TECHNICAL ASSISTANCE REPORT

GEORGIA

Data for Decisions (D4D) Residential Property Price Index (December 2–6, 2024)

APRIL 2025

Prepared By

Jens Mehrhoff

Authoring Department:

IMF Statistics Department

DISCLAIMER

Table of Contents

Acronyms

D4D	Data for Decisions (Fund)
Geostat	National Statistics Office of Georgia
HI	Hedonic Imputation (Index)
MoU	Memorandum of Understanding
RPPI	Residential Property Price Index
TA	Technical Assistance

Summary of Mission Outcomes and Priority Recommendations

1. **Technical assistance (TA) was provided to the National Statistics Office of Georgia (Geostat) during December 2–6, 2024 to extend the scope and coverage of the residential property price index (RPPI) using web-scraped data.** This was the latest mission to Georgia in a series since 2018 to specifically focus on the RPPI but the third one to be conducted under the Data for Decisions (D4D) Fund. Earlier missions have allowed Geostat to produce and disseminate an RPPI for new buildings using the information from one website.

2. **Reliable property price indexes and other indicators of real estate markets are critical ingredients for policymakers to assess the state of the real estate market and its potential impact on macroeconomic and financial stability.** Rapid increases in property prices may indicate the presence of a bubble that increases vulnerabilities in the financial system, while rapid decreases in property prices could signal an impending recession as households and businesses cut consumption and investment. The indexes are also used by policy makers as an input into the design of macroprudential policies, (i.e., policies aimed to reduce systemic risks in the financial sector) and to evaluate the effectiveness of monetary policy transmission.

3. **Previous missions assisted the authorities to analyze, process and aggregate data collected by the authorities from a web-based property website.** In addition, the hedonic approach to index calculation was applied to the data. This approach adjusts for the varying mix of properties in the sample from period to period, allowing for the calculation of a constant quality price index. Initial experimental indices were developing.

4. **Geostat should continue using web-scraped data to adjust for the property mix using a wide range of characteristics, thus improving the quality of the index.** The mission assisted in the development of a fully automated system without the need for user intervention or any manual cleaning or adjustment of the scraped results. The scraper covers one website (www.ss.ge) and both new and existing dwellings (flats and detached houses) for sale in Tbilisi.

5. **The new index production system will be used for the first time with the publication of the RPPI for the first quarter of 2025.** In particular, the mission helped to modernize the index calculation process and update the methods for index calculation. This included tidying the scraped data, removing outliers (and errors), model selection and calculating weights and medians, and implementing the hedonic imputation approach. The previous approach was largely outdated and would have needed a review even without a major upgrade of the data sources.

6. **A second website shares their database directly with Geostat.** They will need to establish a Memorandum of Understanding (MoU) with the new website (myhome.ge) in order to firmly agree upon the data delivery as well as fix the structure (fields, names, temporal coverage, etc.). At the same time, Geostat should inquire with ss.ge whether their database can be shared too. A follow-up mission will further assist Geostat in extending the data sources.

7. **Once the MoU with myhome.ge is in place, their data should be integrated as a second source into the RPPI.** This requires an in-depth data analysis, similar to the steps for ss.ge. Further, information on rental prices is currently being collected but not used for producing a rental price index. Likewise, this requires analyzing the data in order to build both the rental price index as well as the

valuation indicator price-to-rent ratio. Lastly, second-hand properties are not yet included in the RPPI. More testing is required to ascertain whether this subset can form part of the RPPI without undermining its quality.

8. To support progress in the above work areas, the mission made the following priority recommendations:

TABLE 1. Priority Recommendations

Target Date	Priority Recommendation	Responsible Institution
April 2025	Publish the RPPI based on the new index production system and update the technical documentation	Geostat
Early 2025	Negotiate an MoU with myhome.ge on the data delivery and data structure and enter into negotiations with ss.ge	Geostat
Mid-2025	Analyze the collected information on rental prices whether a rental price index can be established	Geostat
End-2025	Test the available information on old properties whether the data can be integrated in the RPPI	Geostat

9. Further details on the priority recommendations can be found under Detailed Technical Assessment and Recommendations.

Section I. Detailed Technical Assessment and Recommendations

A. INTRODUCTION

10. Web-scraping refers to the reverse engineering of the database hosted by the website owner by extracting the data from selected national web portals. Web-scraping has the potential to expand price collection and the outlet sample, with wide product coverage and detailed product descriptions. But at the same time, it has technical challenges and potential legal issues. Algorithms developed using web-scraping software, such as R, are data-agnostic and can be used for house prices but also for consumer prices too.

