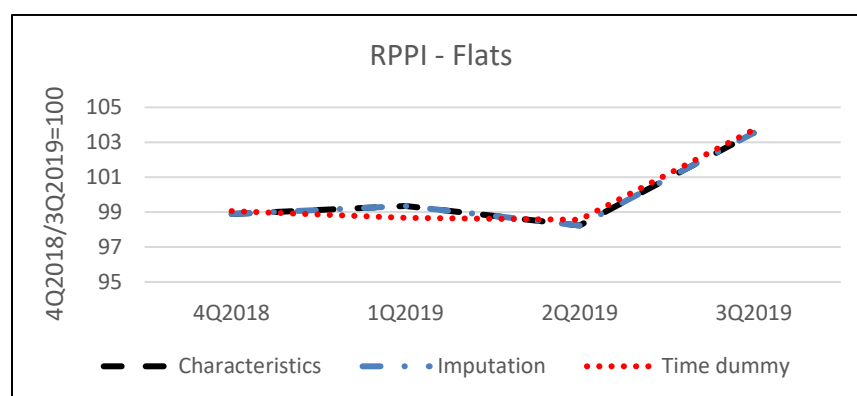
20. **Thresholds for the number of rooms and the number of floors were built and these variables were categorized.** These can be seen in the following table.

Table 5. Thresholds and Categories for Flats

Number of Rooms	Number of Floors
[1,2]	[1,4]
[3,4]	[5,9]
[5,7]	[10,16]
	> 17

21. **The RPPIs for the strata within the Flats are increasing particularly in the third quarter of 2019.** This is likely in consequence of the 15 per cent depreciation of the Georgian currency to the dollar. House prices are set in dollars although mostly published in local currency. Thus, house prices record an elasticity to the exchange rate close to one.

Figure 4. RPPI for Flats and Sub-Indices



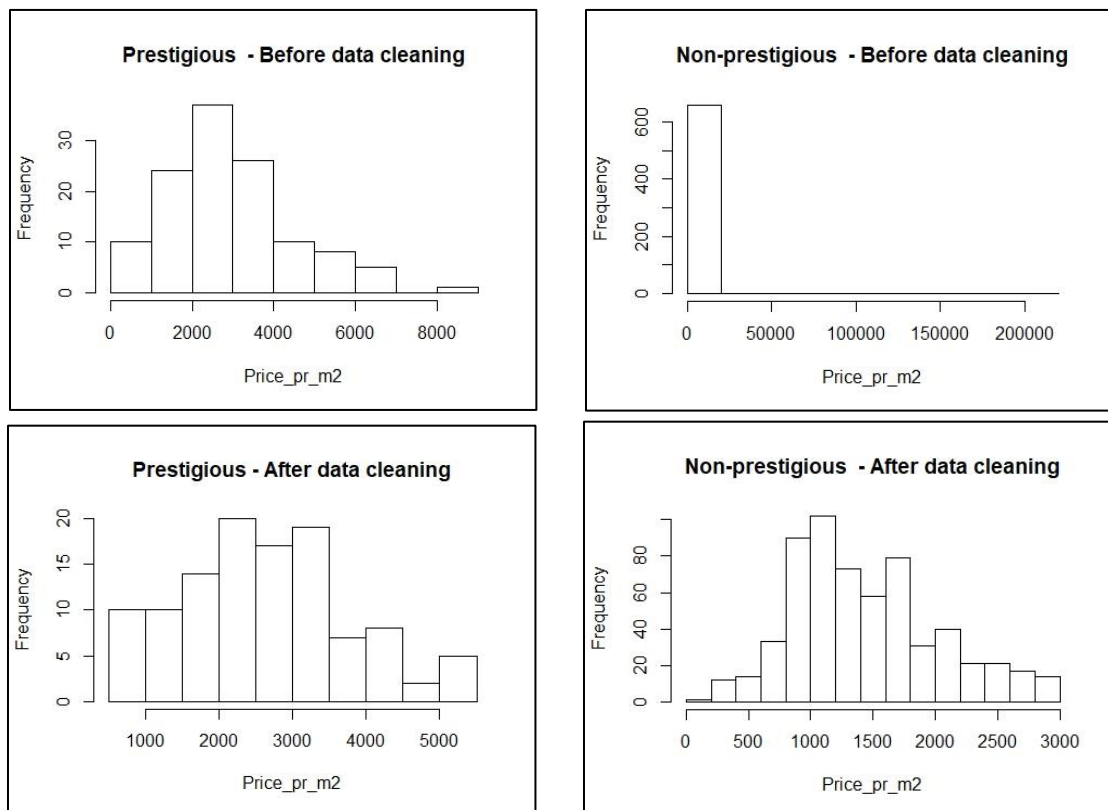
C. RPPI for Detached Houses

22. The data for detached houses were also analyzed to treat the outliers, the missing values, and further consistency checks. The number of rooms was limited to 15 and area per room is 200 sqm and land area between 15 and 6000. The table and figures with the histograms below show the data analytics before and after data treatment. Around 30 per cent of the observations were ruled out of the compilation which indicating good quality of the data.

