Compilation of Armenia’s Residential Property Price Index (RPPI)

Tigran Baghdasaryan (Central Bank of Armenia), Kristine Karapetyan (Central Bank of Armenia), Gor Lazyan (Central Bank of Armenia), Norbert Pfeifer (University of Graz, Austria), Barra Casey (IMF)

International Conference on Real Estate Statistics - Tokyo - February, 2026

ABSTRACT

Understanding house price dynamics is essential for decision-making in financial markets and serves important information for macroeconomic policy interventions. This paper presents the methodology for Armenia's Residential Property Price Index (RPPI), developed by the Central Bank of Armenia (CBA) in collaboration with experts from the International Monetary Fund (IMF). Administrative data from the Cadastre Committee of Armenia (CCA), which covers all residential transactions since 2018, are used to estimate a hedonic, time-dummy regression model to construct quality-adjusted house price indices. The paper addresses practical challenges in the process of developing a first-time house price index, including data acquisition, hedonic specification, and index aggregation. We demonstrate the compilation of separate indices for apartments and individual houses, with further disaggregation into regional sub-indices for apartments. Our analysis show that the hedonic methodology provides superior quality adjustment and produces more stable indices as compared to the stratified median approach. Finally, remaining data challenges including the limited property characteristics for individual house, and the absence of information that allows to distinguish between new and existing dwellings are presented. The CBA intends to publish the proposed indices on a quarterly frequency.

Keywords: RPPI, Real Estate Market, Time-dummy Hedonic Method, Index Aggregation, House Price Index, Armenia

Contents

Abbreviations

RPPI – Residential Property Price Index

CBA – Central Bank of Armenia

CCA - Cadastre Committee of Armenia

IMF – International Monetary Fund

TDH – Time dummy hedonic

SPAR – Sales price appraisal ratio

Introduction

The Residential Property Price Index (RPPI) has important applications across three main areas. First, the RPPI serves policymakers as a critical indicator for financial stability, it supports monetary policy decisions, and provides information on inflation. Second, statistical offices use the RPPI as a deflator in national accounts and to integrate owner-occupied housing into the consumer price index. In addition, the RPPI helps measure the housing component of household wealth, as housing is usually the largest component of household wealth. Lastly, the RPPI provides valuable information to the public, including both private home buyers and institutional investors, to support decision-making in residential property markets.

Hence, understanding housing dynamics is essential for both market participants and macroeconomic policymakers. Given the Central Bank of Armenia (CBA) role in ensuring financial stability, the CBA reflects these considerations, providing reliable market information while meeting analytical requirements for monetary policy and financial stability assessment.

Prior to this work, Armenia lacked an internationally comparable RPPI. To address this gap, the CBA, in collaboration with IMF experts, initiated the compilation of a nationwide index using administrative data from the Cadastre Committee of Armenia (CCA). The CCA database provides comprehensive coverage of all registered residential transactions and includes detailed property characteristics, which enable the application of hedonic methods that are consistent with international standards (IMF, 2020; OECD, 2013).

Unfortunately, data availability did not permit the creation of strata based on a distinction between new and existing properties.

The main objective of this paper is to present an internationally comparable RPPI for Armenia that can support policymakers and investors in assessing inflation-related risks and identifying vulnerabilities in the banking system. The paper also details the specific challenges of constructing a comparable property price index for Armenia. Section two reviews different RPPI construction approaches that are commonly used in the RPPI production. The third section describes the CCA database in detail and describes the available characteristics. Section four outlines the RPPI construction methodology chosen by the CBA, and the fifth section presents the results, at both the aggregate and disaggregate levels. Finally, the paper concludes with insights gained and with recommendations for future work to steadily improve the quality of the index.

Methodology

Since a property price index needs to meet different requirements depending on its intended use, different methodologies have been developed to address this demand. Moreover, not only the audience that uses the index matters, but also the availability and quality of data vary considerably across countries, leading to diverse methodological approaches to construct robust and reliable indices. However, to support this process of developing an appropriate index, international guidelines have been developed that provide well-documented methodologies for compiling Residential Property Price Indices (RPPIs). Namely, the Handbook on Residential Property Price Indices (OECD, 2013) and the Practical Guide on the Compilation of a Residential Property Price Index (IMF, 2020).

The Handbook (OECD, 2013) outlines four main approaches to index compilation:

1. Stratification or Mix Adjustment Methods
2. Hedonic Regression Methods
3. Repeat Sales Methods
4. Sales Price Appraisal Ratio (SPAR)

Although these approaches are discussed in detail in the Handbook, alongside practical applications, we provide a brief overview of each method in the following section and highlight its advantages and disadvantages.

Stratification or Mixed-Adjustment

The most straightforward approach to temporarily measuring changes in house prices is to use a measure of central tendency from the distribution of transaction prices—typically the mean or median —and track this value over time. Given the positive skewness of house price distributions and the heterogeneous nature of property transactions, the median is often preferred over the mean. Nevertheless, both measures share a key limitation: they can produce noisy and potentially biased estimates of price developments because they do not adequately account for changes in the composition or quality of properties sold over time.

In the context of residential property price indices, the stratification approach can be viewed as an extension of the simple mean or median method. It involves dividing the total sample of properties into several sub-samples (strata) that share common characteristics, so that the average value within each stratum reflects properties of similar quality. This extension significantly improves the quality of the index by comparing like with like over time. By tracking average prices for each stratum and aggregating them with weights that reflect their relative importance in the market, compilers dramatically improve quality compared to simply using all the data at once. Therefore, by pooling comparable properties into strata and aggregating the weighted averages of these strata, a quality-adjusted price index can be obtained.

The effectiveness of the stratification approach, however, depends primarily on the choice of the stratification scheme. A very detailed stratification scheme based on many property characteristics can increase homogeneity within strata and, at the same time, reduces the compositional bias from changes in quality mix. Yet, there is a trade-off: increasing the number of strata reduces the number of observations per stratum, potentially increasing the overall RPPI's standard error. Therefore, detailed stratification requires sufficiently rich data and well-defined classification criteria for each stratum. Below, we show how the stratification index can be compiled.

With M different strata, the stratification index can be represented in mathematical form[1] as follows:

$$P^{ot} = \sum_{m=1}^{M} w_m^0 P_m^{ot} \qquad \textit{(1)}, \text{ where}$$

P^{ot} denotes the index for stratum m, which compares the mean (median) price in the current period, or more generally a comparison period t, with the mean (median) price in an earlier or base period 0.

w_m^0 indicates the corresponding weight of stratum m based on total expenditures.

advantages

- ✓depending on the choice of stratification variables, the method adjusts for compositional change of the dwellings,
- ✓is reproducible, conditional on an agreed list of stratification variables, and relatively easy to explain to users.

disadvantages

- ✓cannot deal adequately with depreciation of the dwelling units unless age of the structure is a stratification variable,
- ✓cannot deal adequately with units that have undergone major repairs or renovations,
- ✓requires information on housing characteristics so that sales transactions can be allocated to the correct strata.
- ✓The quality control is limited as for each stratum, sufficient observations in each quarter are needed.

Hedonic Regression Methods

The hedonic regression method is not a single index compilation approach; moreover, it provides different applications depending on data availability. In its core, it addresses a key limitation of the stratification method - the manual selection of appropriate levels for disaggregation. Choosing too many strata leads to better quality control, as homogeneity within each stratum increases, but at the cost of fewer observations per stratum, thereby weakening statistical precision. The hedonic approach addresses this trade-off by incorporating multiple

[1] Handbook title Eurostat, 2013.

variables simultaneously, whereas its most straightforward approach retains the full sample of transactions in a single estimation.

House prices are influenced by a range of attributes, such as location, size, and building type, among others, making it difficult to extract a single measure of their development. This is particularly true as no two properties are exactly alike. In addition, only a small share of the housing stock is transacted within a given period, making it impossible to match identical units across time. Therefore, the common index used in consumer or producer price indices is not applicable to the real estate market. Hence, to obtain a measure that reflects a constant-quality price change, a more sophisticated adjustment technique is required.

The hedonic regression methodology provides such a framework. It treats each dwelling as a bundle of observable characteristics, with the transaction price representing the combined implicit value of these attributes. Although the individual attribute prices cannot be directly observed, they can be inferred through econometric estimation—so-called shadow prices. By controlling for variations in these attributes across periods, the hedonic approach isolates pure price movements from changes in quality, allowing compilers to derive quality-adjusted property price indices.

advantages	disadvantages
✓can effectively incorporate the main characteristics of properties if sufficient information on these attributes is available, ✓by appropriately segmenting the sample, it becomes feasible to create price indices for various types of dwellings and location ✓it generelly requires fewer observations per (adjusted) quality characteristic than the stratification approach	✓adequately controlling for location can be challenging, especially when property prices and price trends vary significantly across specific geographic areas, ✓demands a substantial amount of data, as it necessitates information on all relevant property attributes, which can make it relatively costly to implement, ✓It may require a significant amount of metadata

In practice, three approaches are commonly used for applying hedonic regression: the time dummy method (TD), the (average) characteristic method, and the imputation method. Since these approaches account for the majority of applications used by statistical offices and central banks, we briefly describe them below.

Time-dummy hedonic method

The time-dummy hedonic method is particularly suitable when compilers have access to detailed microdata that include numerous explanatory characteristics but lack a large number of transactions per period. By pooling transactions across multiple periods, this approach efficiently uses all available information while controlling for quality differences. The most common functional form is the log-linear specification, in which the logarithm of the transaction price is the dependent variable. The logarithmic transformation helps to approximate a more normally distributed dependent variable and reduces heteroskedasticity. The regression can be written as:

$$ln\, p_n^t = \beta_0 + \sum_{t=1}^{T} \delta^t D_n^t + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t \qquad (2),$$

where $ln\, p_n^t$ denotes the log transaction price of property n in period t, $\mathrm{Z}_{\mathrm{n,k}}^{\mathrm{t}}$ captures the K explanatory characteristics (such as location, size, number of bathrooms, etc.), for property n, and β_k are the corresponding characteristics shadow prices. β_0 is the intercept, and $\mathrm{D}_{\mathrm{n}}^{\mathrm{t}}$ is a matrix with only zeros, except for each row, it has one entry indicating the transaction period with a one. ε_n^t is an observation specific error term with mean zero.

The coefficients on the time dummies ($\hat{\delta}_t$) measure the pure price changes between the base period and period t, holding property characteristics constant. Exponentiating these coefficients yields the quality-adjusted price index, from which the final index values I_t are derived as follows:

$$I_t = exp(\hat{\delta}_t) * 100 \qquad (3)$$

Average Characteristics

The characteristics (hedonic) method measures price changes by tracking the evolution of a "typical" property whose attributes reflect the average characteristics of the properties sold in each period. This approach is generally preferred over the time-dummy method when a sufficiently large number of transactions are available, allowing separate hedonic regressions to be estimated for each time point.[2]

The concept of a "typical" property is defined by the mean or representative values of its characteristics within a given period or stratum. For example, if 100 dwellings are sold in a given quarter and their average floor area is 92 square metres, the representative or "typical" dwelling for that period is assumed to have a size of 92 m^2. For categorical variables, such as the number of bedrooms or districts, the typical property reflects the relative outcome of each attribute. For example, if one-third of the dwellings have one bedroom and two-thirds have two bedrooms, the typical property is characterised by 33 % one bedroom and 67 % two

[2] Guide title IMF, 2020.

bedrooms. Hence, the price in this period for the typical property is estimated with this bundle of characteristics.