11. Previously, Geostat applied web-scraping (with the permission of the website owner) using a Google Chrome extension with a simple point-and-click interface. However, the dynamic website programming led to a large number of errors, making time-consuming manual adjustments necessary to clean the scraped results. The mission deployed an approach that is running fully automated without the need for user intervention or any manual cleaning or adjustment of the scraped results. The mission also implemented a new production system, revised modelling and compilation procedures, based on web-scraping the existing website (www.ss.ge). To simplify the data collection, Geostat should inquire with ss.ge whether their database can be shared.

12. After the end of the previous mission, an unforeseen issue developed with the web-scrapers, resulting in complications with the results from a second website (www.myhome.ge). Geostat inquired directly with the owners and their database can be shared instead. Geostat will need to establish a Memorandum of Understanding (MoU) with myhome.ge in order to firmly agree upon the data delivery and the data structure e.g. fields, names, temporal coverage, etc. A follow-up mission will further assist Geostat in extending the data sources.

B. TIDYING THE SCRAPED DATA

13. **The scraped data from the three months of the most recent quarter are loaded and joined, at the monthly level, with the exchange rate of the Georgian lari against the US dollar obtained from the National Bank of Georgia.** On this basis, the prices of all houses and flats are denominated in Georgian lari, either because the scraped data were already denominated in lari, or the US dollar amount was converted using the end-of-month exchange rate. Together with Geostat, the mission decided against the use of daily exchange rates since the advertised houses and flats are not sold immediately and this would induce artificial variability in the US dollar prices.

14. **Houses and flats are filtered where the condition indicates that they are old with repairs.** The completion status has already been filtered in the web-scraper to newly built or under construction. For flats, the project type is restricted to “non-standard”, which is a catch-all term for other than Soviet type (for houses this variable is mostly missing and thus not used for filtering in-scope observations).

15. **In the part that follows, duplicates are removed.** It would be incorrect for a house or flat with the same address, say, to appear on the dataset multiple times for the same period, e.g., data from multiple listings. However, it is often not straightforward to identify duplicates. Yet, identifying and managing duplicates is crucial for maintaining clean and accurate data, especially in large datasets. This is because, duplicates give implicitly higher weight to one house or flat than to all others.

16. **First, perfect duplicates, or records that are exactly the same in every field, are removed.** Here, always the newest record is kept – the same record might be scraped three times in any given quarter, i.e., every month. It should be noted that this case is not strictly considered a duplicate since time is a very significant variable in itself; henceforth, these will be referred to as “repeated” observations. The last advertised price might be considered close to the transaction price. Based on the first two months of the fourth quarter of 2024, this removes 1,722 of 7,417 houses (23 percent) and 8,246 of 82,439 flats (10 percent).

17. **Then, two sets of partial duplicates, which occur when some, but not all, of the data points match, are removed.** Even for actually identical houses or flats, it will be seldom the case that all characteristics matches. In this case, searching for partial duplicates is the way to go. However, it will be somewhat arbitrary which variables to choose here. The mission tested the results for different combinations of variables. One set focuses on the numerical variables such as area and the condition and completion status, the other set focuses on the price, numerical variables and categorial variables, e.g., whether the flat has central heating. Again, based on the October and November 2024 data, this removes further 1,529 of the remaining 5,695 houses (27 percent) and 37,866 of 74,193 flats (51 percent).

18. **The next part re-labels the condition, completion status, project type, and district from Georgian to English in order to simplify the further data processing.** This is followed by stratifying the data according to the status of the districts into prestigious and non-prestigious, separately for houses and flats – although with considerable overlap in which districts are considered (non-)prestigious. Also, a price per square meter variable is added. Finally, the tidied quarterly file is saved for use in the next step.

C. REMOVING OUTLIERS (AND ERRORS)

19. An error refers to a mistake or inaccuracy in data collection, entry, or processing. Errors can occur due to various reasons like faulty equipment, human mistakes, or data corruption. Errors need to be corrected or removed to ensure the accuracy of the data analysis. An outlier is a data point that significantly differs from other observations in the dataset. Outliers can occur naturally due to variability in the data or might indicate something unusual or interesting. Not all outliers are errors; some might represent valid but rare events.

20. However, outliers can sometimes be the result of errors, so it is important to investigate their cause. One way to identify outliers is to determine a plausible range of values and track all observations that fall outside of the range. This range can either be developed using a variety of statistical techniques or based on the compiler's judgement and knowledge. The limits should be defined according to the data characteristics of each stratum (and per period). The mission tested different limits for the interquartile range before setting the final values to be used in the outlier detection routine.

