

Article

Regional Housing Price Heterogeneity and RPPI Stratification in Uzbekistan

Dildora Akramovna Tog'ayeva

1. Independent Researcher, Institute for Staff Advanced Training and Statistical Research, National Statistics Committee of the Republic of Uzbekistan

*Correspondence: dildar.akramovna@gmail.com

Abstract: This study examines the magnitude and persistence of regional housing asking-price heterogeneity in Uzbekistan and evaluates its implications for geographic stratification in residential property price index (RPPI) compilation. The empirical panel contains 420 region-segment-quarter cells: 14 territorial units, three housing segments and 10 quarters from 2024Q1 to 2026Q2. Apartment prices are measured per square metre of floor area, while the available indicator for individual houses is price per square metre of land; the two dwelling types are therefore analysed separately. Regional heterogeneity is assessed using the coefficient of variation, the maximum-to-minimum price ratio, the Tashkent City price relative to the median of the other regions, and Spearman rank correlation of regional price rankings. In 2026Q2, cross-regional coefficients of variation were 50.65% for primary-market apartments, 36.71% for secondary-market apartments and 95.56% for individual houses. After Tashkent City was excluded, the corresponding values remained 29.07%, 22.59% and 36.38%. Spearman correlations between regional rankings in 2024Q1 and 2026Q2 were 0.662, 0.833 and 0.916, respectively. The results show that spatial price heterogeneity is persistent and cannot be reduced to a capital-versus-rest distinction. They support a two-stage spatial architecture for RPPI production: broad geographic-market stratification followed by more granular location controls within strata. The reported asking-price levels are not quality-adjusted RPPIs and are interpreted as diagnostics for index design rather than as estimates of pure price change.

Keywords: online listings, spatial dispersion, housing submarkets, coefficient of variation, official statistics

Citation: Tog'ayeva, D. A. Regional Housing Price Heterogeneity and RPPI Stratification in Uzbekistan. American Journal of Economics and Business Management 2026, 9(9), 543-555.

Received: 15th Jul 2026
Revised: 10th Aug 2026
Accepted: 30th Aug 2026
Published: 17th Sep 2026



Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Residential property prices are intrinsically spatial. A dwelling is fixed in location, and accessibility, neighbourhood quality, infrastructure, environmental conditions and local labour-market opportunities are capitalised into its price. Consequently, two dwellings with similar physical characteristics can have very different market values when they are located in different cities, districts or neighbourhoods. This spatial heterogeneity creates a direct measurement problem for residential property price indices: if the geographic composition of observed properties changes from one period to the next, an unadjusted average can move even when prices of comparable dwellings do not change [1].

International RPPI guidance therefore treats changes in dwelling quality and sample composition as central methodological issues. The Handbook on Residential

Property Price Indices distinguishes stratification, hedonic regression, repeat-sales and appraisal-based approaches as alternative or complementary ways of reducing quality-mix effects (Eurostat et al. 2013; European Commission–Eurostat 2017). Hedonic methods are especially useful because they relate price to measurable characteristics, including location (Rosen 1974; Hill 2013). Stratification addresses the same problem at a different level: observations are first divided into relatively homogeneous cells, and price change is measured within those cells before aggregation. In practice, stratification and hedonic modelling are often complements rather than substitutes [2].

Location matters not only because mean price levels differ across space but also because housing markets may consist of economically distinct submarkets. Bourassa, Hoesli and Peng (2003) show that market segmentation can improve hedonic prediction when submarkets are meaningfully defined, while Goodman and Thibodeau (2003) demonstrate the importance of spatially differentiated housing submarkets for prediction accuracy. More recent work emphasises geospatial information and flexible spatial surfaces in house-price measurement (Hill and Scholz 2018). Within the Regional Statistics literature, Safaralizadeh et al. (2024) likewise document substantial spatial variation in housing prices and show why geographically differentiated analysis can reveal patterns obscured by a single pooled model [3].

The rapid growth of online housing platforms creates new opportunities for such analysis. Listing data are typically more timely and geographically detailed than many traditional sources, and they can support experimental indicators of supply, demand and asking prices. At the same time, online data have important limitations, including duplicate advertisements, selective platform coverage and a gap between asking and transaction prices. Loberto, Luciani and Pangallo (2022) show both the analytical potential and the quality-control challenges of housing advertisement data. Eurostat (2026) has also examined web-scraped information as a potential source for more timely house-price and price-level indicators. These developments are particularly relevant for countries where complete transaction microdata with rich property characteristics are not yet available for every region and period [4].

Uzbekistan provides a useful setting for studying these issues because housing prices differ sharply between the capital, large regional centres and smaller regional markets. A national RPPI that ignores spatial heterogeneity risks mixing genuine price change with shifts in the territorial composition of observed listings or transactions. Yet a purely conceptual argument for geographic stratification is not enough. The case for stratification becomes stronger when the magnitude, persistence and structure of regional price differences are documented empirically, including after the dominant capital-city market is removed [5].