Therefore, the regression in $lnp_n^t = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t$ (4) is run separately for each period t, from 1 to T as follows:

$$lnp_n^t = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t \qquad (4)$$

The estimated coefficients from $lnp_n^t = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t$ (4) are then used to determine the price evolution of the representative (average) property over time.

The imputation method

The imputation hedonic method is closely related to the characteristics (average property) method, as it also estimates separate hedonic regressions for each period using the same model specification as in $lnp_n^t = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t$ (4). However, rather than predicting the price for a single representative property, this method imputes a price for every transaction. Specifically, for each property, a price is predicted twice: once using the coefficients estimated for period t, and once using the coefficients estimated for period t+1.

This procedure effectively yields two predicted prices for each transaction—one reflecting period t conditions and another reflecting period t+1 conditions—allowing the calculation of individual price changes between the two periods. These predicted price changes are then aggregated, typically using the arithmetic mean, geometric mean, or median, to derive the overall quality-adjusted price index.

The imputation method thus provides a flexible framework that fully exploits available microdata while ensuring that all observations contribute to the estimation of period-to-period price changes. Compared to the time-dummy approach, it avoids imposing a constant relationship between characteristics and prices over the entire sample period, while still maintaining the advantage of explicit quality adjustment.

The repeat sales method

The repeat sales method was first proposed by Bailey, Muth, and Nourse (1963). It relies on information from properties that have been sold more than once and is therefore classified as a matched-pairs approach. Because it measures price changes for the same property between two sales, the method implicitly controls for all time-invariant quality characteristics, eliminating the need for explicit quality adjustment.

However, this approach typically uses only a subset of all transactions, since most properties are sold only once during the observation period. Consequently, the sample size is often substantially lower, limiting the representativeness of the resulting index measure. Furthermore, the subset of repeatedly sold properties may differ from the full housing stock, potentially introducing sample-selection bias. For instance, properties sold more frequently may differ in quality or location, and may include a disproportionate share of lower-quality dwellings.

Another challenge arises from potential quality changes between sales. Renovations or improvements can lead to an overestimation of price growth, while depreciation or physical deterioration between sales can bias results downward. The simple repeat sales model assumes constant quality across transactions, which may not hold in practice.

To mitigate these issues, several refinements to the repeat sales approach have been proposed, including hybrid hedonic-repeat sales models that explicitly account for quality differences over time. Despite its limitations, the repeat sales method remains a valuable tool for housing market analysis. A well-known application is the S&P CoreLogic Case–Shiller U.S. National Home Price Index, which is based on a refined implementation of the repeat sales framework.

advantages

- ✓does not require extensive information about property characteristics, unlike the hedonic method,
- ✓are relatively straightforward to apply compared to other methods

disadvantages

- ✓relies exclusively on data related to properties that have been sold multiple times within the specified sample period,
- ✓there is a high probability of encountering sample selection bias issues in repeat sales data,
- ✓it may be not usable due to insufficient market sales
- ✓can be unrepresentitative for the whole market

Sales Price Appraisal Ratio (SPAR)

The Sales Price Appraisal Ratio (SPAR) method compares the actual transaction price to the appraisal value to determine price changes. The index is constructed by tracking the evolution of the average ratio of sale price to appraisal value over time. The fundamental assumption underlying this approach is that appraisal values tend to adjust more gradually than market prices, allowing movements in the ratio to reflect market-driven price dynamics.

A major benefit of the SPAR method is that it does not require repeat sales, nor does it require detailed information on individual property characteristics. It can therefore be implemented in markets where mass appraisal systems or administrative valuation databases exist and are regularly updated. This makes the SPAR method particularly practical for countries where official property valuation systems cover a large share of the housing stock, such as in several

European countries, including Denmark, the Netherlands, and Sweden, as well as outside Europe (e.g. New Zealand).

The accuracy and reliability of a SPAR-based index depend critically on the consistency and quality of the appraisal data. Changes in valuation methods or update frequencies can distort the index by capturing methodological revisions rather than genuine market price changes. Furthermore, systematic biases in appraisal values—such as those varying by region, property type, or value range—can affect the representativeness of the results.

Despite these limitations, the SPAR method provides a cost-effective and transparent approach to measuring residential property price developments, especially in settings where transaction data are limited but administrative appraisal records are comprehensive and of consistent quality.

advantages

- ✓it is essentially rooted in the standard matched model methodology and aligns with traditional index number theory
- ✓the computation procedure is straightforward
- ✓detailed data about property characteristics is not required

disadvantages

- ✓It does not account for changes in the quality of properties.
- ✓The results heavily rely on information from previous period assessments.

Published in 2020, the Practical Compilation Guide for Residential Property Price Indices (IMF, 2020) focuses primarily on stratification and hedonic regression approaches, while also addressing the repeat sales and SPAR methods more briefly.

When discussing the main methodological approaches, it is essential to consider international experience. The absence of full harmonisation across countries in the compilation of RPPIs possess challenges for users and analysts, as methodological differences hinder direct international comparisons and cross-country studies.

Institutional arrangements for the compilation and dissemination of RPPIs also vary considerably. In some countries, the National Statistical Institute (NSI) is responsible for producing and publishing the index, whereas in others, the National Central Bank (NCB) or a government department carries this mandate.

The Bank for International Settlements (BIS) provides a valuable service by collecting and disseminating national RPPIs in a harmonised global database. The BIS currently publishes more than 300 series from 61 countries, representing the indicators with the broadest national coverage available for each jurisdiction. In selecting the series, the BIS collaborates closely with national central banks and considers the methodological guidance provided in the OECD Handbook on Residential Property Price Indices (2013). As a result, the BIS residential property price dataset offers a valuable source of internationally comparable information, even though underlying data sources and compilation methods differ across countries.

Using the methodological documentation provided by the BIS for selected residential property price series, Table 1 briefly presents the methodology of countries where central banks are responsible for compiling housing price indices. [3]

Table 1. International experience in Housing Price Indices

Countries	*Methods*
Austria	Hedonic Regression
Chile	Mixed stratification
Columbia	Hedonic regression
Germany	Hedonic regression
Greece	Mixed stratification
India	Mixed stratification
Moldova	Characteristics hedonic method
Morocco	Repeat-sales
N. Macedonia	Hedonic regression
Thailand	Hedonic regression (time dummy)
Turkey	Hedonic regression

As Table 1 illustrates, the hedonic regression approach is the dominant method among the listed countries, consistent with the recommendations of international handbooks and guidelines. However, the specific type of hedonic approach applied depends largely on the availability of detailed data and the technical and institutional resources of the compilers.

[3] The information of Moldova was taken from the website of National Bank of Moldova.

Data Sources

Given the specifications for tracking property prices and the intended use of RPPIs, it is crucial to select an appropriate data source for constructing these indices.

In the literature, various sources are used to compile an RPPI. The most widely used data sources include land registry data, data from real estate platforms, appraisal data, real estate agencies data, and other administrative databases. Statistical agencies and various entities typically rely on land registry data to create RPPIs, which are derived from house price data generated through a country's legal and administrative procedures related to property transactions. These datasets often encompass all transactions nationwide and are cost-effective, thus, placing minimal burden on respondents.

However, one limitation of administrative data sources is their potential for time delays. In some countries, there may be a lag of three months or more between the actual transaction and its recording. This process can lead to a substantial lag in the index compilation. Fortunately, in Armenia, property sales are promptly recorded by the Cadastre Committee of Armenia (CCA), minimizing time-related distortions and providing an ideal data source for RPPI compilation.

The CCA database contains all registered residential property transactions in Armenia, along with detailed residential property characteristics. This dataset is considered high-quality, with the declared price serving as a reliable indicator of the price negotiation between the seller and the buyer. While transactions in Armenia are often conducted in US dollars, recorded prices in Armenia are denominated in Armenian Dram.

Furthermore, the CCA dataset is very timely, with transactions for the reference quarter typically transmitted to CBA within 28 days. Although the dataset is provided quarterly, it includes the registration date, enabling monthly index calculations based on transaction volume and user needs. Additionally, the dataset contains essential property characteristics, including floor area, region/district, property zone, construction material, floor level for apartments, and year of construction. Because the characteristics of residential houses differ from those of apartments, separate processing is necessary for each dwelling type. Tables below depict the available property characteristics for apartments (Table 2) and for individual houses (Table 3).

Table 2. Main characteristics of CCA database for apartments

Region	Zone
Street	Construction material
Address	Maximum stage of building
Floor area	Floor level
Price	Year of construction
Date of transaction	

Table 3. Main characteristics of CCA database for individual houses

Region	Land area
Street	Price
Floor area	Date of transaction

The dataset provided by CCA goes back to the first quarter of 2018 and are frequently transmitted on a quarterly basis. *Figure 1*, Figure 2 and Figure 3 depict the number of transactions for apartments, individual houses, and all types of residential properties accordingly.