21. In order to gauge the historical variability of the data, several quarters of data are loaded. Previously, Geostat performed the outlier detection manually by cutting off extreme prices per square meter in the lower and upper tails of the distribution if these observations did not sit well with the rest of the observations. Essentially, the interquartile range follows the same strategy but allows to automatize this step of the process. This algorithm drops all observations in a given quarter and district status (the time and stratum dimensions, respectively) that lie outside the range of the lower quartile minus a multiple of the interquartile range and the upper quartile plus a multiple of the interquartile range. Since the lower and upper quartiles represent the 25^{th} and 75^{th} percentile of the empirical distribution, the interquartile range encompasses 50 percent of the observations. The multiple of the range allows for some additional variation. One of the major advantages of this approach is that is depends on the distribution of the actual data, i.e., if there is a widespread, observations further away from the "middle" (the median, to be precise) are not deleted automatically, while if the spread is more narrow, this approach will identify observations more easily as outliers.

22. The thresholds were established experimentally. For houses, the multiple of the interquartile range is 1.0 in prestigious districts and 2.0 in non-prestigious districts. For flats, the multiple is 3.0 in both strata. In all cases, the calculations were performed on the logarithmic scale to account for the typical positive skew of prices that may arise from values far above the mean, such as high-value houses and flats. The whiskers in the chart below reflect the range of permissible values in the upper and lower tail, respectively.

Distribution of prices per square meter by strata
(Georgian lari, log scale)

Sources: ss.ge, Geostat; and IMF staff calculations.

23. Putting this to work at the quarterly level, it is worthwhile to check how the allowed range of values as well as the proportion of outlying observations moves over time, noting that price levels not only differ across strata but also fluctuate over time. Since this approach is typically accompanied by Cook's distance after the hedonic regression step, only gross outliers should be removed at this stage. Out of 1,135 houses in prestigious districts, 60 (five percent) are removed, and out of 3,031 houses in non-prestigious districts, 63 (two percent) are removed. For flats, the numbers are in prestigious districts 28 out of 14,706 (0.2 percent), and in non-prestigious districts 137 out of 21,620 (zero-point six percent). The majority of outliers being lower outliers, i.e., where the price is less than the lower whisker, most likely due to data errors. The data set with outliers removed for the most recent quarter is saved for use in the next step.

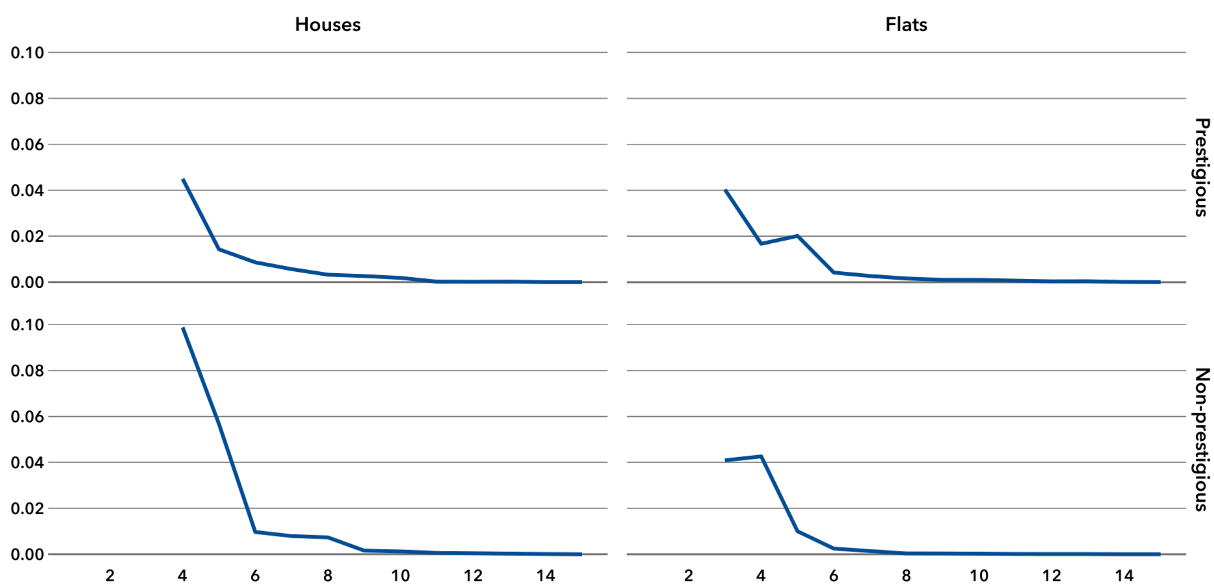
D. MODEL SELECTION, WEIGHTS, AND MEDIANS