This paper therefore asks four questions. First, how large is cross-regional asking-price dispersion within each housing segment? Second, does material dispersion remain after Tashkent City is excluded? Third, is the relative ranking of regional price levels stable over time? Fourth, do primary- and secondary-market apartment price relationships vary sufficiently across regions to justify retaining market-type distinctions inside the geographic structure? The answers are used to draw implications for RPPI stratification and location controls rather than to estimate the RPPI itself.

The contribution is deliberately diagnostic. The paper separates regional asking-price levels from quality-adjusted price change. It does not claim that cross-sectional price dispersion is an RPPI, nor does it claim that high dispersion proves the exact size of composition bias. Instead, it uses a compact set of spatial diagnostics to evaluate whether the conditions under which geographic stratification is useful are present. This distinction reduces the risk of interpreting descriptive price-level differences as causal evidence about index growth and makes the results directly relevant to official-statistics design [6].

1.1 Conceptual background and literature.

A residential property price index seeks to measure change in the price of a constant-quality dwelling or a representative bundle of dwellings. The difficulty is that properties sold or advertised in consecutive periods are rarely identical. Observed average prices can therefore change for two reasons: the prices of comparable properties change, or the composition of the observed sample changes. If expensive properties account for a larger share of observations in one quarter, a simple mean may rise even when market values are unchanged. The opposite occurs when lower-priced properties become more prevalent. This composition problem is one of the main reasons why housing-price measurement requires methods beyond elementary averages [7].

Stratification reduces this problem by dividing the market into groups that are more homogeneous than the national market as a whole. The groups can be based on geography, dwelling type, new-versus-existing status, size class or combinations of these variables. Within a well-designed stratum, the variation in unobserved quality should be smaller and the composition of observations should be more stable. The RPPI Handbook notes, however, that very fine stratification creates its own problem: some cells may contain too few observations for stable index calculation (Eurostat et al. 2013). An operational design must therefore balance homogeneity against sample adequacy [8].

Hedonic regression provides a second layer of control. A standard hedonic model expresses the logarithm of property price as a function of structural and locational characteristics and, in time-dummy applications, period indicators. The approach is attractive because it can adjust for observed differences in floor area, number of rooms, building characteristics, floor, age, plot size and location. Rosen (1974) provides the economic foundation for interpreting observed product prices as the joint valuation of characteristics, while Hill (2013) surveys the main hedonic approaches used for residential price indices. The important point for the present study is that geography can enter both before estimation, through stratification, and inside the regression, through location variables [9].

These two layers are not redundant. A national model with a small number of coarse location dummies may be unable to capture different parameter relationships across local markets. Conversely, a very fragmented set of strata without hedonic adjustment may leave substantial within-cell quality variation uncontrolled. Bourassa et al [10]. (2003) conclude that housing submarkets matter for prediction and that location is a central reason why they matter. Goodman and Thibodeau (2003) similarly link spatial market segmentation to hedonic prediction accuracy. This literature supports an architecture in which broad submarkets are separated first and more granular location information is then used within the model.

Geospatial data can further improve this architecture. Administrative boundaries are convenient, but they do not always correspond to actual housing-market boundaries. Hill and Scholz (2018) show that geospatial information and flexible spline surfaces can improve hedonic house-price indices by allowing location value to vary continuously over space. Such approaches are especially useful in large metropolitan areas where a single city dummy is too coarse. The same logic applies to Tashkent City: treating the capital as one spatial unit may be adequate for national stratification, but a mature RPPI system can later test district, neighbourhood, distance-to-centre or other geocoded controls within the city [11].

Regional Statistics has published research that illustrates the broader relevance of spatial housing analysis. Safaralizadeh et al. (2024), using Tehran as a case study, compare ordinary least squares and geographically weighted regression and show that housing-price relationships vary across space. The present paper addresses a different question: it does not model the determinants of individual property prices, but evaluates whether regional price-level differences are sufficiently large and persistent to matter for

price-index design. The common theme is that spatial heterogeneity is not a nuisance to be ignored; it is an empirical feature that should shape model specification and interpretation.

Online listings add another dimension. They can provide high-frequency observations and detailed property characteristics, but they describe seller asking prices rather than completed transaction prices. Loberto et al. (2022) emphasise that listing datasets can be highly informative if duplicate ads and other platform-specific distortions are controlled. Eurostat (2026) similarly considers web-scraped data promising for experimental house-price level indicators. For official statistics, a prudent strategy is therefore to use listing data where they provide unique timeliness or detail, while clearly separating experimental asking-price indicators from transaction-based or quality-adjusted official indices. The present study follows that principle [12].

2. Materials and Methods

The empirical source is the regional price-level table contained in the author's RPPI methodology workbook. The table covers 14 territorial units of Uzbekistan, three housing segments and ten quarters from 2024Q1 through 2026Q2. The segments are primary-market apartments, secondary-market apartments and individual houses. The full analytical panel therefore contains $14 \times 3 \times 10 = 420$ region-segment-quarter cells. The analysis is conducted at this aggregated cell level; it does not re-estimate the underlying hedonic model from individual advertisements.

