Characteristics of a Sufficient Statistic to Measure City Housing Prices

Prepared for the NBER CRIW Conference on Measurement of Housing and the Housing Sector, Alexandria, VA, March 12–13, 2026

Daniel Broxterman*
Florida State University

William Larson†
US Department of the Treasury

Anthony Yezer‡
George Washington University

March 2, 2026

Abstract

For a variety of empirical purposes, it is important to be able to characterize the levels of and changes to housing prices in cities, whether measured using rents or asset values. This task is complicated by the heterogeneity of the housing stock, the fact that neighborhood is consumed jointly with housing, and differences in accessibility. This paper concentrates on the issue of intra-city location which, based on economic theory, is systematically related to housing prices. The final conclusion is that a sufficient statistic to describe both the level of and change in the average housing price requires that prices be aggregated from relatively homogeneous market areas and weighted by characteristics such as units or interior space. Commonly used repeat-sales and hedonic measures of price change are generally not weighted in this fashion, but could be modified to do so.

JEL Classifications: R14, R31

Keywords: housing price change, repeat sales, standard urban model

*Florida State University, Tallahassee, FL, USA (dbroxterman@fsu.edu)

†US Department of the Treasury, Washington, DC, USA (william.larson@ofr.treasury.gov). Views and opinions expressed are those of the authors and do not necessarily represent official positions or policy of the Office of Financial Research or the USDepartment of the Treasury.

‡George Washington University, Washington, DC, USA (yezer@gwu.edu)

1 Introduction

Hedonic and repeat-sales indexes are the two principal approaches used in OECD countries to address the fundamental index-number problem of separating price change from quality change in metropolitan house price measurement. A substantial literature evaluates the strengths and limitations of these approaches and proposes refinements to the empirical estimator.

In contrast, this paper derives from urban spatial theory index-construction conditions that are required to recover average citywide price change. Spatial equilibrium implies two distinct restrictions: one governing spatial aggregation of price changes within a city, and another governing the definition of the housing quantity to be indexed.

In the standard monocentric city model with constant marginal commuting costs, the iso-utility condition implies that housing price changes vary systematically with distance from the city center. Unless highly restrictive preference assumptions are imposed, appreciation rates cannot be spatially invariant.

Extensive empirical evidence confirms this prediction. Using granular transaction and rental data, Ahlfeldt et al. (2023) show that housing prices and rents can evolve very differently across locations, even over short distances. Related studies document persistent differences in appreciation across central-city and suburban neighborhoods in US cities during the 2000s and 2010s, with faster growth and greater volatility in more central locations (Glaeser et al., 2012; Malone and Redfearn, 2018; Bogin et al., 2019a,b; Edlund et al., 2022; Seagraves and Gatzlaff, 2025).

Spatial heterogeneity in housing price dynamics is also evident in responses to localized shocks. Infrastructure investments, energy price changes, monetary policy shocks, and the COVID-19 pandemic have generated pronounced within-city variation in appreciation patterns (Molloy and Shan, 2013; Larson and Zhao, 2020; Fischer et al., 2021; Liu and Su, 2021; D'Lima et al., 2022). Similarly, Baum-Snow and Han (2024) document substantial differences across census tracts in both appreciation rates and supply elasticity, reinforcing the link between intra-urban price dynamics and underlying spatial economic structure.

Such heterogeneity has direct implications for aggregation. When transaction probabilities are correlated with location-specific appreciation, commonly estimated city-level indexes need not recover average citywide housing price change and therefore fail to constitute sufficient statistics for that object. We show that average citywide price change must be constructed as a weighted aggregation of submarket price indexes (formulated here as Laspeyres indexes for illustration), where submarkets satisfy an appreciation-homogeneity condition and aggregation weights are proportional to each submarket's share of the housing stock. These restrictions follow from spatial equilibrium rather than from adjustments to the estimation procedure.

Despite these theoretical and empirical findings, most widely used hedonic and repeat-sales indexes estimate metropolitan-level appreciation using a single model applied to all transactions within the metro. Explicit geographic or compositional weighting is typically introduced only at higher levels of aggregation—most commonly when regional indexes are combined into a national series. As a result, within-city variation in appreciation is absorbed into a single metropolitan price change rather than explicitly aggregated across neighborhoods or housing types. When appreciation patterns differ systematically within a city and are correlated with transaction propensity, the resulting metropolitan estimate may reflect the spatial distribution of transactions rather than that of the housing stock. Implicitly, such approaches assume either spatially invariant appreciation or that observed transactions are proportionally representative of the housing stock. Urban theory provides little justification for the former assumption, and transaction-based sampling provides no guarantee of the latter. Consequently, commonly used metropolitan indexes may fail to satisfy the sufficient-statistic conditions implied by spatial equilibrium unless explicit stock-based spatial weighting is imposed.