Figure 1. Total number of transactions for apartments in each period

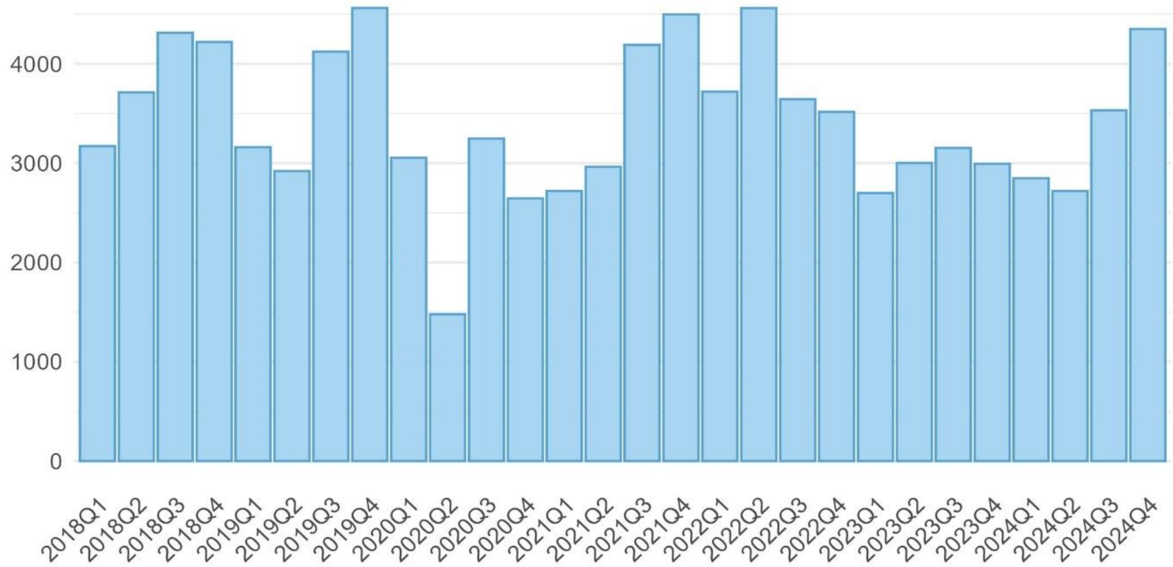


Figure 2. Total number of transactions for individual houses in each period

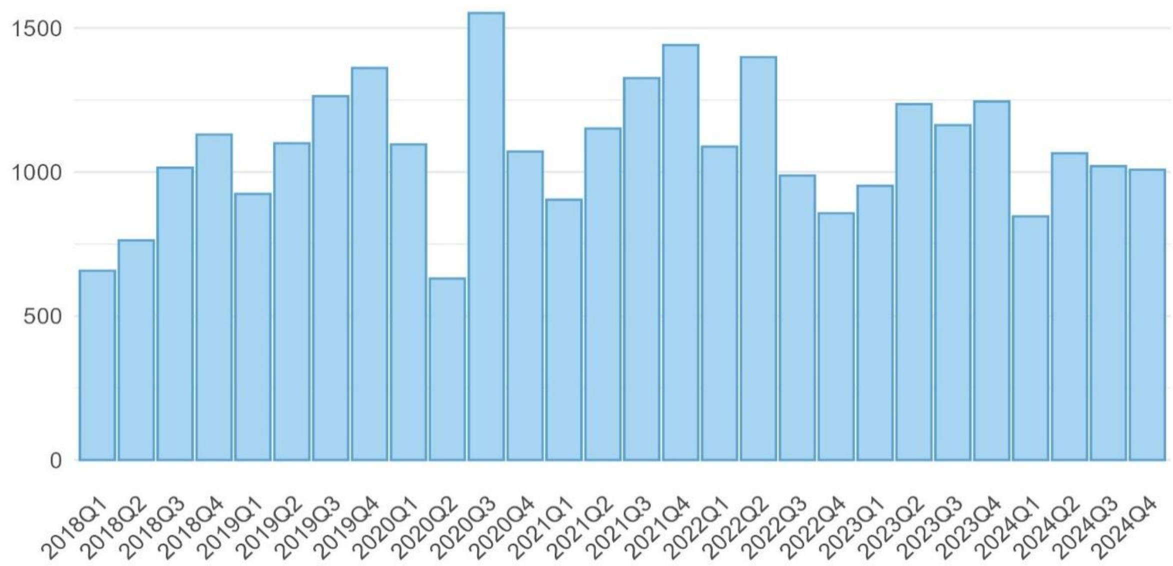
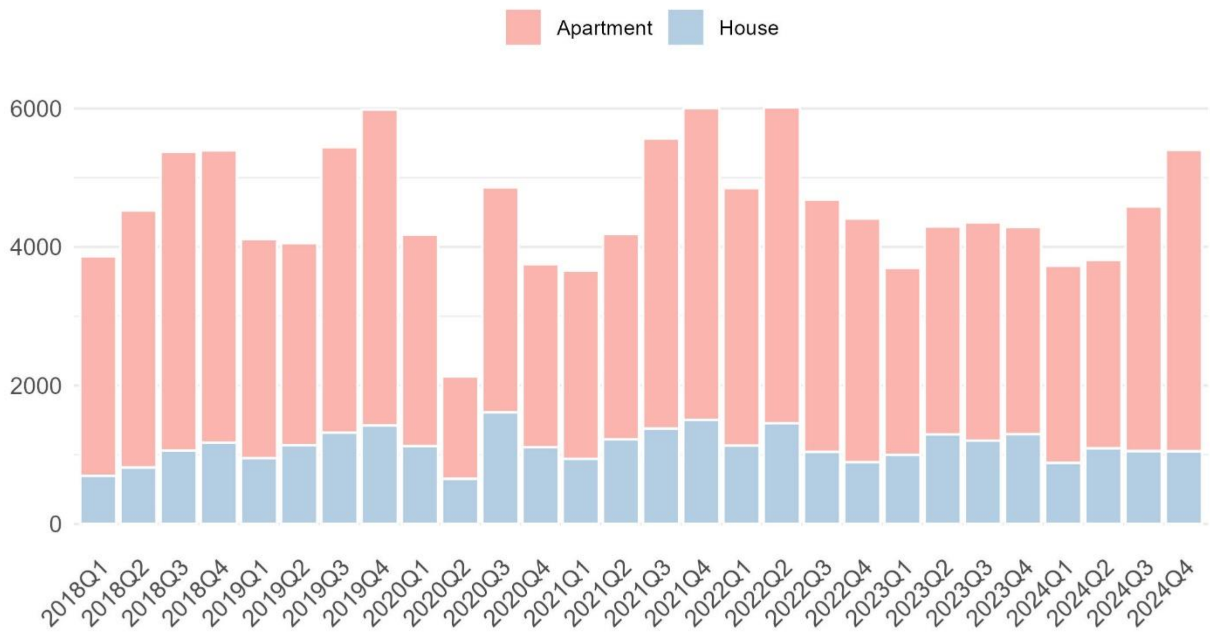


Figure 3. Total number of transactions for all type of residential properties in each period



As can be seen, the number of transactions exhibit some fluctuations between 2018Q1 and 2024Q4, with overall a higher number of apartment sales than for individual houses. Given the economic disruptions caused by the COVID-19 pandemic particularly in the second quarter of 2020, we see for both property types a sharp decline in transaction volume.

The Compilation of the RPPI in Armenia and Results

In line with the current literature, two primary methods are recommended for compiling residential price indices: the stratified median (or mix adjustment stratification) and the several hedonic regression methods.

The first step in constructing an RPPI is to prepare the raw transaction data to ensure the reliability and consistency of the source data. In this process, the data undergoes cleaning, filtering, and validation to remove errors and non-market transactions from the raw data. Also, outliers (e.g., very unusual transactions) require different treatment to avoid the strong impact of individual transactions on the index. In addition to identifying and treating outliers, it is necessary to address entries with missing information. Given that apartments and houses differ in their characteristics and market behavior, processing is carried out separately for each property type.

The hedonic regression methods discussed in section 2 encompass three approaches: the time dummy hedonic, the average characteristics, and the imputation method. Due to data availability and specificity in Armenia, the median with stratification and the time dummy hedonic (TDH) methods are used in this study.

The next section provides a brief overview of our methodological approach for compiling the indices.

Data processing

After selecting the data sources and before constructing an appropriate index, it is crucial to assess the data quality and undertake pre-processing steps to prepare the data for index calculation. As mentioned earlier, the two types of residential properties have different types of characteristics and specifications, so the data processing is done separately and separate models are run for each type. This section outlines the process of transforming raw data and presents the steps involved in its processing. The index is processed and compiled using the R statistical program.

Apartments

The initial step involves **visually inspecting the data to identify potential issues** and comparing the current dataset with the historical data to ensure consistency over time. To assess the quality of the CCA dataset and explore relationships among characteristics, we use various visualization techniques, including histograms, bar charts, and scatter plots. Descriptive statistics, including mean and median prices per region, are also employed to provide further information on the dataset's quality and characteristics.

In the second step, **non-market transactions and entry errors are filtered out from the dataset**. To achieve this, price per square meter thresholds were established for each region separately during the assessment process. These thresholds exclude observations that would otherwise distort the index. Specifically, for market transactions within Yerevan, the price per square meter was set to fall within the range of 50,000 to 5,000,000 AMD. For the Outside of Yerevan, the range was set from 15,000 to 1,000,000 AMD per square meter. Additionally, thresholds were set for the total transaction price, with a lower limit of 1,500,000 AMD and an upper limit of 600,000,000 AMD. Floor area boundaries were defined as follows: apartments with less than 20 square meters and above 400 square meters (above 500 square meters in Yerevan) were excluded from the dataset. These criteria help ensure the dataset contains only relevant and reliable market transactions for further analysis and index calculation.

In the third step, **transactions associated with the social housing program are to be excluded from the RPPI.** Social housing programs, which offer affordable housing to specific groups, are not directly related to house price development in Armenia and should therefore be omitted from RPPI compilation. Transactions linked to social housing programs are typically characterized by a unified price per square meter. To identify such transactions, observations involving multiple transactions within the same building with identical square meter prices were removed from the dataset. During the 2018–2024 period, approximately 5,000 of the approximately 100,000 (5.1%) observations were excluded using this approach. While this method may risk misidentifying transactions as social housing, because developers set similar prices for multiple apartments, the number of observations classified as social housing remains relatively small.

The fourth step involves analysing the **frequency distribution of each variable** to assess its potential influence on price. Variables with insufficient variation (e.g., if only one dwelling possesses a cellar) are excluded, as they lack sufficient variability for reliable estimation of the coefficients.

After selecting appropriate variables for the index construction, **duplicated entries are identified and removed** from the dataset. It is crucial to eliminate cases with duplicated addresses and dates during this stage to ensure data accuracy.

Subsequently, missing characteristics are filled up by using information from other transactions within the same building. Some building specific details such as the year of construction, building material, and building level may be absent from the dataset, in which cases other transactions within the same building are utilized to fill in these missing values. Through this approach, approximately 8,000 out of 9,000 observations with incomplete information were imputed. Only a small number of non-building specific information were missing in the dataset. Table 4 presents the number of observations with missing information for apartments that were filled using the approach. This imputation approach is only applicable to apartments, as multiple units within the same building share common characteristics, unlike neighbouring houses.

Table 4. Number of observations with missing values by variables for apartments

Variables	*Before*	*After*
Construction material	8948	1233
Building level	8911	1216
Construction year	9116	1233

Certain property characteristics are merged into homogeneous groups to address the low number of transactions per quarter for some categories. This grouping is particularly applied to (i) year of construction, (ii) building material, (iii) building level, and (iv) zones.

These characteristics are grouped into meaningful categories, particularly based on historical significance and common price segments. For property zones with only a few observations, they are grouped with the next-closest property zone based on median price per square meter.

After grouping the property characteristics, descriptive statistical analysis is conducted. Table 5 presents the summary statistics. A key aspect of this step is to identify significant differences between the maximum and minimum selling prices, which may indicate the presence of outliers in the dataset.

Table 5. Descriptive statistics of apartments (2018Q1-2024Q4)

Variable	*Min.*	*1st Q*	*Median*	*Mean*	*3rd Q*	*Max.*
Price, AMD	1,500,000	10,000,000	19,500,000	24,038,708	30,000,000	571,356,000
Log of price	14.22	16.12	16.79	16.65	17.22	20.16
Area, sqm	20.0	50.30	66.00	69.38	80.70	398.80
Price per sqm, AMD	15,168	166,362	316,432	339,128	446,512	3,415,444

Note: Table 5 provides the descriptive statistics of apartment transactions before outlier cleaning.