The price concept differs between apartments and individual houses. Apartment observations are expressed as asking price per square metre of total floor area. For individual houses, the available workbook indicator is asking price per square metre of land. The levels of these two measures are not directly comparable, and the paper never interprets a house price per square metre of land as equivalent to an apartment price per square metre of floor area. Individual houses are used only for within-segment comparisons across regions and over time. This measurement boundary is important because the house series displays the highest regional dispersion.

The regional cell values are geometric-mean asking-price levels. Geometric means are useful for positive, right-skewed price distributions because they reduce the influence of very large observations compared with arithmetic means. Nevertheless, the resulting regional levels are still descriptive. They are not constant-quality indices and do not control for all differences in the characteristics of properties advertised in each region. For that reason, the diagnostics are interpreted as evidence about the need for spatial controls, not as direct estimates of spatially adjusted price change [13].

The first diagnostic is the coefficient of variation. For segment s in quarter t , $CV_{st} = 100 \times SD(P_{rst}) / \text{Mean}(P_{rst})$, where P_{rst} denotes the regional price level and the standard deviation is calculated across regions r . The CV is scale-free and therefore permits comparison of relative regional dispersion across segments with very different price levels. Two versions are calculated: one for all 14 regions and one for the 13-region set after Tashkent City is excluded. The second version is a simple but powerful sensitivity check because it shows whether the national result is driven entirely by the capital.

The second diagnostic is the maximum-to-minimum ratio $H_{st} = \max(P_{rst}) / \min(P_{rst})$. The ratio has an intuitive economic interpretation: it shows how many times the highest observed regional price level exceeds the lowest within the same segment and quarter. Unlike the CV, it is based on two extreme regional observations and is therefore more sensitive to outliers. It is used as a complementary measure rather than as the principal dispersion statistic.

The third diagnostic isolates the capital-city premium. For each segment and quarter, $K_{st} = P_{\text{Tashkent},s,t} / \text{median}(P_{\text{other regions},s,t})$. The median of the other 13 regions is used rather than their mean so that the benchmark is not overly influenced by a few relatively expensive regional markets. A value of 1 would mean that Tashkent City is at the median of the other regions; values materially above 1 indicate a capital premium in the observed asking-price level.

The fourth diagnostic measures persistence in the spatial ordering of regions. Spearman's rank correlation is calculated between the 14 regional price levels in 2024Q1 and the corresponding levels in 2026Q2 for each segment. The statistic evaluates whether regions that were relatively expensive at the beginning of the period tended to remain relatively expensive at the end. High rank persistence matters for index design because it suggests that regional differences are not merely temporary noise. Spearman correlation is interpreted descriptively and does not establish a causal mechanism.

For apartments, a fifth diagnostic compares the primary and secondary markets. The 2026Q2 primary-to-secondary ratio is calculated for each region. If this ratio were nearly identical everywhere, separate market treatment might add little to the spatial structure. If it varies substantially, national pooling can conceal regional differences in the relative pricing of new and existing apartments. Individual houses are excluded from this comparison because their measurement denominator differs.

No universal threshold is imposed for an 'acceptable' CV or max/min ratio. Such cut-offs would be arbitrary without a loss function that links dispersion to index bias, revision size or prediction error. Instead, the decision logic is comparative. Evidence for retaining spatial controls becomes stronger when four conditions occur together: dispersion is economically large, dispersion remains material after the capital is removed, regional rankings are persistent, and market-segment relationships vary across regions. The present analysis evaluates this joint pattern [14].

The methodology also distinguishes diagnostic evidence from validation evidence. A diagnostic can show that spatial heterogeneity exists and is persistent. It cannot quantify how much the national RPPI would change under alternative strata. The stronger validation exercise, proposed for future research, is to recompute the full RPPI under competing geographic partitions while holding other components of the index system constant, and then compare bias-variance indicators, model fit, revision magnitude and sample adequacy.

3. Results and Discussion

The results indicate substantial cross-regional asking-price heterogeneity in every housing segment. Table 1 reports the coefficient of variation at the beginning and end of the study period, with and without Tashkent City, and the 2026Q2 maximum-to-minimum ratio. The capital is clearly important, but it is not the only source of spatial dispersion.

Table 1. Cross-regional dispersion of asking-price levels, 2024Q1–2026Q2

Segment	CV 2024Q1 (%)	CV 2026Q2 (%)	CV* 2024Q1 (%)	CV* 2026Q2 (%)	Max/min 2026Q2
Primary-market apartments	51.96	50.65	30.59	29.07	3.94
Secondary-market apartments	46.23	36.71	27.48	22.59	3.08
Individual houses	100.92	95.56	37.55	36.38	9.72

Note: CV excludes Tashkent City (13 regions). Source: author's calculations.*

For primary-market apartments, the all-region CV was 51.96% in 2024Q1 and 50.65% in 2026Q2. Excluding Tashkent City reduced the corresponding values to 30.59% and 29.07%. The decline caused by excluding the capital is substantial, yet a 29.07% CV among the remaining 13 regions still represents meaningful spatial heterogeneity. The 2026Q2 maximum-to-minimum ratio was 3.94, so the highest regional primary-apartment asking-price level was almost four times the lowest.