24. To apply the hedonic approach, we need to select a regression model for each stratum The authorities chose the following strata: (i) houses in prestigious districts, (ii) houses in non-prestigious districts, (iii) flats in prestigious districts and (iv) flats in non-prestigious districts. Firstly, we should prepare the data for model selection. The condition of houses and flats are grouped into three frame types, particularly "black frame" for those which need redecoration, "white frame" for those currently being redecorated, and "green frame" for those that are (newly) redecorated. Also, numerical variables are filtered to allowable ranges using thresholds developed experimentally, e.g., the floor area, the number of rooms, or the area per room. A complication arose for houses: the plot area might refer to the garden only, excluding the footprint of the building. Further, e.g., the number of rooms was aggregated for a greater number of observations in each group, increasing the reliability of estimates. Last, a new variable with the administrative districts was created, aggregating the individual districts into groups. Since the R codes are separate for model selection, weights, and medians (and the hedonic imputation approach), these calculations are repeated in each individual code.

25. **There is no free lunch in model selection: no one model dominates all others over all possible datasets.** On a particular dataset, one specific model may work best, but some other model may work better on a similar but different dataset. What works well for houses does not necessarily work well for flats and vice versa, what works well for prestigious districts does not necessarily work well for non-prestigious districts and vice versa, and what works well for one dataset (website) does not necessarily work well for another dataset (website) and vice versa. Hence, it is an important task to decide for any given set of data which model produces the best results. Selecting the best approach can be one of the most challenging parts of compiling RPPIs in practice. For more technical details on model selection see Annex 1.

26. **As can be seen from the chart below, adding the fourth characteristic for both prestigious and non-prestigious houses have quite a positive impact on the fit of the models i.e. using the metric of adjusted R^2.** The fourth characteristic is the condition for houses in prestigious districts and the administrative district for those in non-prestigious districts. The fifth characteristic still contributes significantly, and it turns out to be administrative district for prestigious houses and condition for non-prestigious ones, making the two models identical (by chance). The contributions to adjusted R^2 of the subsequent variables are negligible so that the final model in both cases includes (i) floor area, (ii) plot area, (iii) the administrative district, and (iv) the condition (time dummies will not be included in the model used for producing the RPPI). For flats, the mission, together with Geostat, settled for adding three additional variables in both cases to floor area. In prestigious districts (in order of relevance), the administrative district, the number of floors, and the condition, and in non-prestigious districts, the condition, the administrative district, and whether the flat has central heating. Both for houses and flats, the adjusted R^2s of models using more characteristics were quite similar. The rationale here is that if a set of models appear to be more or less equally good, then we might as well choose the simplest model—that is, the model with the smallest number of predictors.

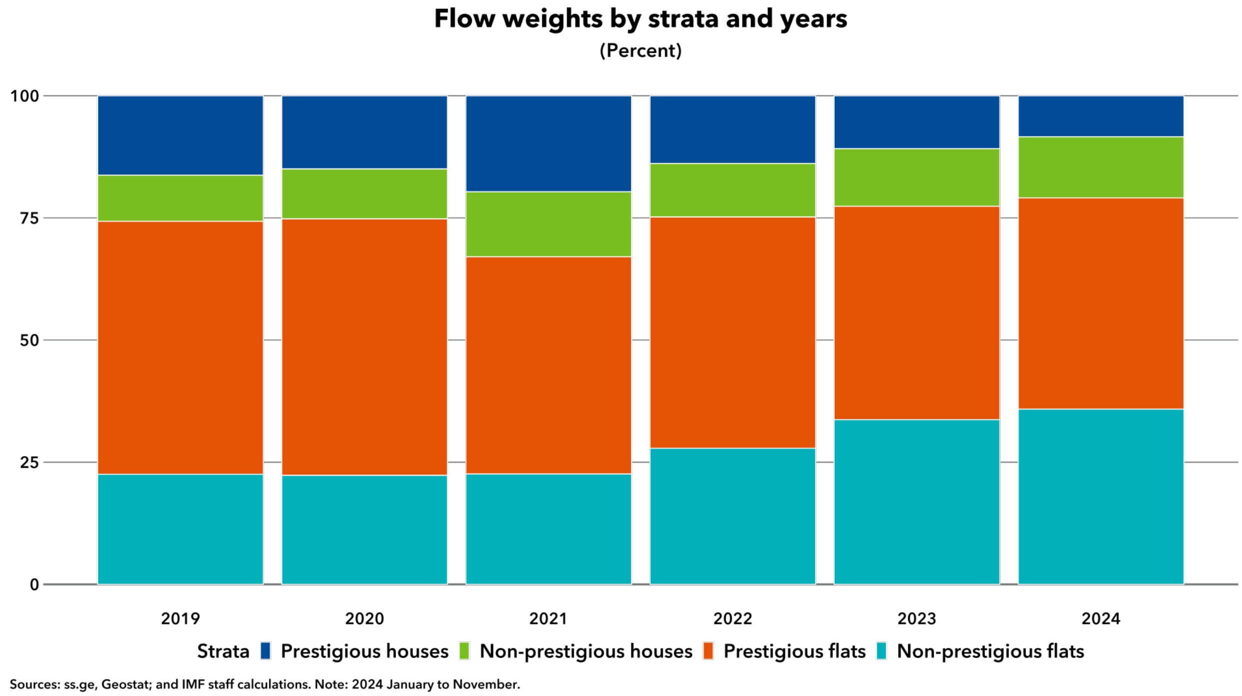
Change in adjusted R^2 statistic by strata
(As a function of the number of characteristics)