A distinct restriction concerns the definition of the indexed housing quantity. Because housing services are not directly observable, empirical work must rely on measurable proxies.

The prevailing practice is to index housing units. Urban spatial theory, however, implies that this choice affects measured responsiveness and appreciation patterns. Liu (2018) shows that when housing consumption adjusts along the intensive margin, supply elasticity measured in square footage exceeds elasticity measured in housing units. Because households trade off commuting costs against the price per unit of space, not the price per dwelling, unit-based indexes understate effective supply responsiveness and may overstate the price effects of demand shocks. A space-based index is therefore more consistent with spatial equilibrium, except in applications focused narrowly on household formation.

Interior living space provides a tractable proxy for housing services. Space is observable, central to production costs, and closely tied to land use in standard urban models. Because unit size varies systematically across locations and over time, indexing housing units conflates price appreciation with shifts in the size distribution of the housing stock. A space-based index better isolates the price of housing on the intensive margin and aligns measurement with spatial equilibrium conditions.

More broadly, units and space capture only part of the attribute bundle that defines housing services. Home values reflect numerous observed and unobserved characteristics—including lot size, school quality, amenities, and environmental conditions—so housing services are inherently multidimensional and latent. Nevertheless, a housing-service-based index can be implemented within a Laspeyres framework.

Implementation of these theoretical conditions requires confronting heterogeneity in tenure, unit size, and other attributes. Rental prices are not observed for owner-occupied housing, and asset prices are not generally available for rental units. Unit size varies substantially across space, so price per room or per square foot may diverge markedly from unit-level prices. Because the distribution of rental and owner-occupied housing—and the distribution of unit sizes—is highly uneven within cities, aggregation that ignores these dimensions may fail to reflect the spatial distribution of housing services implied by theory.

These considerations may justify constructing separate indexes for rental and owner-occupied housing or choosing between indexes based on unit prices versus those based on housing space. The appropriate approach is ultimately empirical and depends on city-specific characteristics, including the prevalence of rental housing and the degree of variation in unit size. Using publicly available data, this paper illustrates how these definitional choices affect housing price indexes across US metropolitan areas. Although the empirical application focuses on the US, the underlying measurement issues apply broadly to city-level index construction in other countries.

Complementary evidence indicates that measurement and aggregation choices materially affect inferred housing price appreciation. In rental markets, Ambrose et al. (2023) show that rent growth measured using newly signed leases diverges sharply from estimates based on surveys of existing tenants. In owner-occupied markets, Anenberg and Laufer (2017) demonstrate that repeat-sales indexes constructed from contract prices can produce different appreciation paths than those based on closing prices. Most directly related, Contat and Larson (2024) show that alternative aggregation of tract-level repeat-sales indexes yields materially different citywide appreciation measures when submarket growth rates differ. These findings underscore that aggregation is not a secondary technical detail but central to accurate measurement.

This paper makes three primary contributions. First, it derives sufficient-statistic conditions for metropolitan house price indexes from the standard monocentric city model of Alonso (1964), Mills (1967), and Muth (1969). Second, it shows that both spatial aggregation and the definition of the housing good (units versus space), together with tenure heterogeneity, affect the construction of theoretically consistent price indexes. Third, it presents empirical examples illustrating that aggregation consistent with theory yields materially different measures of appreciation across metropolitan areas. Even modest differences in aggregation and quantity definition can meaningfully alter measured appreciation for individual cities.

The remainder of the paper is organized as follows. The next section formalizes the sufficient statistic requirements for a Laspeyres index. Section 3 reviews the strengths and limitations of hedonic and repeat sales indexes. Section 4 shows that, under standard urban equilibrium conditions, price changes must be modeled as a function of distance from the city center and weighted by the housing stock at each distance band; deviations implicitly assume zero income elasticity of housing demand. Sections 5 and 6 present empirical comparisons that aggregate subindexes using theoretically justified weights, drawing on data from the American Housing Survey and the Federal Housing Finance Agency. These comparisons reveal substantial differences in appreciation by tenure and centrality. Section 7 concludes with implications for the empirical measurement of house price change.

2 Sufficient statistics for a Laspeyres index

Housing price measures can be based on either rental or asset prices and the basis of quantity measurement can be the housing unit or housing services provided, i.e., a measure such as interior space which may be adjusted for quality. These details of measurement are consequential but not the focus of this paper. Regardless of the price or quantity measure chosen, the index number problems we identify are mathematically identical.