Secondary-market apartments were less dispersed than primary-market apartments by the end of the period. Their all-region CV declined from 46.23% in 2024Q1 to 36.71% in 2026Q2. The CV excluding Tashkent City declined from 27.48% to 22.59%. The 2026Q2 maximum-to-minimum ratio was 3.08. This pattern suggests some narrowing of regional differences in the secondary apartment segment, but the residual dispersion remains large enough that a pooled national price level would be highly sensitive to shifts in the geographic composition of listings [15].

Individual houses show the largest relative dispersion. Their all-region CV was 100.92% in 2024Q1 and 95.56% in 2026Q2. Removing Tashkent City reduced the measure sharply, to 37.55% and 36.38%, respectively. The 2026Q2 maximum-to-minimum ratio was 9.72. These magnitudes should be interpreted cautiously because the house indicator is defined per square metre of land, making it particularly sensitive to regional differences in plot values and settlement structure. Even so, the persistence of a 36.38% CV outside the capital indicates that the result cannot be explained by a single Tashkent outlier.

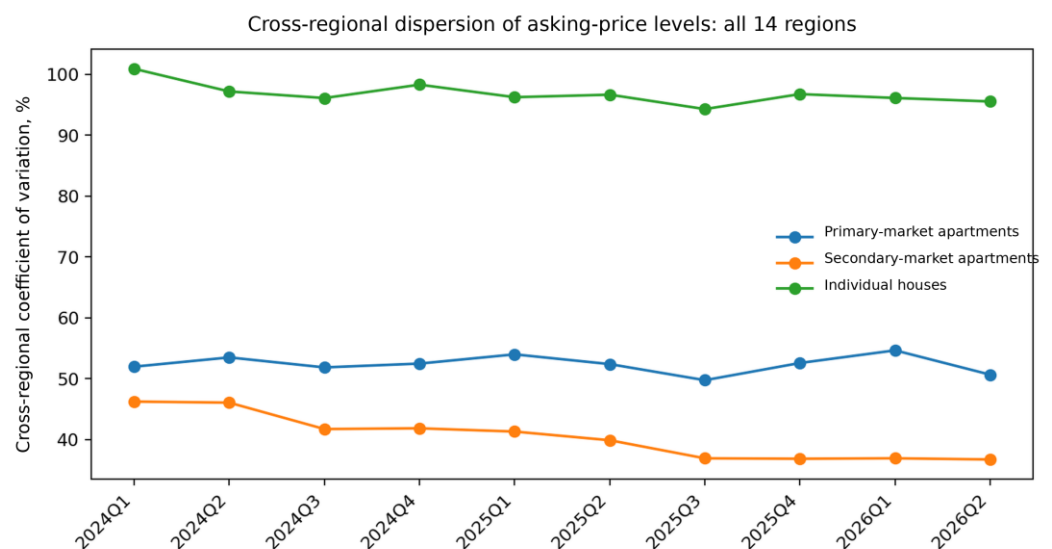


Figure 1. Cross-regional coefficient of variation of asking-price levels, all 14 regions, 2024Q1–2026Q2.

Source: author's calculations.

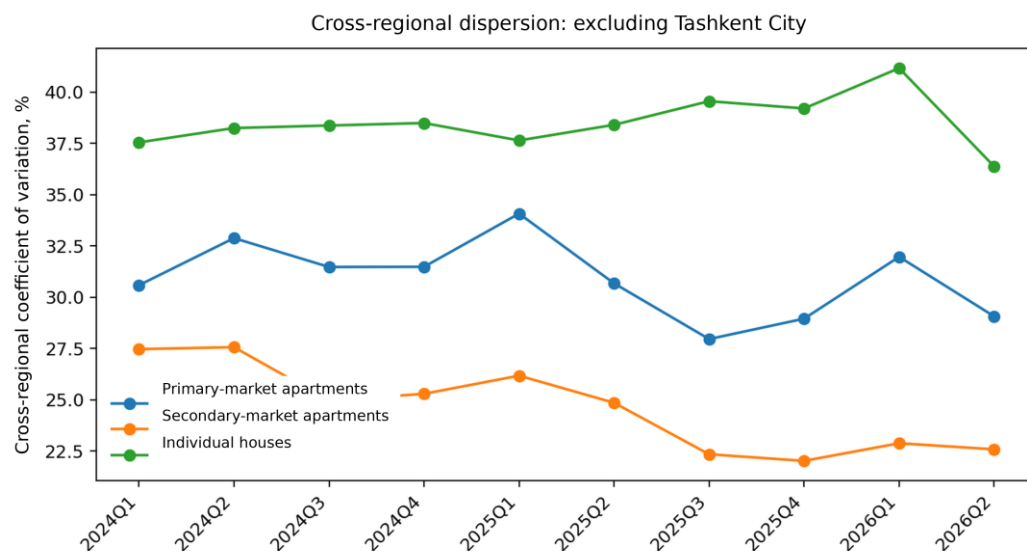


Figure 2. Cross-regional coefficient of variation of asking-price levels after excluding Tashkent City, 2024Q1–2026Q2. Source: author's calculations.