Sources: ss.ge, Geostat; and IMF staff calculations.

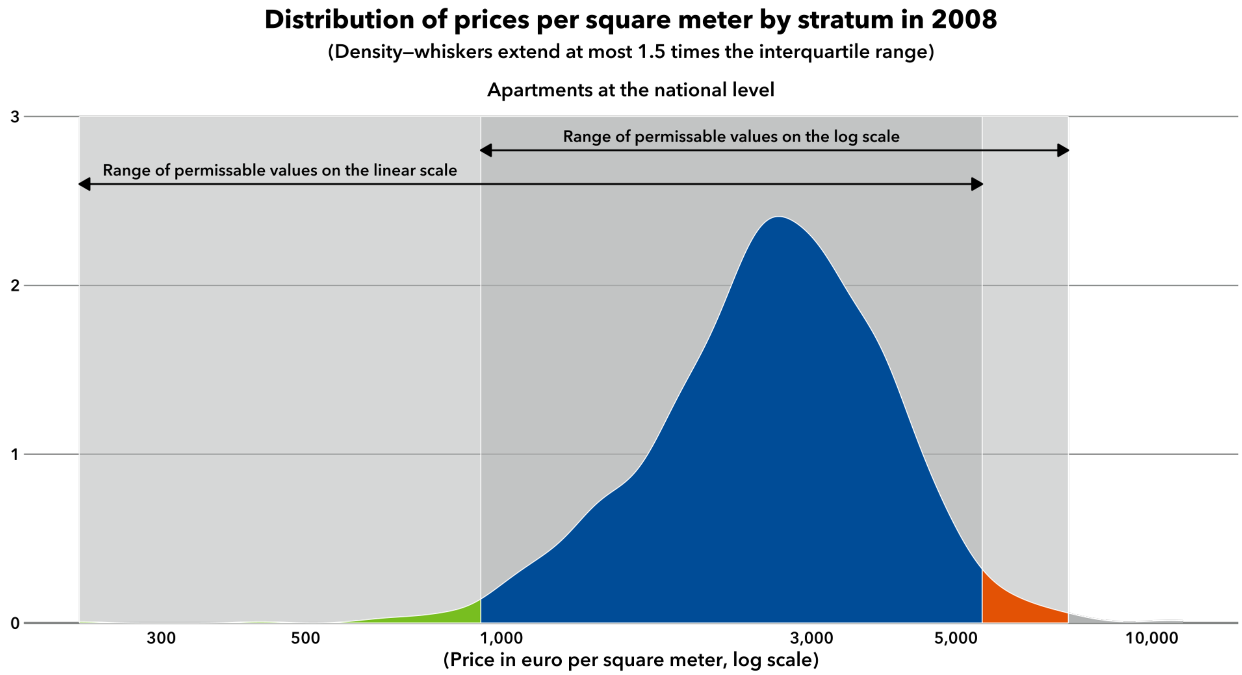
27. Before turning to the index calculation based on these model selection results, weights are used to aggregate indices from the stratum level to higher levels and the aggregate level. Explicit weights relate to weights calculated on the basis of the value of the houses and flats. These weights will reflect the relative importance of each stratum. Flow weights reflect changes from one period to the next, i.e., the value of all houses and flats purchased during a certain period. Flows can be derived from the same source used for prices if it is deemed reflective of overall market activity. It is best practice to update the weights annually and chain-link the indices.

28. The flow weights based on values can simply be derived by summing up the purchase prices, which would be equivalent to multiplying the number of purchases by the average price. Flow weights are typically not restricted to one year of data. In fact, it is recommended to update the weights annually. The chart below shows that, over the years, the relative importance of the strata varies. It is good practice to update the weights annually so that the weights continue to be representative of the residential property market.