Figure 4. Distribution of price per sqm, AMD (Apartments, 2018Q1-2024Q4)

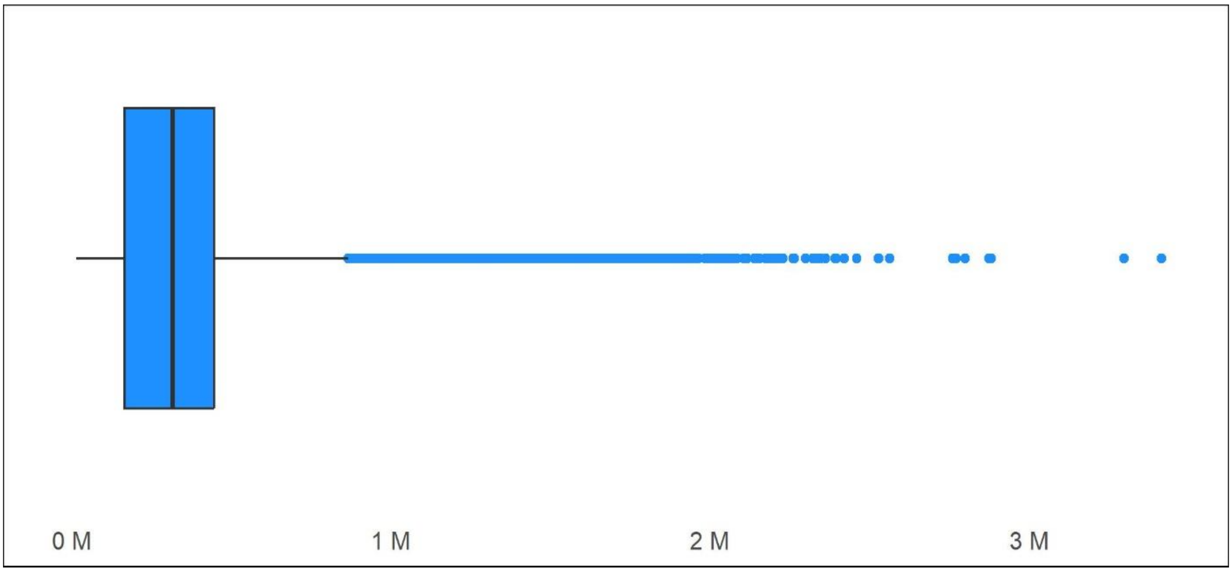
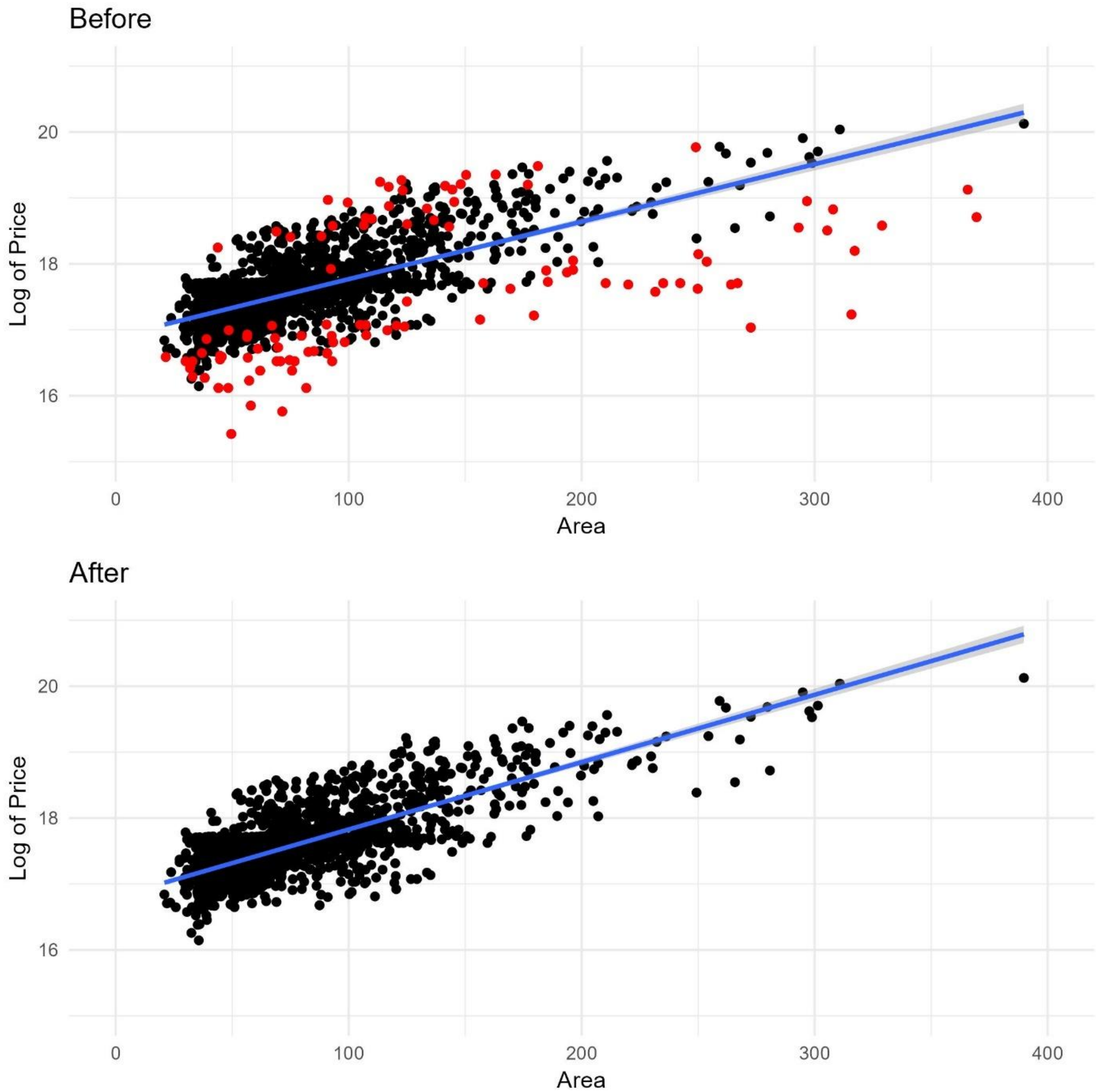


Table 5 presents the distribution of values for the observed variables. Notably, there is significant difference between minimum and maximum values, which suggests that outliers may indeed be present in the dataset. Outliers can have a substantial impact on statistical analysis and may skew results, thus warranting further investigation and potential treatment.

For outlier detection, we use the **Cook's Distance method**, which involves running a regression model as presented in $lnp_n^t = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t$ (4) that incorporates all relevant characteristics for the price determination. This regression model is rerun with all observations and with all except one observation (for each observation in the dataset) to calculate the impact of each transaction on the regression outcome. The threshold for determining outliers is typically set to 4 divided by the number of observations. This outlier detection procedure is applied separately for each stratum and rolling window, comprising 4 quarters of data. Based on a subset of our data, the percentage of observations that are identified as outliers falls between four and seven percent.

Figure 5 presents a scatter plot of the relation of area and log-price for a sample data before and after applying Cook's Distance for outlier detection. As can been seen from Figure 5, most outliers (marked in red) cluster around the distribution boundaries, exhibiting either low or high price per square meter relative to property size. This visualization enables a clear comparison of the data distribution and the impact of outlier removal through Cook's Distance method.

Figure 5. Applying Cook's Distance for outlier detection using sample data for apartments



Individual houses

The same steps of data processing were done for individual houses. Correspondingly the same indicators are calculated in this section which are presented in the following table and figures. The main differences that need to be highlighted is that in case of individual houses the transactions are generally concentrated in marzes (outside Yerevan). Additionally, it is important to note that, due to the limited availability of characteristics for houses compared to apartments, a national index was compiled without further regional breakdown.

Table 6. Descriptive statistics of individual houses (2018Q1- 2024Q4)

Variable	Min.	1st Q	Median	Mean	3rd Q	Max.
Price, AMD	1,500,000	5,860,418	11,400,000	19,127,974	21,979,500	481,140,100
Log of price	14.22	15.58	16.25	16.29	16.91	19.99
Floor area, sqm	40.0	101.6	159.6	177.3	234.9	499.8
Land area, sqm	40.0	307.0	600.0	827.4	1,087.0	12,618.0
Price per sqm, AMD	15,001	36,451	71,248	127,638	155,103	4,788,119

Figure 6. Distribution of price per sqm, AMD (Individual houses, 2018Q1-2024Q4)