The temporal profiles in Figures 1 and 2 reinforce the endpoint comparison. With all 14 regions included, the three segments occupy clearly different dispersion ranges, and individual houses remain the most heterogeneous. After Tashkent City is excluded, the absolute dispersion falls in every segment, but none of the series collapses toward zero. The sensitivity analysis therefore changes the magnitude of the result without changing its substantive interpretation: regional heterogeneity is a national feature, not only a capital-city feature.

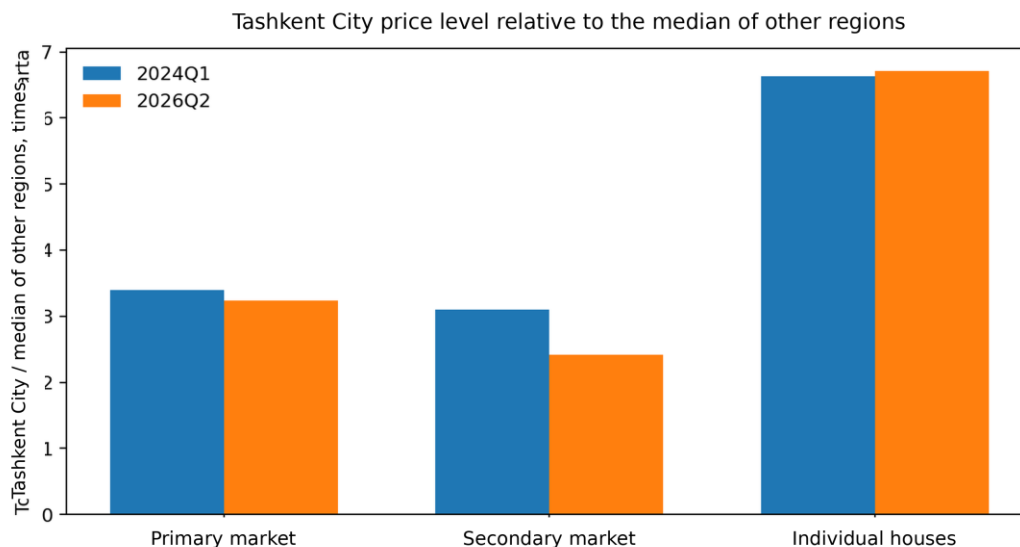


Figure 3. Tashkent City asking-price level relative to the median of the other regions, 2024Q1 and 2026Q2. Source: author's calculations.

The capital premium gives the same message in a more intuitive form. In 2026Q2, the Tashkent City asking-price level was 3.24 times the median of the other 13 regions for primary-market apartments and 2.42 times for secondary-market apartments. For individual houses, the ratio reached 6.71. Figure 3 compares the capital premium at the beginning and end of the period. The purpose of this indicator is not to label the capital as an outlier to be removed from the index. Tashkent is economically important and should remain in national measurement. Rather, the indicator shows why national averages can be strongly affected by changes in the share of observations originating from the capital.

Table 2. Highest and lowest regional asking-price levels by segment, 2026Q2

Segment	Highest three regions, 2026Q2	Lowest three regions, 2026Q2	Unit
Primary-market apartments	Tashkent City – 21.82; Samarkand – 12.51; Tashkent Region – 12.11	Fergana – 5.54; Andijan – 6.11; Syrdarya – 6.24	UZS million/m ² floor area
Secondary-market apartments	Tashkent City – 16.31; Samarkand – 11.55; Tashkent Region – 9.26	Karakalpakstan – 5.30; Jizzakh – 6.07; Syrdarya – 6.29	UZS million/m ² floor area
Individual houses	Tashkent City – 9.63; Samarkand – 2.92; Khorezm – 2.61	Karakalpakstan – 0.99; Surkhandarya – 1.13; Navoi – 1.25	UZS million/m ² land

Note: apartment prices are per square metre of floor area; individual-house prices are per square metre of land. Source: author's calculations.

Table 2 identifies the highest- and lowest-price regions in 2026Q2. Tashkent City recorded the highest level in all three segments. Samarkand was also in the high-price group for primary and secondary apartments, while Tashkent Region ranked among the highest apartment markets. The lowest primary-market apartment levels were observed in Fergana, Andijan and Syrdarya; the lowest secondary-market levels were in Karakalpakstan, Jizzakh and Syrdarya. For individual houses, Karakalpakstan, Surkhandarya and Navoi occupied the lower end of the observed distribution. These rankings illustrate that the spatial structure is more complex than a single capital/non-capital split.