29. In addition to the quarterly indices, Geostat also publishes median prices per square meter. For newly built houses, these are broken down by the ten administrative districts. The same is true for newly built flats, but in addition a second breakdown by the three frame types is provided.

30. As already explained, outliers were identified for houses and flats by prestigious and non-prestigious districts and on the logarithmic scale. Previously, the outlier identification was performed on the linear scale. Due to the positive skew of prices per square meter, this has potentially a sizeable effect on the medians. Unlike indices, which measure change over time, medians measure levels. On the linear scale, applying the interquartile range will filter out high-value houses and flats. Since the effect cannot easily be identified given the mismatch between outlier identification on (non-)prestigious districts and publication of administrative districts, this is exemplified in the chart below using a synthetic dataset that the IMF provides for training and educational purposes.



Sources: Synthetic House Price Dataset, CSO Ireland; and IMF staff calculations.

31. The improved methodology for calculating median prices will be explicitly documented. It should be noted, however, that the previous time series will remain unchanged and will not undergo any revisions. Additionally, the enhanced method for index calculation will be detailed in the updated methodological manual, which will be made available on the Geostat website.

E. INDEX COMPILATION–HEDONIC APPROACH

32. Price indices are generally measured on a matched-model basis. The goal is to compare prices over time, ensuring like-for-like comparisons. Detailed specifications are used to facilitate this process. When the characteristics of sampled varieties change, we quality-adjust so that our price comparison is “like for like”. We want a constant-quality property price index, but properties are heterogeneous, very different in terms of location and other price-determining characteristics, properties are sold infrequently and the quality-mix of types of properties transacted changes from period to period, and further, the data sources used are often secondary sources, and not designed for RPPI compilation.

33. Hedonic regressions recognize that a dwelling is a bundle of characteristics (structure attributes and location). These characteristics cannot be sold separately and so their prices cannot be independently observed. Hedonic regression techniques can estimate “shadow” or marginal valuations for these characteristics. We can use these shadow prices in a number of different ways to compile constant-quality price indexes. A set of K characteristics of the dwellings are identified and data over N dwellings are collected. A hedonic regression of the log of the price of dwelling i, ln p_i, on its set of quality characteristics $z_{k,i}$ is given by:

$$\ln p_i = \beta_0 + \sum_{k=1}^{K} z_{k,i} \cdot \beta_k + \varepsilon_i$$

where the β_ks are estimates of the shadow prices (or marginal valuations) the data ascribes to each characteristic.

34. The authorities should use the Hedonic Imputation Approach for index compilation. This method is outlined in more detail in Annex 2. A second layer for detection of unusual observations after fitting a hedonic regression is Cook's distance. This will ensure the parameters are not unduly influenced by these unusual observations and thus the index is not affected. "Unusual" can have two meanings in this context. Outliers are observations for which the response is unusual given the predictors. In contrast, observations with high leverage have unusual values for the predictors.

35. The new index production system will be used for the first time with the publication of the RPPI for the first quarter of 2025. In particular, the mission helped to modernize the index calculation process and update the methods for index calculation. This included tidying the scraped data, removing outliers (and errors), model selection and calculating weights and medians, and implementing the hedonic imputation approach. The previous approach was largely outdated and would have needed a review even without a major upgrade of the data sources.

F. OUTLOOK AND NEXT STEPS

36. Three areas of work were identified for further research. First, once the MoU is in place with myhome.ge, their data should be analyzed with the aim to integrate this second source into the RPPI. Second, information on rental prices for flats are currently already scraped, this information should be used in a follow-up mission to establish both a rental price index as well as to derive the valuation indicator price-to-rent ratio, a very useful tool in detecting housing bubbles. Third, old properties are not included so far in the RPPI, further scrutiny is needed as regards the non-representative number of observations, missing variables such as age, and unreliable results due to overall higher variability.

G. OFFICIALS MET DURING THE MISSION

Name	Institution	Position
Gogita Todradze	Geostat	Executive Director
Giorgi Tetrauli	Geostat	Head of Price Statistics Department
Ketevan Bedinashvili	Geostat	Senior Specialist
Khatuna Aptsiauri	Geostat	Head of Consumer Price Statistics Division

H. ANNEX 1: MODEL SELECTION

37. It is natural to want to quantify the extent to which the model fits the data. The quality of a linear regression fit is typically assessed using the R^2 statistic. It takes the form of a proportion—the proportion of variance explained—and so it always takes on a value between zero and one and is independent of the scale of the response. It turns out that R^2 will always increase when more variables are added to the model, even if those variables are only weakly associated with the response. When adding a variable leads to just a tiny increase in R^2, this is evidence that the variable can be dropped from the model. Essentially, it provides no real improvement in the model fit, and its inclusion will likely lead to poor results due to overfitting.