For ease of exposition and for formalizing our propositions in Section 4, we reduce the dimensionality that characterizes locations in a city. Rather than index each individual neighborhood or track myriad characteristics, we operate under the simplifying assumption that housing at each radial distance k from the central business district (CBD) is identical. This allows for differentiation/differences and integration/summation over distance rather than across individual neighborhoods.

Laspeyres (1871) would answer the index number problem for measuring the average housing price and its rate of change by recognizing that price and quantity varies with location and constructing a price index treating housing in each location as a distinct good. Suppose that r_t^* is an index number indicating the average price per unit of housing at time

t, and that housing at a distance k from the center city is homogeneous such that

$$r_t^* = \sum_k {r_t(k)h_t(k)} \big/ {H_t(k)}, \tag{1}$$

where $r_t(k)$ and $h_t(k)$ are, respectively, the price and quantity of housing at time t and distance k from the city center, and $H_t(k) = \sum_k h_t(k)$. The Laspeyres measure of the rate of change in housing price between $t-1$ and t is then

$$\frac{\Delta r_t^*}{r_{t-1}} = \frac{\sum_k r_t(k)h_{t-1}(k) \big/ H_{t-1}(k)}{\sum_k r_{t-1}(k)h_{t-1}(k) \big/ H_{t-1}(k)}. \tag{2}$$

The sufficient statistic implications of Equations (1) and (2) are demanding: The marginal distributions of both housing prices and quantities over various distances, k, must be known.

There are two assumptions that may simplify the computation of (2). First, if the percentage change in the price of housing is constant, in expectation, across all k, then $r_t(k)/r_{t-1}(k) = \theta$ and it follows from (2) that a sufficient statistic for measuring the rate of housing price increase is a measure of θ which may occur at any k (Contat and Larson, 2024). For example, in small cities with relatively flat appreciation gradients in many periods, this assumption may commonly hold. The spatial distribution of housing may vary, but it is not consequential to the computation of a sufficient statistic for measurement of the percentage change in housing cost.

A second assumption may also yield the targeted Laspeyres index. If the weights used in the index construction are proportional to $h_{t-1}(k)/H_{t-1}(k)$, then the resulting index will be a Laspeyres index. For example, this circumstance can occur if an index is calculated using observed transactions, and transaction shares are proportional to housing stock shares.

The discussion thus far has assumed agreement on the fundamental nature of the housing good being priced. While it is common to construct indexes based on the price of housing units, this approach treat a highly heterogenous good as uniform. In particular, the quantity of space in housing units varies substantially over time and across location. If the objective is to measure the price of housing services or housing space, then the definitions of h and H in the preceding equations should reflect a measure of housing services, rather than assuming that each housing unit provides the same quantity of housing space. The empirical examples in later sections will illustrate the importance of differences in weighting by measures of space and units.

3 Conventional house price indexes

Constructing a price index requires comparing prices for identical assets across time. In real estate markets, this task is complicated by two well-recognized challenges. First, properties are highly heterogeneous, and many price-relevant attributes—especially locational amenities—are imperfectly observed, making it inherently difficult to fully control for cross-sectional variation in asset quality. Second, even when the same property transacts multiple times, its price-determining characteristics may change over time due to depreciation, renovation, or alteration. Isolating true market-wide price movements therefore requires accounting for both unobserved heterogeneity across properties and quality evolution within properties over time.[1]

3.1 Weighted repeat-sales methods

In the United States, residential property price indexes are overwhelmingly based on the repeat-sales method, which addresses unobserved heterogeneity by comparing prices for the same unit across multiple transactions, under the assumption that unit characteristics are

[1] An additional complication arises when transaction prices reflect the option value of land rather than the value of the existing structure. In dense or supply-constrained urban markets, properties are often purchased with the intention of demolition or substantial redevelopment, such that observed prices primarily capitalize land value, zoning constraints, and redevelopment options rather than current housing services—e.g., see Rosenthal and Helsley (1994); Clapp et al. (2012); and Gedal and Ellen (2018). In these cases, price changes may reflect shifts in land or option values rather than market-wide housing appreciation, complicating interpretation of standard house price indexes.

time invariant. A repeat-sales index is constructed by estimating a pooled regression of log price differences for repeat-sale pairs on time-period indicators that take the value +1 in the resale period, −1 in the original sale period, and zero otherwise. Because error variance increases with the interval between transactions, observations with longer holding periods are typically down-weighted to account for heteroskedasticity (Case and Shiller, 1987; Calhoun, 1996). The price index is recovered by exponentiating the estimated time-period coefficients.