Rank persistence provides additional evidence. Spearman correlation between the regional price ranking in 2024Q1 and 2026Q2 was 0.662 for primary-market apartments, 0.833 for secondary-market apartments and 0.916 for individual houses. The lower coefficient for primary apartments is consistent with a segment that can be more sensitive to new construction, project launches and changing development locations. The very high coefficients in the secondary-apartment and individual-house segments show that their relative regional price hierarchy was strongly persistent over the study horizon.

Table 3. Primary-to-secondary apartment asking-price ratio by region, 2026Q2

Region	Primary / secondary asking-price ratio
Fergana	0.841
Andijan	0.929
Kashkadarya	0.988
Navoi	0.992
Syrdarya	0.993
Namangan	0.999
Khorezm	1.003
Bukhara	1.022
Surkhandarya	1.040
Samarkand	1.083
Jizzakh	1.087
Karakalpakstan	1.284

Tashkent Region	1.307
Tashkent City	1.337

Note: median across 14 regions = 1.013; only apartment segments are compared. Source: author's calculations.

The primary-to-secondary comparison further demonstrates spatial variation in market structure. The median regional ratio in 2026Q2 was 1.013, suggesting that in a typical region the two apartment price levels were close. However, the regional range was wide. Fergana had a ratio of 0.841, meaning that the primary-market asking-price level was below the secondary-market level. At the other end, the ratio reached 1.337 in Tashkent City, 1.307 in Tashkent Region and 1.284 in Karakalpakstan. A single national assumption about the premium of new apartments over existing apartments is therefore not valid across all regions.

Taken together, the diagnostics satisfy the joint decision logic set out in the methodology. Cross-regional dispersion is large; it remains material outside the capital; regional rankings are persistent, especially in the secondary-apartment and house segments; and the relationship between primary and secondary apartment prices varies across space. The evidence is therefore consistent with retaining geographic and market-type controls in RPPI compilation.

3.1 Discussion and implications for RPPI design.

The first implication is that geographic stratification should be treated as a structural component of the measurement system rather than as a cosmetic reporting breakdown. When regional price levels differ by factors of three, four or more within the same housing segment, changing regional sample shares can materially affect an unadjusted national average. Stratification reduces this risk by comparing observations inside more homogeneous geographic-market cells before aggregation. The result is conceptually aligned with international guidance on controlling quality-mix change (Eurostat et al. 2013; IMF 2019).

The second implication is that a national capital-versus-rest split would be too coarse. Tashkent City has a very large premium, but the exclusion exercise shows that considerable dispersion persists among the other 13 regions. An RPPI architecture should therefore recognise multiple regional markets rather than treating every non-capital observation as belonging to one homogeneous pool. This does not mean that every administrative region must always be a separate stratum. Where sample sizes are small, economically similar regions can be combined, but the decision should be based on explicit homogeneity and sample-adequacy criteria.

The third implication concerns the interaction between stratification and hedonic modelling. Geographic strata can control broad differences between markets, while hedonic location variables control remaining heterogeneity within strata. This two-stage architecture is preferable to an either-or choice. In a large city, for example, the national system can first separate Tashkent City from other markets and then include district, neighbourhood or geocoded location variables in the Tashkent hedonic model. Hill and Scholz (2018) show why continuous geospatial information can be valuable when simple administrative dummies do not capture local price gradients.

The fourth implication is that primary and secondary apartment markets should not automatically be pooled. The national median primary-to-secondary ratio is close to one, but regional ratios differ substantially. In some regions, primary-market asking prices are below secondary-market levels, while in others they carry a sizeable premium. Pooling the two segments without interactions could impose an average relationship that is inappropriate for several regional markets. Separate strata or market-type interactions allow this heterogeneity to be represented more transparently.

The fifth implication concerns weights. Stratification only improves national measurement if stratum indices are aggregated with representative weights. A high-price region with a large transaction share should influence the national index more than a small market, but this influence should enter through an explicit weighting system rather than through accidental changes in the number of online listings. The distinction is crucial when listing intensity differs across platforms, regions or quarters. Transaction counts and appropriate value proxies are therefore preferable to raw advertisement shares when constructing upper-level weights.

The sixth implication is operational. A production RPPI needs a rule for reviewing strata over time. The diagnostics used here can form part of such a review dashboard. For each region and market segment, compilers can monitor sample size, price-level dispersion, coefficient stability, prediction error and the effect of alternative cell definitions. A stratum that becomes too sparse can be merged with an economically similar neighbour, while a very heterogeneous or rapidly growing urban market can be split when observations are sufficient. The decision and its effective date should be documented in metadata to preserve transparency.

The seventh implication is that price-level indicators and the RPPI should remain separate statistical products. A regional asking-price table can be useful for market monitoring, spatial analysis and early warning because it is intuitive and timely. But it answers a different question from a quality-adjusted index. The price-level table asks where observed asking prices are high or low; the RPPI asks how the price of comparable housing changes over time. Publishing or internally monitoring both measures can improve interpretation as long as their concepts are clearly labelled.