38. It is natural to wonder which predictors are related to the response. It is possible that all of the predictors are associated with the response, but it is more often the case that the response is only associated with a subset of the predictors. The task of determining which predictors are associated with the response, in order to fit a single model involving only those predictors, is referred to as variable selection. Various statistics can be used to judge the quality of a model, to determine which model is best. The adjusted R^2 statistic is a popular approach for selecting among a set of models that contain different numbers of variables. For a least squares model with K variables, the adjusted R^2 statistic is calculated as:

$$\text{Adjusted } R^2 = 1 - \frac{\text{RSS}/(N-K-1)}{\text{TSS}/(N-1)}$$

where RSS is the sum of squared residuals, the residual is the difference between the observed response value and the response value that is predicted by our linear model, and TSS is the total sum of squares, measuring the total variance in the response—the natural logarithm of prices.

39. While RSS always decreases as the number of variables in the model increases (and consequently R^2 always increases), RSS / (N – K – 1) may increase or decrease, due to the presence of K in the denominator. The intuition behind the adjusted R^2 is that once all of the correct variables have been included in the model, adding additional noise variables will lead to only a very small decrease in R^2. Since adding noise variables leads to an increase in K, such variables will lead to an increase in RSS / (N – K – 1) and consequently a decrease in the adjusted R^2. Therefore, in theory, the model with the largest adjusted R^2 will have only correct variables and no noise variables. Unlike the R^2 statistic, the adjusted R^2 statistic pays a price for the inclusion of unnecessary variables in the model.

40. One classical automated and efficient approach to choose a smaller set of models is forward selection. We begin with the null model—a model that contains an intercept but no predictors. We then fit K simple linear regressions and add to the null model the variable that results in the highest adjusted R^2. We then add to that model the variable that results in the highest adjusted R^2 for the new two-variable model. This approach is continued until all of the predictors are in the model. Forward stepwise selection involves fitting a total of 1 + K × (K + 1) / 2 models.

41. With this dataset, we cannot add predictors to the model one-at-a-time but need a variant of the textbook approach instead. This is because, for example, the characteristic condition is actually depicted by two dummy variables (representing the three frame types). We either add the two dummy variables or leave out the characteristic condition altogether. While we technically could have added just the dummy variable for one condition, say, this would be counterintuitive. In order to account for the differences in number of variables added by adding one characteristic to the regression model, the mission used adjusted R^2 in order to select the best model with one additional characteristic. It is key to understand the distinction between number of characteristics and number of variables for what follows.

42. **Still, we must identify the best model among a set of models with different numbers of characteristics or variables.** The adjusted R^2 curves may be quite flat. In this setting, we can select a model based on more subjective grounds, like parsimony. The rationale here is that if a set of models appear to be more or less equally good, then we might as well choose the simplest model—that is, the model with the smallest number of predictors. This is done by eyeballing the change in the adjusted R^2 curve and looking for a point at which the change in adjusted R^2 drops to zero. This inflection point is often referred to as an elbow. However, this type of visual analysis is inherently ad hoc.

43. **Models using price and price per square meter as responses are observationally equivalent (if the model contains floor area).** The only difference is the coefficient on the log of floor area. As a consequence, we will always include log of floor area as the first characteristic in all subsequent models. For houses, log of plot area is also always included. Further, time dummy parameters shift the hedonic surface upwards or downwards and measures the effect of "time" on the log of price. This means that the baseline model for houses already includes three characteristics, while that for flats two.

I. ANNEX 2: HEDONIC IMPUTATION APPROACH

44. **The hedonic imputation approach that works at the level of individual dwellings – its rationale lies in the matched-model method, comparing like for like.** In this approach, the log of prices is regressed on the price-determining characteristics separately for each period:

$$\ln p_i^t = \beta_0^t + \sum_{k=1}^{K} z_{k,i}^t \cdot \beta_k^t + \varepsilon_i^t$$