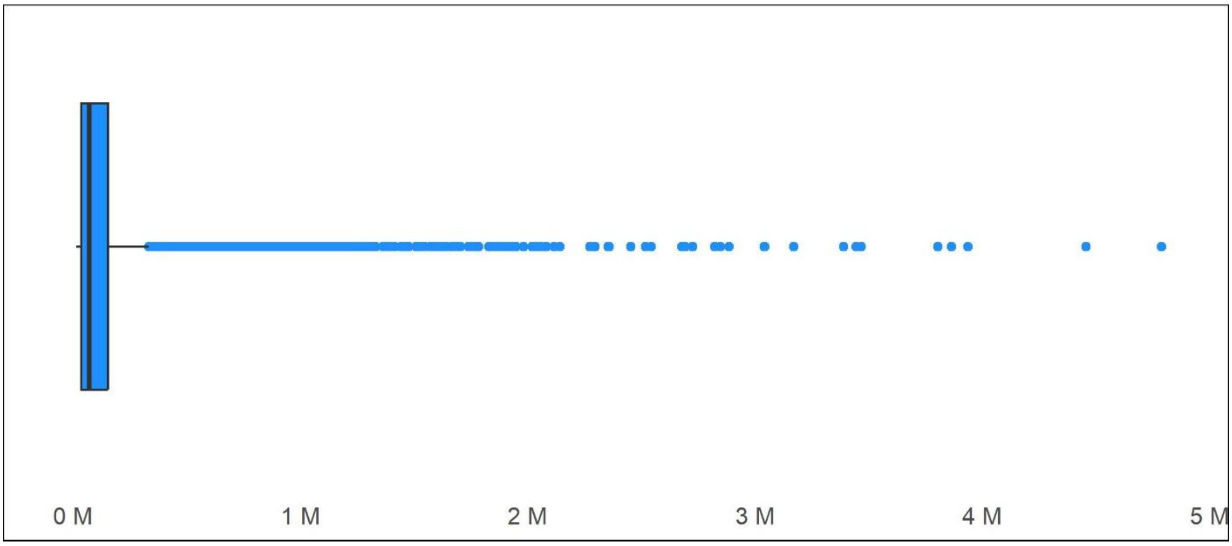
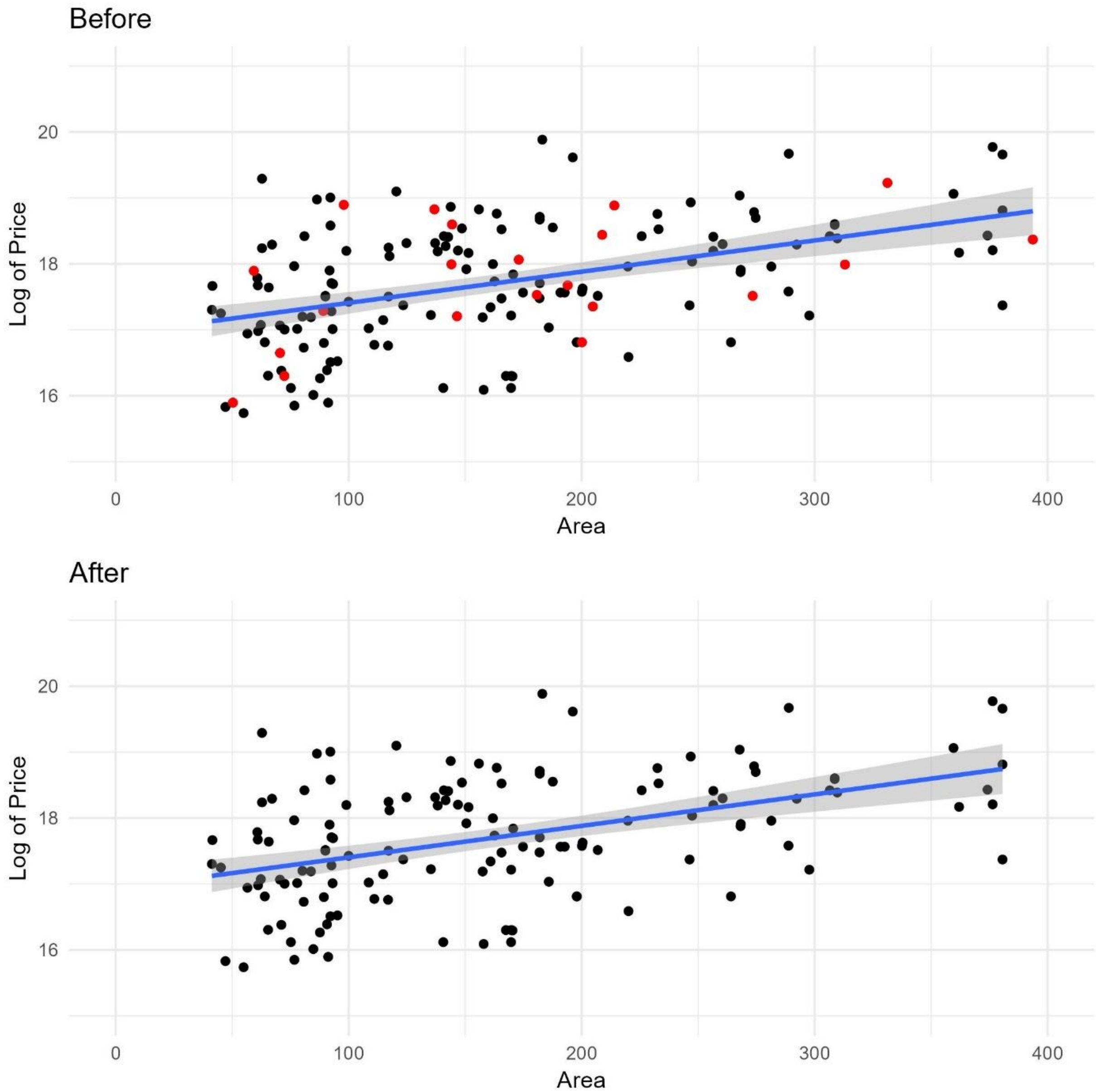


Figure 7. Applying Cook's Distance for outlier detection using sample data for houses

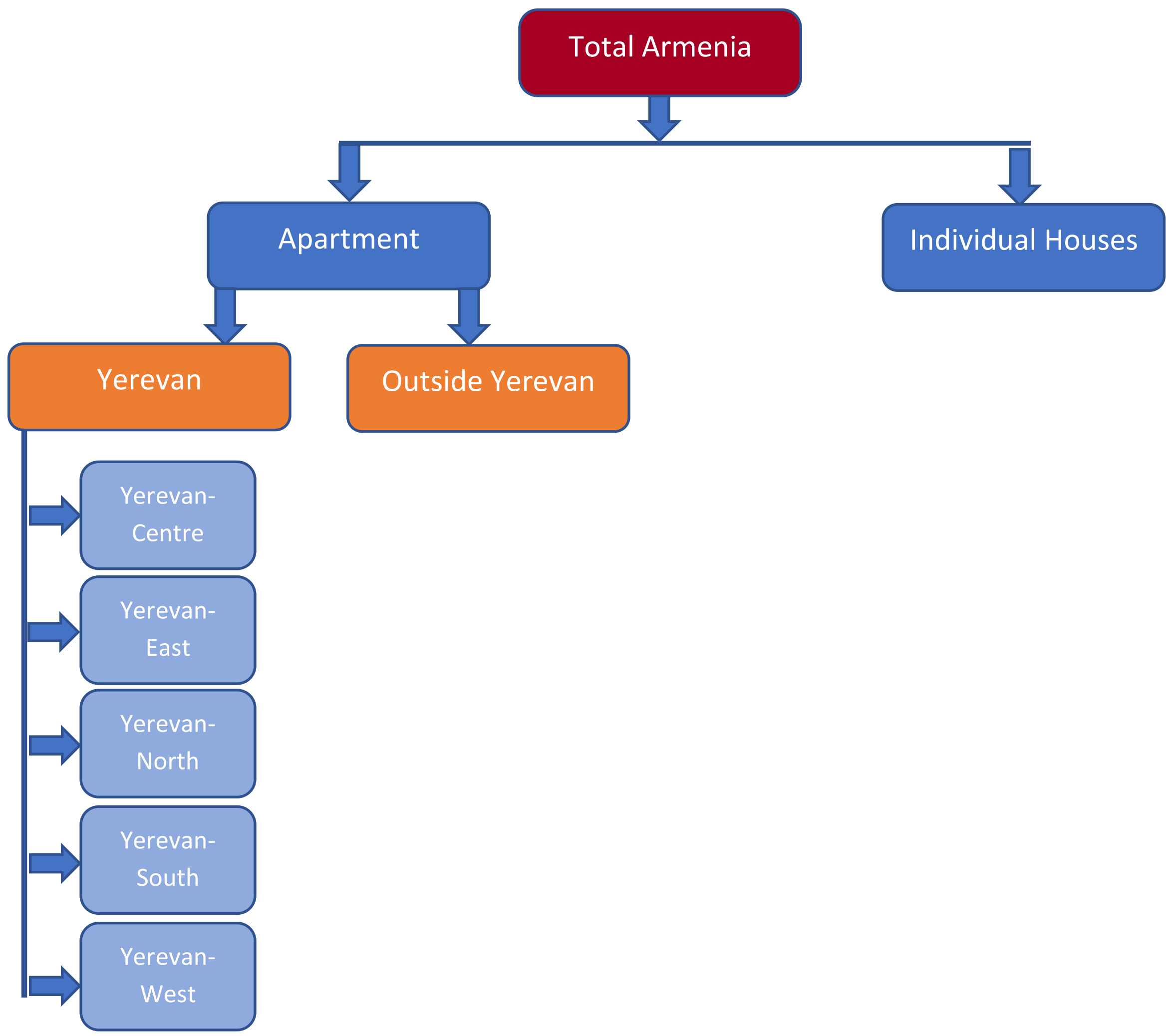


Stratification and Weighting

Stratification is fundamental to RPPI compilation, as it enables the construction of sub-indices that capture regional price dynamics while providing the granularity needed for targeted policy analysis. The stratification scheme for Armenia balances three considerations: data availability, regional market heterogeneity, and sample size requirements for reliable estimation.

As the data for apartments allow a more granular stratification scheme, the regions were split into **six strata, that were developed for the compilation of regional sub-indices in Armenia**. Within Yerevan, five regional strata were established for apartments. Due to the absence of enough information is the reason that the individual houses are not stratified and presented as one stratum (Individual Houses).

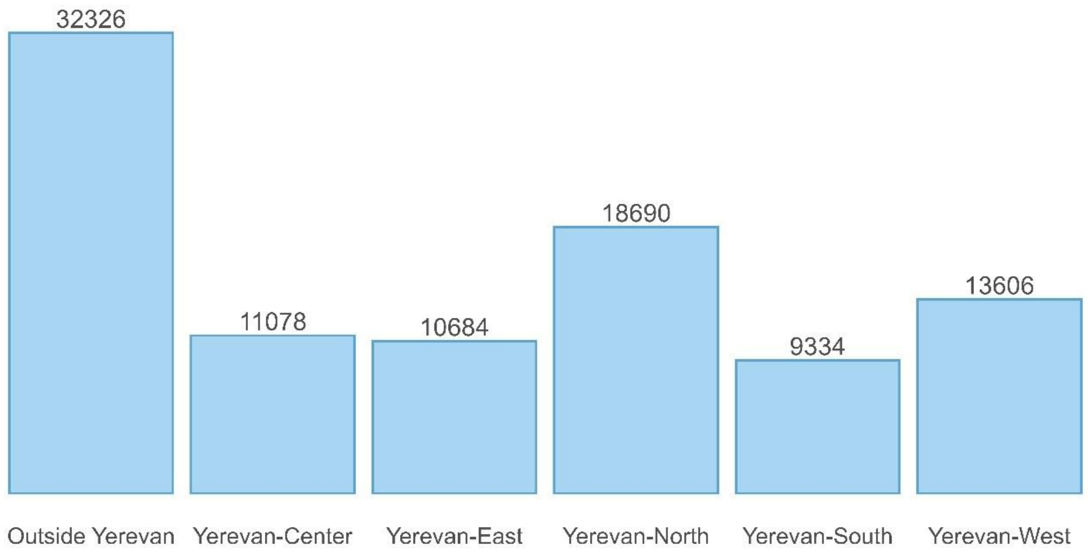
Figure 8. Individual and aggragated indices of RPPI



For constructing strata within Yerevan, neighbouring districts were aggregated into four directional zones (east, north, south, and west), as well as a separate stratum for the inner city. The resulting stratum-wise indices are subsequently utilized to aggregate data for the overall index for apartments both for total Yerevan and for the country as a whole.

Figure 10 provide a visual representation of the number of observations in each stratum. This information shows the distribution of data across the various strata over time.

Figure 9. Number of observations in each stratum (Apartments, 2018Q1-2024Q4)



As we can see from the figure above, the majority of apartment transactions in Armenia occur in the stratum outside of Yerevan. Within Yerevan, the most active stratum appears to be Yerevan-North. This information emphasises the importance of regional variations in transaction activity and market dynamics when analysing real estate trends and compiling property price indices. Understanding the distribution of transactions across different regions helps providing valuable insights into the dynamics of the real estate market in Armenia.