Finally, the results suggest a path toward more advanced spatial validation. Administrative regions are a practical starting point, but future work can compare them with statistically derived submarkets, district-level cells or geospatial surfaces. The objective should not be to maximise geographic detail for its own sake. The objective is to find the spatial structure that produces the most stable constant-quality measurement while retaining sufficient observations and operational simplicity.

3.2 Robustness, limitations and further research

The exclusion of Tashkent City is the principal robustness check in this paper. It addresses an obvious concern: because the capital has much higher price levels, the all-region CV might simply be measuring one extreme observation. The results reject that narrow interpretation. Dispersion falls when Tashkent is excluded, as expected, but remains substantial in all three segments. This strengthens the qualitative conclusion that spatial controls are needed beyond a capital dummy.

A second robustness dimension is temporal persistence. The study does not rely on one quarter alone. The data cover ten consecutive quarters, and the endpoint rank correlations are positive and sizeable in all segments. The 0.833 correlation for secondary-market apartments and 0.916 for individual houses are particularly strong. Persistent ranking does not prove that the same causal factors operate in every period, but it makes it less plausible that the regional price hierarchy is only a transient sampling artefact.

Several limitations remain. First, the source consists of asking prices from online listings rather than verified transaction prices. Sellers may revise asking prices, properties may remain unsold, and platform coverage may differ across regions. Online listing data are valuable precisely because they are timely and granular, but they require quality controls such as duplicate detection and coverage monitoring (Loberto et al. 2022). Results based on asking prices should therefore be interpreted as market signals rather than final transaction valuations.

Second, the regional price-level indicators are not quality-adjusted. Differences in dwelling size, age, renovation status, construction material, floor, plot size or neighbourhood characteristics can contribute to the measured regional dispersion. This limitation does not invalidate the diagnostic role of the results; indeed, systematic differences in housing composition are one reason stratification can be useful. But it means that the CV cannot be interpreted as a pure location effect. A stronger decomposition would combine micro-level hedonic estimates with spatial fixed effects or geospatial terms.

Third, the geographic unit is the region. Important heterogeneity exists within regions and especially within Tashkent City. A single regional mean can conceal sharp differences between central and peripheral districts, transport corridors and neighbourhoods with different infrastructure. Future work should test district and neighbourhood controls where geocoding is sufficiently complete. Spatial cross-validation can then compare administrative dummies with continuous distance measures or flexible surfaces.

Fourth, the house indicator uses land area in the denominator, whereas apartment indicators use floor area. This is why the paper does not compare house and apartment price levels directly. The exceptionally high house dispersion may partly reflect differences in land scarcity, plot size and urbanisation patterns. A future individual-house hedonic model should include both building and plot characteristics and should evaluate whether price-per-land-area is the most informative descriptive indicator for spatial diagnostics.

Fifth, the current study does not recompute the RPPI under alternative geographic structures. This is the most important remaining validation task. A convincing index-level experiment would estimate the same hedonic specification under at least three spatial designs: a highly pooled national model, the operational regional stratification, and a more granular alternative where data permit. The resulting sub-indices could be compared on prediction error, coefficient stability, revision magnitude, effective sample size and national index differences. Such an experiment would directly quantify the bias-variance trade-off of stratification.

Sixth, inferential uncertainty around the descriptive indicators is not estimated. The analytical unit is an aggregated region-segment-quarter price level rather than a probability sample of independent regional observations. Bootstrap or repeated-sampling inference would require access to the underlying listing-level data and a clear resampling design that respects duplicates and temporal dependence. Until that is implemented, the numerical diagnostics should be read as descriptive evidence rather than as formal hypothesis tests.

Further research can also connect stratification design to hierarchical models. Partial pooling can be useful when some regional strata are data-rich and others are sparse. A multilevel hedonic model could preserve regional heterogeneity while borrowing statistical strength across related markets. This may offer a compromise between a fully pooled national regression and completely separate regressions for every small cell. Such models would be especially valuable if the future RPPI expands to district-level geography.

4. Conclusion

This study evaluates the empirical case for geographic stratification in Uzbekistan's residential property price measurement using a ten-quarter panel of regional asking-price levels. The central finding is that housing prices are spatially heterogeneous in a persistent and economically meaningful way. In 2026Q2, the all-region coefficient of variation was 50.65% for primary-market apartments, 36.71% for secondary-market

apartments and 95.56% for individual houses. Excluding Tashkent City reduced these values but did not eliminate dispersion: the corresponding 13-region CVs remained 29.07%, 22.59% and 36.38%.

The spatial hierarchy is also persistent. Spearman correlations between regional rankings in 2024Q1 and 2026Q2 were 0.662, 0.833 and 0.916 across the three segments. Moreover, the relative price of primary and secondary apartments varies by region, with the 2026Q2 ratio ranging from 0.841 in Fergana to 1.337 in Tashkent City. These patterns are inconsistent with the idea that one national pooled market or one capital-versus-rest split can fully represent the spatial structure of Uzbekistan's housing market.