45. **The hedonic imputation (HI) index going from period 0 to period *t* is given as the ratio of average predicted prices in periods *t* and 0, using period 0 characteristics:**

$$I_{HI}^{0:t} = \frac{\exp\left[\frac{1}{N^0}\sum_{i=1}^{N^0}\ln \hat{p}_{i|z^0}^t\right]}{\exp\left[\frac{1}{N^0}\sum_{i=1}^{N^0}\ln \hat{p}_{i|z^0}^0\right]} = \frac{\exp\left[\frac{1}{N^0}\sum_{i=1}^{N^0}\left(\hat{\beta}_0^t + \sum_{k=1}^{K} z_{k,i}^0 \cdot \hat{\beta}_k^t\right)\right]}{\exp\left[\frac{1}{N^0}\sum_{i=1}^{N^0}\left(\hat{\beta}_0^0 + \sum_{k=1}^{K} z_{k,i}^0 \cdot \hat{\beta}_k^0\right)\right]} = \frac{\exp\left[\frac{1}{N^0}\sum_{i=1}^{N^0}\left(\hat{\beta}_0^t + \sum_{k=1}^{K} z_{k,i}^0 \cdot \hat{\beta}_k^t\right)\right]}{\exp\left[\frac{1}{N^0}\sum_{i=1}^{N^0}\ln p_i^0\right]}$$

where ln $p^t_{i|z0}$ and ln $p^0_{i|z0}$ are the predicted logs of prices imputed using period zero characteristics and period *t* and period 0 regression coefficients, respectively. The last equality follows from a feature of linear regression that the average of predicted prices (using characteristics from the same period as the imputation is for) is equal to the average of actual prices. Since the mission used Cook's distance from a preliminary hedonic regression as a measure to filter out influential observations, however, the sample for the imputation step is adjusted, i.e., N^0 here refers to those observations kept in the final hedonic regression (and likewise for the thus estimated coefficients β^0 and β^t).

46. **Before moving to the application of the hedonic imputation method, the mission first reviewed the stability of the coefficient estimates.** But before that, the averages of dwelling characteristics over time were inspected. Developments in the underlying data will be crucial in understanding trends and abrupt changes in the coefficient estimates as well as serve as a data quality

check. Charts visualizing the results can be found in the Appendix. The interpretation of the regression parameters is as follows.

Log of floor / plot area: Elasticity of price. If the coefficient is zero point seven, for example, prices increase by zero-point seven percent if floor/plot area increases by 1 percent (and vice versa). As such, the price per square foot *decreases* by zero-point three percent in this scenario.

Dummies on administrative districts / condition / number of floors / central heating: Average price level difference between each administrative district, condition, etc., and their reference group (for which the coefficient is set to zero, see the results in the Appendix). If the coefficient is zero point two, for example, prices are, on average, 100 × (exp (0.2) – 1) = 22 percent higher than in the reference group.

47. Updating weights each year requires linking together indices spanning shorter sequences of periods that overlap in one (or more) periods to form an index for a long sequence of periods; this resulting index is referred to as a chain index. When using the one-quarter overlap technique, there is for each year a series of five index numbers running from the fourth quarter of the previous year (its index number being equal to 100) to the fourth quarter of the current year. Now these separate five-quarter series can be linked together into a single long-term series.

48. The mission continued with calculating for each year a series of five index numbers using the fourth quarter of the previous year as the reference period, the index links, which do not form a single long-term series:

$$I_{HI}^{4,y-1\,:\,q,y} = \frac{\exp\left[\frac{1}{N^{4,y-1}}\sum_{i=1}^{N^{4,y-1}} \ln \hat{p}_{i|z^{4,y-1}}^{q,y}\right]}{\exp\left[\frac{1}{N^{4,y-1}}\sum_{i=1}^{N^{4,y-1}} \ln \hat{p}_{i|z^{4,y-1}}^{4,y-1}\right]}$$

where 4, y – 1 refers to the fourth quarter of the previous year and q, y to the current quarter. The imputation sample $z^{4,\,y-1}$ is adjusted for influential observations using Cook's distance.

49. Using the flow weights from the previous year, the four strata (houses and flats in prestigious and non-prestigious districts) are aggregated to the all-dwellings level using the Laspeyres formula; these index links still do not yet form a single long-term series.

$$I_{L}^{4,y-1\,:\,q,y} = \frac{\sum_{j=1}^{4} w_j^{y-1} I_{HI}^{4,y-1\,:\,q,y}}{\sum_{j=1}^{4} w_j^{y-1}}$$

50. Finally, the Laspeyres index links in 2025 will be chained on the index last compiled using the old methodology in the fourth quarter of 2024. The index keeps the reference period 2020 = 100.

$$I_{Lc}^{2020\,:\,q,2025} = I_{Lc}^{2020\,:\,4,2024} I_{L}^{4,2024\,:\,q,2025}$$

51. Both the aggregation and chaining steps were previously done using Microsoft Excel, this practice will be continued. The R code produces the hedonic imputation indices for the four strata along with the weights (as well as the medians per administrative district) and this is inputted into Microsoft Excel. To this end, the aggregation and chaining procedure has been updated in the spreadsheet, too.