Figure 10. Number of observations in each stratum by period (Apartments)

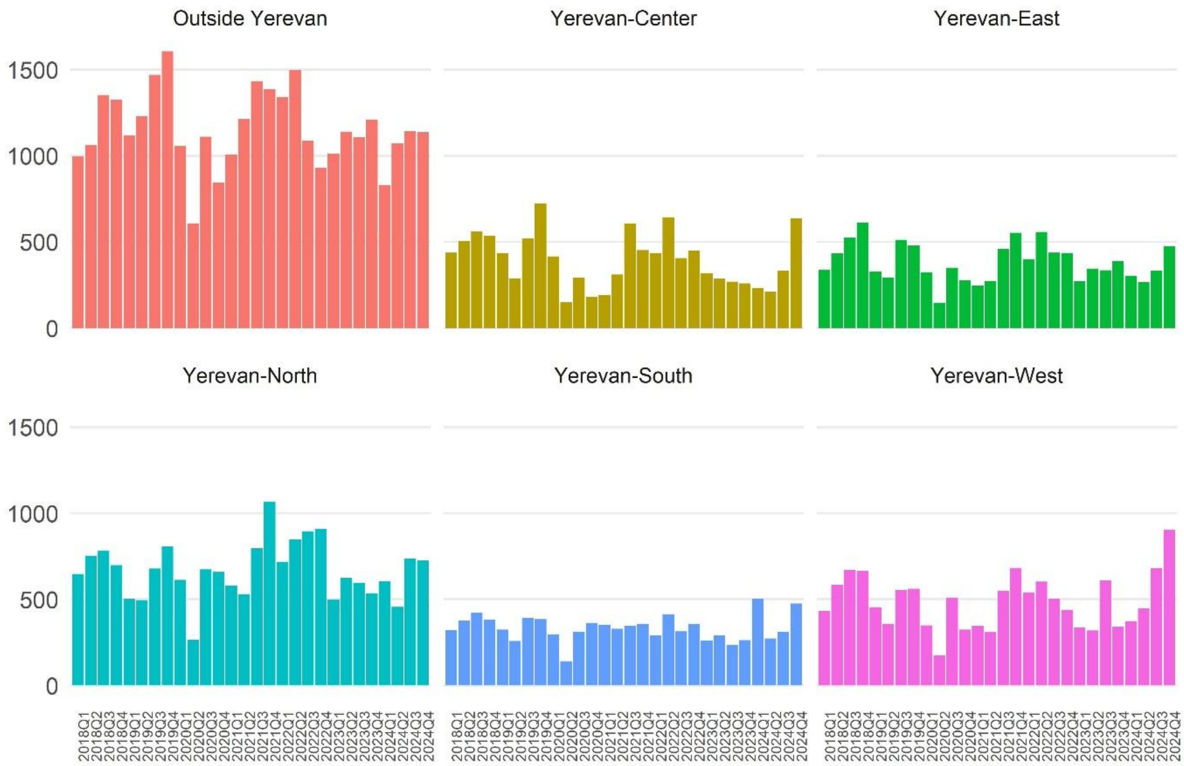
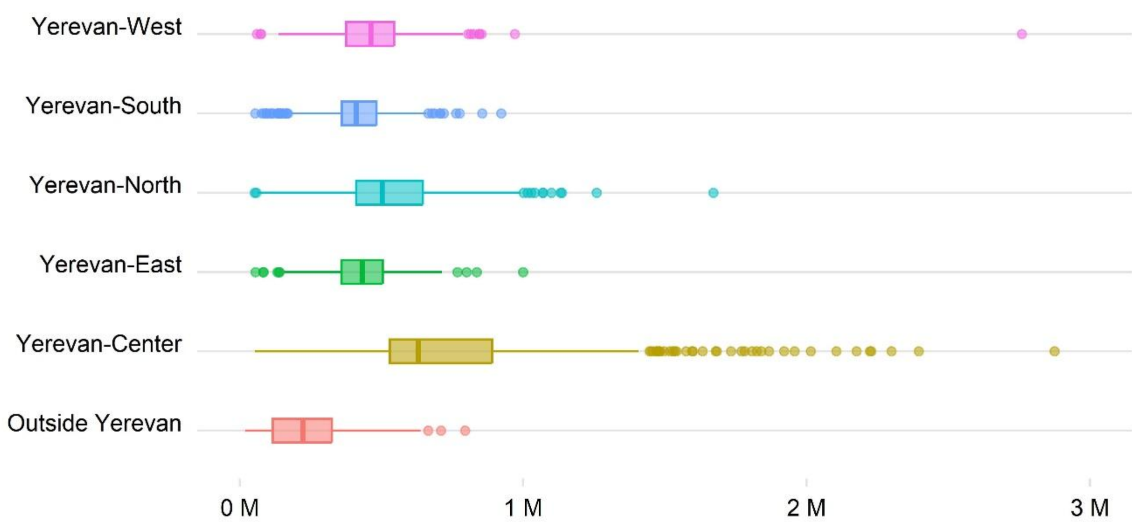


Figure 11 presents the distribution of prices per square meter within each stratum. This visualization provides valuable insights into the variability and spread of prices across different property strata. Analysing the distribution allows for a better understanding of the range and dispersion of prices within each stratum, highlighting any potential patterns or outliers that may exist.

Figure 11. The distribution of price per sqm for each stratum in 2024Q4

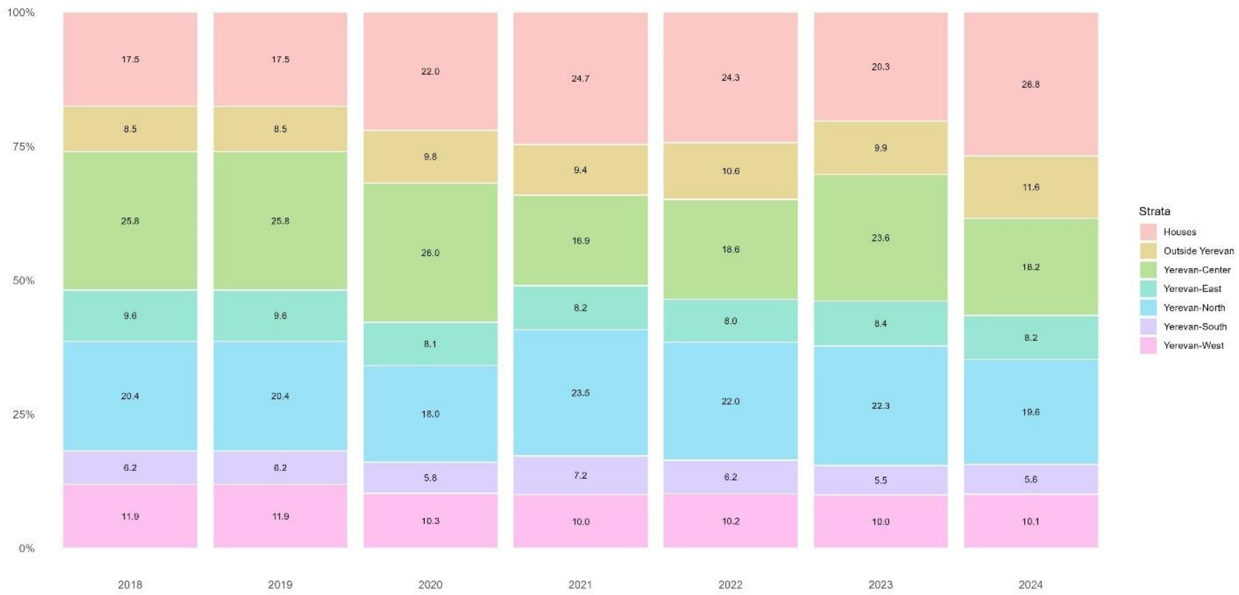


The data displayed in this graph indicate the presence of numerous outliers, highlighting the necessity to detect and address them before utilizing a cleaned dataset for index calculation. The scatter of the distribution observed in particularly in the city centre of Yerevan, suggests a higher degree of variability compared to other stratums. This variability can be attributed to the diverse types of buildings present in the city centre, contributing to a wider range of prices per square meter. Detecting and managing outliers is crucial for ensuring the accuracy and reliability of the index calculation process, especially in regions with greater diversity in property types and prices.

In order to aggregate the stratum-level price indices to an overall RPPI, weights are applied to retrieve the national index. There are two types of weights commonly used in RPPI compilation: stock weights and flow weights. Stock weights reflect the total value of the existing housing stock per stratum, while flow weights reflect actual transaction volumes. For the aggregation of sub-indices in the Armenian RPPI, flow weights were selected for three reasons. First, they better reflect current market dynamics, making them more suitable for financial stability monitoring—the CBA's primary use case. Second, reliable housing stock data by stratum are unavailable in Armenia, making stock weight construction infeasible without additional surveys. Third, flow weights align with international practice for central bank-compiled indices (BIS, 2023). For the representativeness of the weights, data from the previous year are used to derive weights for each year. Figure 12 depicts the weights of apartments correspondingly for the period from 2018 to 2024.

It's important to note that when receiving data from the CCA, stratification for houses must be considered separately from apartments. Apartments are typically concentrated in Yerevan, while houses are predominantly located in rural areas. This distinction ensures that appropriate stratification methods are applied for different types of residential properties.

Figure 12. Weights used for the aggregation of the apartment price index



Most observations are in the Outside Yerevan stratum, but this area has relatively low prices, according to the weights. From 2018 to 2024, Yerevan-Center and Yerevan-North share similar results, both high values and weights, mainly caused due to higher prices compared to other regions. Although Yerevan-Center has fewer observations, its weight remains significantly high, reflecting the overall higher price level in the city center. Another key finding is the growing contribution of single-family houses over time, increasing from about 17% in 2018 and 2019 to 27% in 2024.

The median index with stratification

Stratification[4] involves dividing the entire sample of houses into several sub-samples or strata. Once a measure of the change in the central tendency for each stratum, such as a mean or median price index, is constructed, the aggregate mix-adjusted RPPI is usually computed as a weighted average of indices for each stratum.

Step 1. Following data processing and stratification definition, we compute the median price within each stratum for every period.

Step 2. Individual indices are constructed for each stratum by dividing the median price in the current period (t) by the median price in the reference period (0) for that stratum.

Step 3. The individual indices are then aggregated using a Laspeyres method, employing a weighted average with base year weights. Here the base year weights are derived by dividing

[4] More detailed see in Guide (https://www.imf.org/en/Data/Statistics/RPPI-guide).

the total value of transactions within each stratum by the total value of all transactions across all strata.

Step 4. The reference period for the individual indices corresponds to the initial period, ensuring it spans a full year rather than a quarter.

Step 5. All indices, including stratum indices and the total RPPI for apartments for the current period, are chained by multiplying them with the chained index from the final period of the preceding year.

Time Dummy Hedonic method

The TDH method (also known as the rolling year hedonic regression model) is especially appropriate when compilers have detailed microdata with many explanatory characteristics but relatively few transactions in each period. The CSO of Ireland was among the first to adopt the time dummy hedonic method with a rolling window for compiling a house price index[5]. The selection of the TDH method in this study is primarily based on data availability and its relative simplicity compared to other types of hedonic methods.

Step 1. Dummy variables are created for each period and for categorical variables, as described in detail in

Step 2. Strata are defined in a manner similar to the median with stratification method.

Step 3. "Shadow" prices are calculated by exponentiating the coefficients of the regression with the natural logarithm of price as the dependent variable and all other variables as explanatory variables.

Step 4. Individual indices for each stratum are computed using the following formula:

$$I_t = exp(\widehat{\delta_t}) * 100 \qquad (5)$$

The coefficient δ_t has already described on page 5 in the formula $lnp_n^t = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{nk}^t + \varepsilon_n^t$ (. The coefficients of the other variables are they are only used for quality control and therefore are disregarded when deriving the index.