For RPPI compilation, the evidence supports a two-stage spatial approach: broad geographic and market-type stratification followed by more detailed location controls within strata. The purpose is not to reproduce observed price-level differences in the index, but to prevent shifts in the composition of observed properties from being misread as pure price change. Representative upper-level weights, transparent strata-review rules and clear separation between asking-price indicators and the quality-adjusted RPPI are essential complements to this approach.

The paper remains diagnostic rather than causal. The next research step is to recompute the RPPI under alternative geographic partitions using the same microdata and model framework and to quantify the resulting changes in prediction error, coefficient stability, revision and national index growth. That exercise would convert the present evidence on spatial heterogeneity into a direct index-level sensitivity analysis and provide a stronger empirical basis for the long-run design of Uzbekistan's RPPI.

REFERENCES

- [1] S. C. Bourassa, M. Hoesli, and V. S. Peng, "Do housing submarkets really matter?," *Journal of Housing Economics*, vol. 12, no. 1, pp. 12–28, 2003, doi: 10.1016/S1051-1377(03)00003-2.
- [2] European Commission and Eurostat, *Technical Manual on Owner-Occupied Housing and House Price Indices*. Luxembourg: Publications Office of the European Union, 2017. [Online]. Available: <https://ec.europa.eu/eurostat/documents/7590317/0/Technical-Manual-OOH-HPI-2017/>
- [3] Eurostat, International Labour Organization, International Monetary Fund, OECD, UNECE, and World Bank, *Handbook on Residential Property Price Indices (RPPIs)*. Luxembourg: Publications Office of the European Union, 2013, doi: 10.2785/34007.
- [4] Eurostat, *Feasibility of Monthly House Price Indices and Price-Level Indicators from Web-Scraped Data*, Statistical Working Papers, KS-01-26-025. Luxembourg: Publications Office of the European Union, 2026, doi: 10.2785/5813495.
- [5] A. C. Goodman and T. G. Thibodeau, "Housing market segmentation and hedonic prediction accuracy," *Journal of Housing Economics*, vol. 12, no. 3, pp. 181–201, 2003, doi: 10.1016/S1051-1377(03)00031-7.
- [6] R. J. Hill, "Hedonic price indexes for residential housing: A survey, evaluation and taxonomy," *Journal of Economic Surveys*, vol. 27, no. 5, pp. 879–914, 2013, doi: 10.1111/j.1467-6419.2012.00731.x.
- [7] R. J. Hill and M. Scholz, "Can geospatial data improve house price indexes? A hedonic imputation approach with splines," *Review of Income and Wealth*, vol. 64, no. 4, pp. 737–756, 2018, doi: 10.1111/roiw.12303.
- [8] International Monetary Fund, *Residential Property Price Index (RPPI) Practical Compilation Guide*. Washington, DC, USA: Statistics Department, International Monetary Fund, 2019. [Online]. Available: <https://www.imf.org/en/data/statistics/rppi-guide>
- [9] M. Loberto, A. Luciani, and M. Pangallo, "What do online listings tell us about the housing market?," *International Journal of Central Banking*, vol. 18, no. 4, pp. 325–377, 2022. [Online]. Available: <https://www.ijcb.org/journal/v18n4/what-do-online-listings-tell-us-about-housing-market>
- [10] S. Rosen, "Hedonic prices and implicit markets: Product differentiation in pure competition," *Journal of Political Economy*, vol. 82, no. 1, pp. 34–55, 1974, doi: 10.1086/260169.

-
- [11] E. Safaralizadeh, F. Janakipour, S. Nikkhoo, M. Amimi, and M. Zarghamfard, "Factors affecting housing prices in metropolitan regions: The case of Tehran, 2021," *Regional Statistics*, vol. 14, no. 1, pp. 130–158, 2024, doi: 10.15196/RS140107.
 - [12] M. J. Bailey, R. F. Muth, and H. O. Nourse, "A regression method for real estate price index construction," *Journal of the American Statistical Association*, vol. 58, no. 304, pp. 933–942, 1963, doi: 10.1080/01621459.1963.10480679.
 - [13] K. E. Case and R. J. Shiller, "Prices of single-family homes since 1970: New indexes for four cities," National Bureau of Economic Research, Cambridge, MA, USA, Working Paper 2393, 1987, doi: 10.3386/w2393.
 - [14] W. E. Diewert, J. de Haan, and R. Hendriks, "Hedonic regressions and the decomposition of a house price index into land and structure components," *Econometric Reviews*, vol. 34, nos. 1–2, pp. 106–126, 2015, doi: 10.1080/07474938.2014.944791.
 - [15] M. Silver, "How to better measure hedonic residential property price indexes," International Monetary Fund, Washington, DC, USA, IMF Working Paper WP/16/213, 2016. [Online]. Available: IMF publication