Table 6. Number of Observations and Number of Variables for Detached

Periods	No. Obs before cleaning	No. Obs after cleaning	% of excluded Obs	Total no. of variables
4Q2018	1045	695	33,5	10
1Q2019	1212	857	28,9	10
2Q2019	2335	1556	33,0	19
3Q2019	2522	1741	30,1	19

Figure 5. Histograms for Detached Before and After Data Treatment—4Q 2018: Prestigious and Non-Prestigious



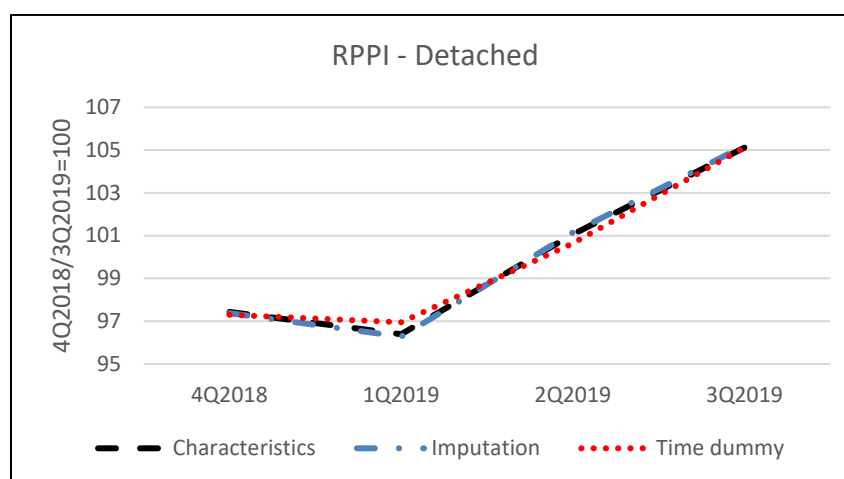
23. Thresholds for the number of rooms were built and this variable was categorized. The categories can be seen in the following table.

Table 7. Thresholds and Categories for Detached

Number of Rooms
[1,5]
[6,10]
[10,15]

24. The RPIs for the strata within Detached are increasing particularly in the third quarter of 2019.

Figure 6. RPPI for Detached and Sub-indices



D. Scanner Data

25. The mission provided some guidance on the use of scanner data (SD) on the CPI compilation.

26. The introduction of SD should be made on a step-wise approach to avoid huge impacts on the CPI and to make it more manageable, reliable, and safe. Food products should be the first group to be covered by SD because electronic systems are well developed in food retailers and also allows to cover and replace a significant part of the price surveys avoiding complex methodology issues as quality adjustments.

27. Negotiation with data providers can take a long time and the first meeting should aim at engaging them in the project. The statistical law should be used as a last resource. On the first meeting with the data providers Geostat should:

- inform of the progress of other countries on this field,
- assure the confidentiality,
- assure that the data will be used only for CPI and not shared with other authorities in the country,
- assure that details on the specific retailer, price levels, and average prices are not published.

In return for their cooperation Geostat can offer a personalized report.

To guarantee the confidentiality at least two retailers should be involved.

28. A memorandum for the meeting should be prepared explaining the aim of the project, the impact of SD on the quality of the CPI and should be very specific on the request. Two years of back data of all products should be requested. Data should be received

automatically weekly for at least the first two weeks of each month. The structure of the scanner data should be as follows:

- EAN/GTIN code
- Retailer category code
- Retailer category label
- Item label
- Sales/ turnover
- Number of units sold/Quantity
- Unit size (kg, lt)
- Package size
- Reference period (Year, month)

29. The classification to the COICOP is most challenging. A learning data set must be built by manually classification of the data sets received from the retailers. While a machine learning process for the use of semantic data is not developed, Geostat should focus on classifying items with high turnover.

30. The methodology proposed by the mission is rather simple to implement and takes good advantage of the SD. It is a matched model approach with a dynamic sample. The methodology is well described in the paper *The use of supermarket scanner data in the Luxembourg CPI*. The data of the current month is merged with the previous month to identify the products available in both months. After sampling is made using the share of each item. A filter for outliers and dumping prices are applied. These prices are imputed as well as the missing prices, for 14 months. Sub-indexes by retailer are compiled, at the lowest level of the COICOP, starting from monthly changes of the items available in two consecutive months; then multiply by the index-based December—of the previous month; finally chain by multiplying by the index-based 2010—of the previous month. The retailers are aggregated with the turnover weights to obtain an SD index per retailer per lower level COICOP that is afterward aggregated with the same level sub-index obtained by the field survey. The products covered by scanner will no longer be included in the field price collection for those retailers.

To support progress in the above work areas, the mission recommended a detailed one-year action plan with the following priority recommendations carrying particular weight to make headway in improving the RPPI:

Priority	Action/Milestone	Target Completion Date
Outcome: Experimental RPPI is compiled		
H	Compile the RPPI with data from the two websites to assess the results.	December 2020
H	Continue negotiations with NAPR to improve the survey	December 2020
H	Obtain all data including secondary market and rest for the country	October 2019
M	Built a comprehensive files management framework	October 2019
M	Methodology should be kept fixed for at least one year	December 2020
M	Merge data from the two websites or compile a sub-index for each	January 2020

E. Officials Met During the Mission

Name	Institution
Gogita Todradze	Executive Director, Geostat
Giorgi Tetrauli	Head of Price Statistics Department, Geostat
Khatuna Aptsiauri	Head of Consumer Price Statistics Division
Revaz Maisuradze	Senior Specialist, Consumer Price Statistics Division, Geostat

Appendix 1. Files Management