Step 5. The aggregation of the individual indices involves a weighted average, utilizing transaction value weights similar to the median with stratification method.

Step 6. All indices are chained similar to the median with stratification method.

[5] https://www.tara.tcd.ie/tara8/server/api/core/bitstreams/428e0f3e-a2ce-49bf-a9a7-fb4bdf52608f/content

Table 7. The characteristics which are used in TDH method for apartments

Characteristics	*Description*
Build_year	The year of construction of a building
Cons_material	Construction material of a building
Floor_Area	Total area in square metres
Region	The marz where the apartment is located
Zone	Appraisal Zones based on characteristics of the geographical area
Floor_level	Floor level of apartment in the building
Quarter	The dummies for the time period

Table 8. The characteristics which are used in TDH method for individual houses

Characteristics	*Description*
Floor_Area	Total floor area in square metres
Land_Area	Total land area in square metres
Region	The marz where the house is located
Quarter	The dummies for the time period

The quality of regressions conducted for the calculation of RPPI is presented in Appendix (Figure A1 and Figure A2).

The figures below compare the discussed methods for the total RPPI of Armenia and two sub-indices, both for 2018 as a base year and year-on-year changes.

Figure 13. The RPPI with TD and Stratification methods for Total Armenia (2018Q1-2024Q4), 2018=100, %

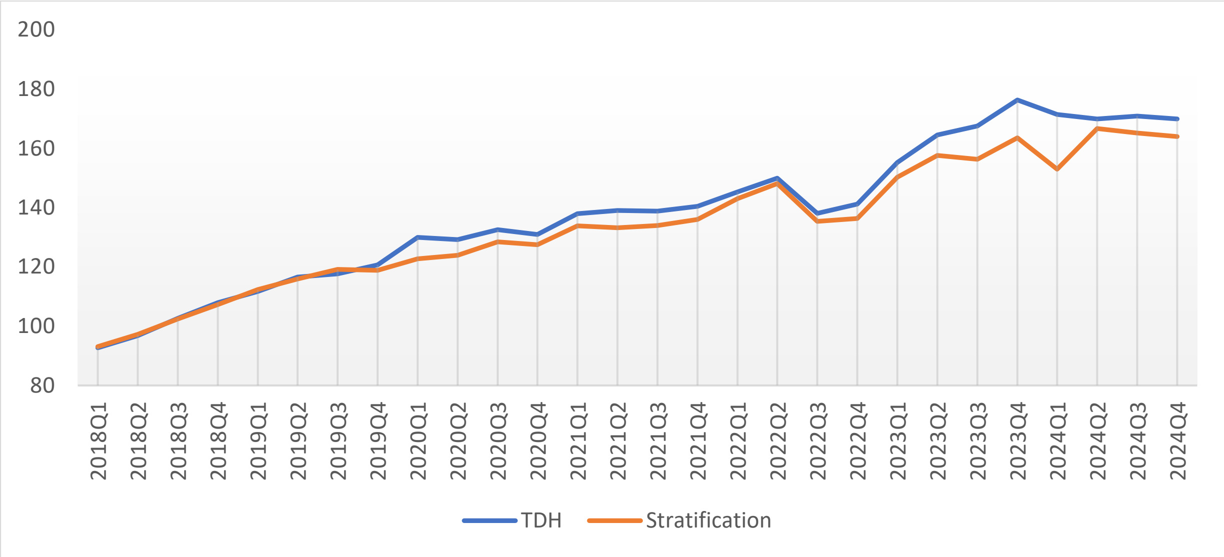
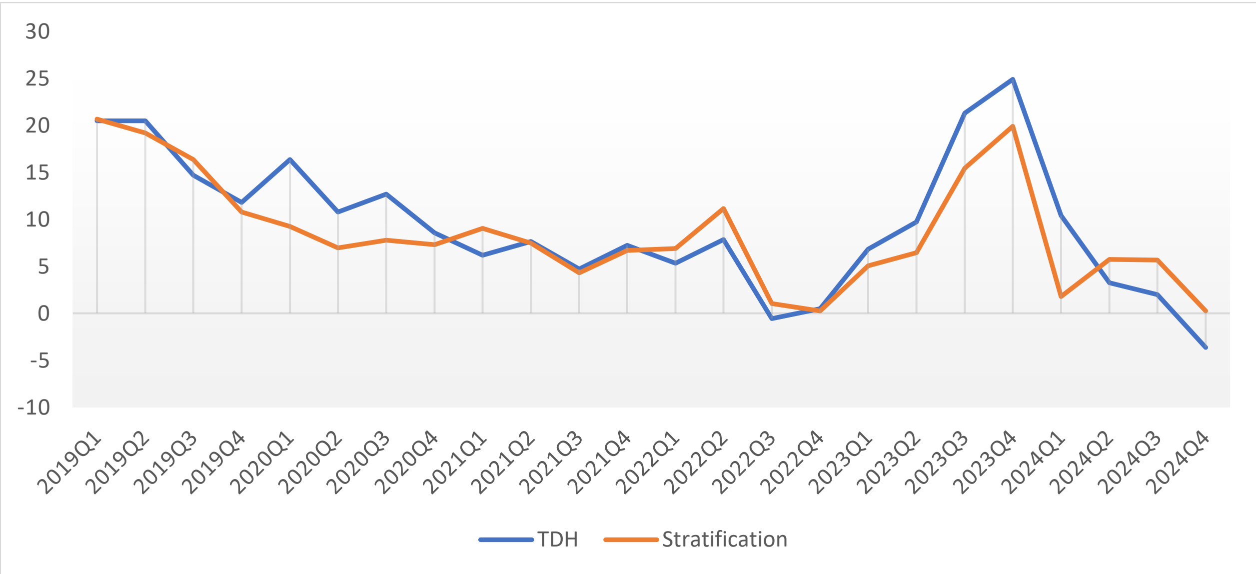


Figure 14. The RPPI with TDH and Stratification methods for Total Armenia (2019Q1-2024Q4), Year-on-year, %



Although the TDH-based index exhibits higher volatility (the variances of 12-month indices are 51.3 (TDH) and 33.9 (stratification method)) compared to the stratification-based index, it better captures quality differences over time. Therefore, the CBA has decided to adopt the TDH method over the stratification approach.

Conclusion and Future Work

The residential property price index holds significant importance for decision-making across various sectors of the economy, particularly given their interconnection with the financial system. Consequently, the Central Bank of Armenia has prioritized the development of this index.

This document shows the compilation process of a methodology for constructing the RPPI for Armenia. It emphasis the consideration of the specific nature of the domestic real estate market while considering recent methodological guidelines. Among various available data sources, the CCA dataset has been selected because of its full coverage, including all transactions, a rich set of appropriate characteristics for apartments and individual houses, and the timeliness, that this dataset is available for computation within 28 days of the reference quarter.

While the Handbook of RPPI proposes several methods, the focus has been on two main approaches: stratified median and time-dummy hedonic regression methods. The later was selected because it incorporates a larger set of characteristics and thus offers a stronger ability to capture for quality changes over time.

The national index is further split into two sub-indices for apartment and individual houses. Due to the large number of observations and the better quality characteristics for apartments, this segment is additionally split into two main stratums: Yerevan and Outside of Yerevan. Within Yerevan, the index is further stratified into four directional strata and a separate stratum for the inner city of Yerevan.

Moving forward, as more characteristics for individual houses becomes available, the CBA aims to compile also separate indices for houses at the region level.

The CBA’s ongoing efforts to improve data collection also aim at enabling the compilation of separate sub-indices for new and existing residential properties, thereby increasing the analytical value of the RPPI—particularly from a financial stability perspective.

References

1. Balsa J. J., Vásquez J. (2022). “Housing Price Index Central Bank of Chile”, Central Bank of Chile.
2. Brunauer W., Feilmayr W., Wagner K. (2017). “The Austrian residential property price index: Methodological note”, Oesterreichische National Bank.
3. Case B., Wachter S. (2005). “Residential real estate price indices as financial soundness indicators: methodological issues”, 2005, vol. 21, pp 197-211 from Bank for International Settlements.
4. El Mahmah A. (2013). “Construction a real estate price index: the Moroccan experience”, Bank for International Settlements.
5. EUROSTAT, ILO, IMF, OECD, UNECE, the World Bank (2013). Handbook on Residential Property Prices Indices (RPPIs).
6. Laferrère A. (2005). “Hedonic housing price indexes: the French experience”, 2005, vol. 21, pp 271-287 from Bank for International Settlements.
7. IMF (2020). RPPI Practical Compilation Guide.
8. Quality report on residential property price indices on the basis of data provided by bulwiengesa AG (2020), Bundesbank.
9. O’Hanlon N. (2011). “Constructing a National House Price Index for Ireland”, Journal of the Statistical and Social Inquiry Society of Ireland Vol. XL, Central Statistics Office and Centre for Policy Studies UCC.
10. Selected residential property price series – data documentation (2023), Bank for International Settlements.
11. Stojanova D., and etc. (2008). “Real Estate Prices in the Republic of Macedonia”, Munich Personal RePEc Archive.

Appendix

Figure A1. Dynamics of R squared for each stratum and rolling window for apartments

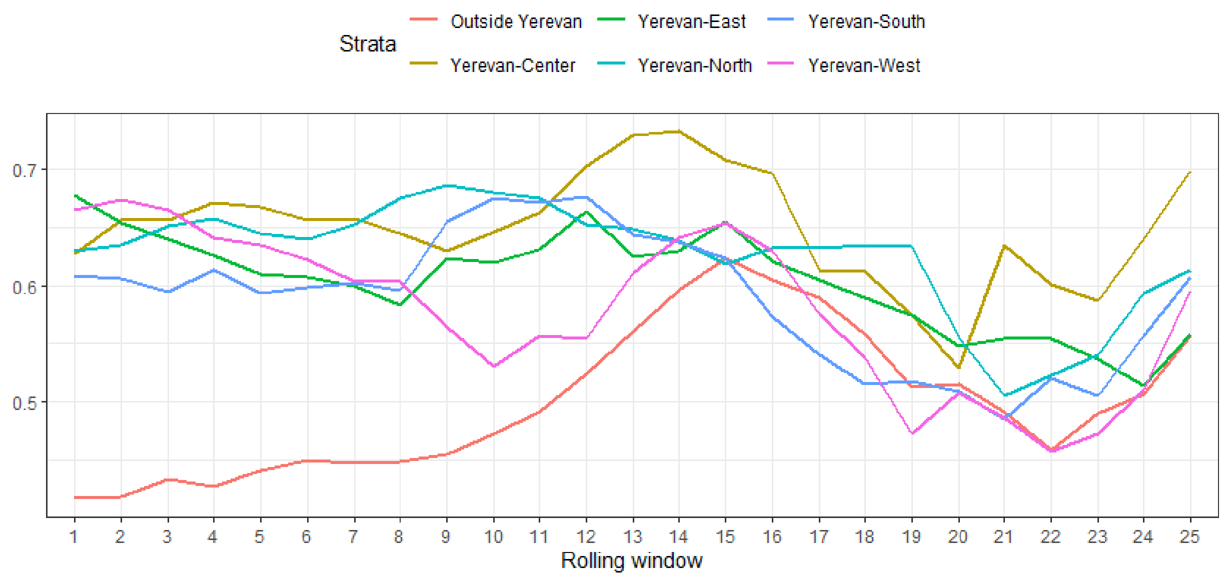


Figure A2. Dynamics of R squared for each stratum and rolling window for individual houses

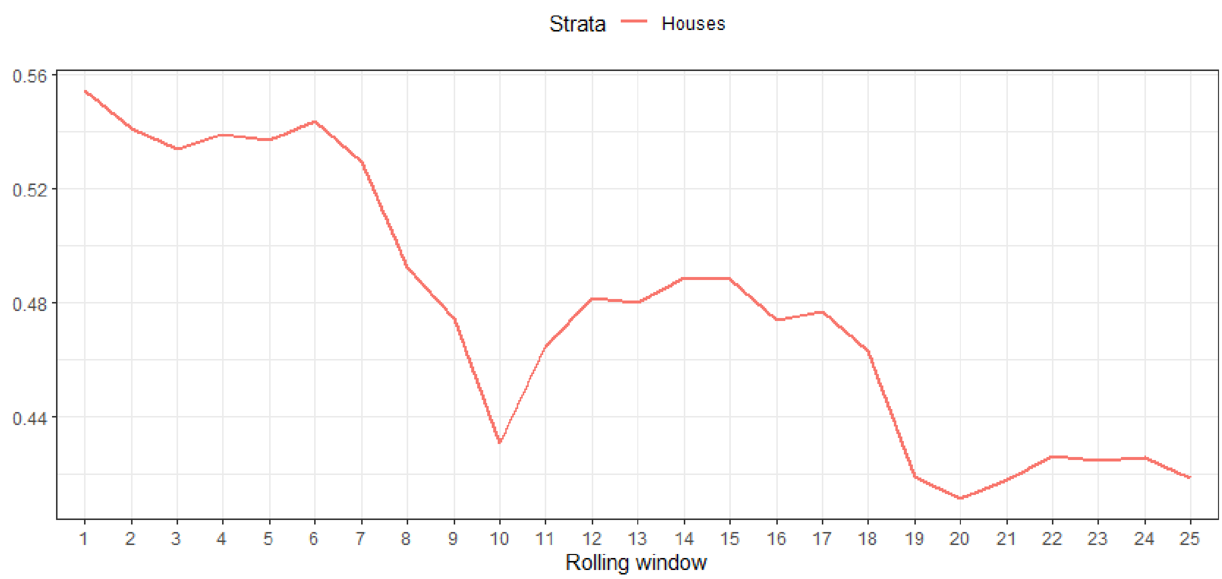


Figure A3. The RPPI for Apartments with TDH and Stratification (2018Q1-2024Q4), Base year: 2018, %

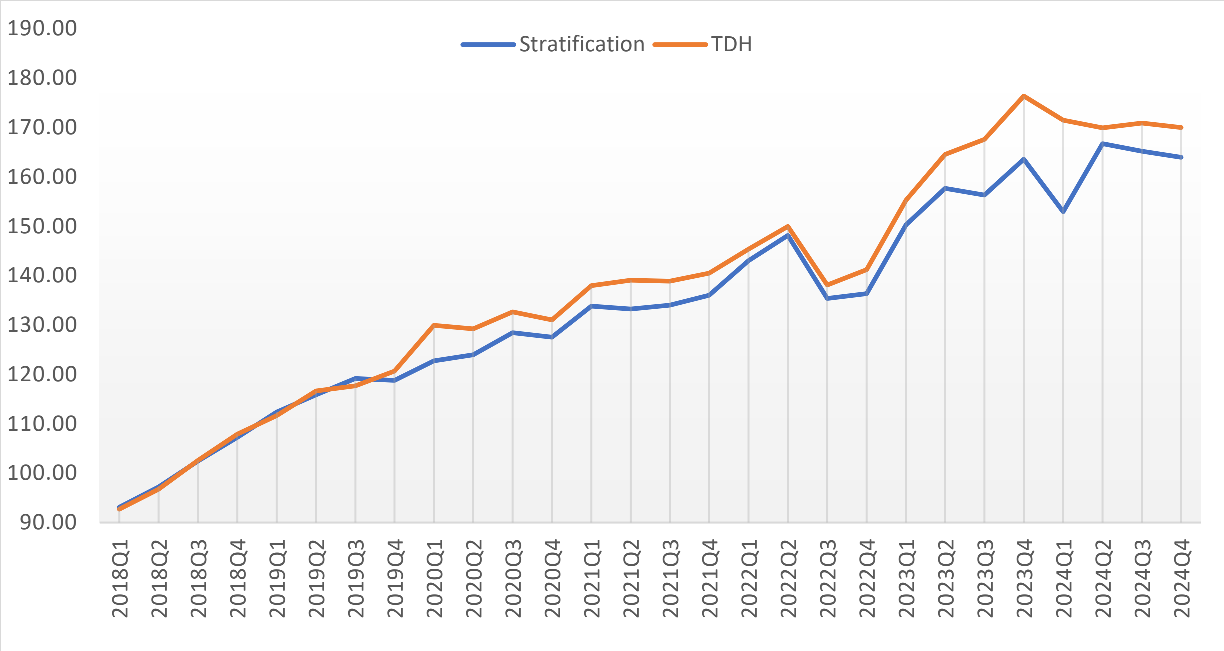


Figure A4. The RPPI for Apartments with TDH and Stratification (2019Q1-2024Q4), Year-on-year, %

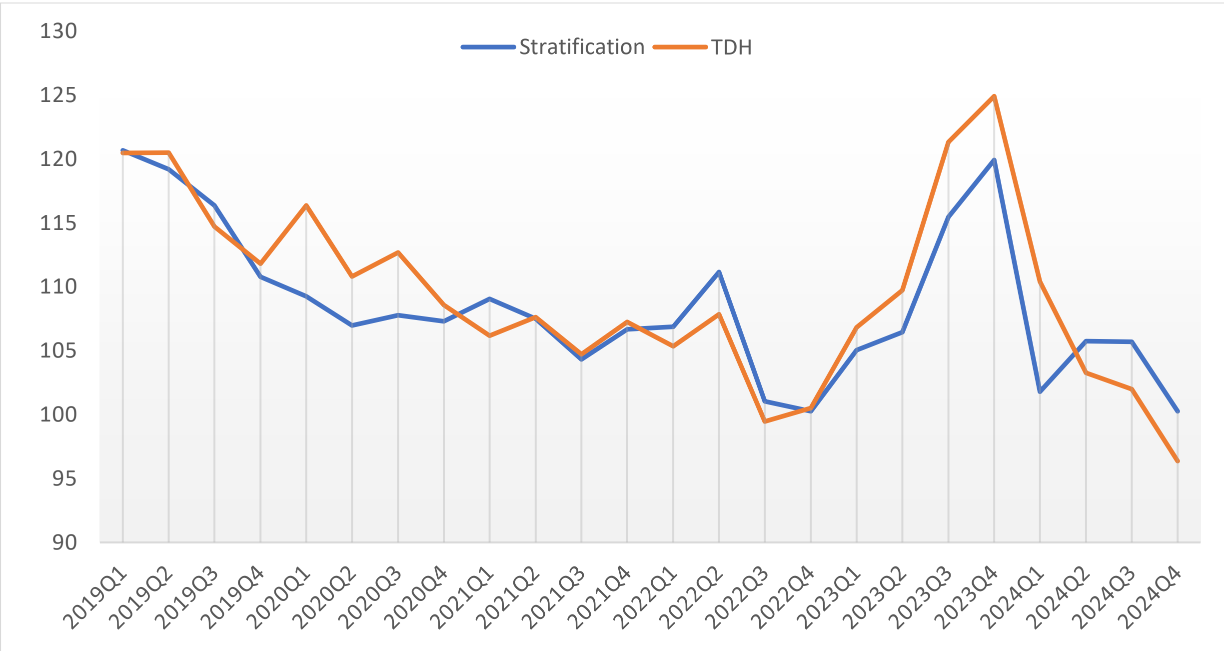


Figure A5. The RPPI for Individual Houses with TDH and Stratification methods (2018Q1-2024Q4), Base year: 2018, %

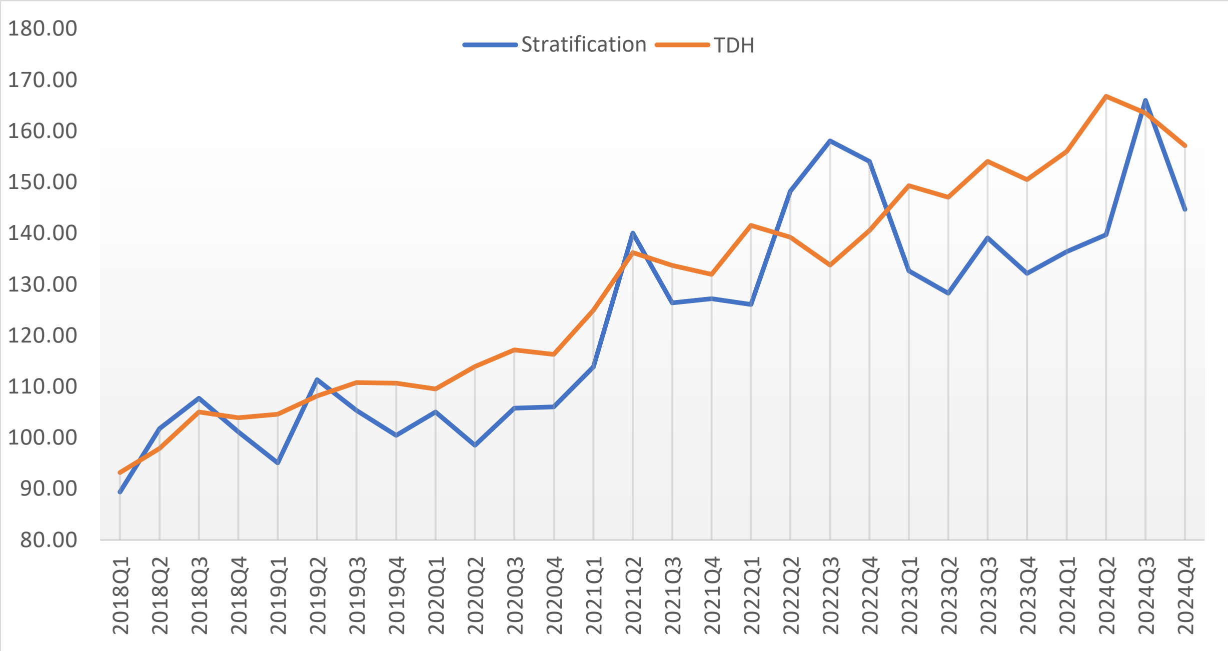


Figure A6. The RPPI for Individual Houses with TDH and Stratification methods (2019Q1-2024Q4), Year-on-year, %