Although Bailey et al. (1963)—Bailey, Muth, and Nourse (1963; hereafter BMN)—provide the foundational regression formulation of the repeat-sales estimator, the underlying structure is closely related to the geometric mean index proposed by Jevons (1865). The principal contribution of BMN is to demonstrate that price relatives from arbitrary combinations of periods can be embedded within a regression framework, producing a geometrically aggregated index. The adoption of repeat-sales house price indexes in the US followed the seminal contribution of Case and Shiller (1987), which extended the BMN framework and established the methodology used in practice today.[2]

The primary advantage of the repeat-sales approach is its minimal data requirements: only a property identifier, transaction price, and sale date are needed. To the extent that difficult-to-measure attributes, such as locational amenities, are time invariant, their effects are differenced out by construction. However, repeat-sales indexes face several well-documented limitations. Most fundamentally, observed price changes between transactions may reflect unobserved quality evolution—depreciation, renovation, or alteration—rather than pure market appreciation, a threat to identification.

Additional concerns relate to sample representativeness and temporal stability. Properties that transact repeatedly may differ systematically from the broader housing stock, giving rise to selection bias and limiting external validity. Moreover, because properties that sell

[2] The realization that standard repeat-sales price indices are essentially Jevons indices invites analysis of the representativeness of the implicit transaction (and holding period) weighting in a house price index for a large area (e.g., city or state). Under certain proportionality and/or homogeneity conditions, these regression-based Jevons-like formulations may produce unbiased estimates of geometric Laspeyres or Tornqvist indices.

only once do not contribute to estimation, a large share of transactions is excluded, reducing statistical efficiency, especially in thin markets. Finally, because repeat-sale pairs are only observed upon resale, new transactions reveal information about past market conditions, resulting in continual historical revision of the index.[3]

A further issue concerns the index-number characteristics of the repeat-sales estimator. The standard repeat-sales framework constitutes an elementary price index, meaning it aggregates price relatives across transactions without explicit expenditure or quantity weights. Because each transaction pair enters the regression as a single observation, the estimator implicitly assigns equal weight to price relatives, apart from interval-based adjustments. As Diewert (2012) notes, such elementary indices can be inconsistent with economic index theory unless strong proportionality or homogeneity conditions hold; Keynes (1930, p. 57) similarly criticized unweighted geometric indices for their lack of proportional weighting. In the housing context, this implies that standard repeat-sales indexes are transaction-weighted rather than weighted by the distribution of housing units or housing services, raising concerns about representativeness when constructing metropolitan or national measures of appreciation.

The two primary US indexes—the S&P Cotality (formerly CoreLogic) Case-Shiller Home Price Indices and the Federal Housing Finance Agency (FHFA) House Price Index (HPI)—are each released as suites of indexes spanning multiple levels of geographic aggregation. While both indexes rely on repeat-sales methods to control for unobserved, time-invariant heterogeneity, they differ along several dimensions, including data coverage, the treatment of heteroskedasticity, and weighting schemes used to aggregate indexes.

The FHFA HPI is constructed from repeat transactions of single-family homes financed with conforming mortgages that are purchased or securitized by Fannie Mae and Freddie Mac. Because the index relies exclusively on conforming mortgage transactions acquired

[3] A substantial literature developed following Case and Shiller (1987), that addresses sample selection bias and issues related to transaction interval. This literature is reviewed in Nagaraja et al. (2014). Practitioner-oriented discussions of sample size and coverage issues appear in Eurostat (2013). Additional research examining index revision and temporal instability includes Clapp and Giaccotto (1999), Clapham et al. (2006), and Deng and Quigley (2008).

by the GSEs, it excludes nonconforming (e.g., Alt-A, jumbo, and subprime) loans and cash purchases and therefore does not represent the full universe of housing transactions. Price changes are estimated using generalized least squares, with observations spanning longer intervals between transactions down-weighted to account for increasing error variance over time, resulting in a geometrically weighted index. The national index is formed by combining separate indexes for the nine Census divisions using fixed weights based on the distribution of the housing stock. See Calhoun (1996) for details.

The S&P Cotality Case-Shiller indices are constructed from a wider set of arms-length housing transactions obtained from public deed records and are estimated using an arithmetic repeat-sales framework. Individual sales pairs receive weights proportional to the initial transaction value, with additional adjustments that account for greater dispersion in price changes over longer holding periods.[4] The national index is formed by aggregating the nine Census division indexes using estimates of the aggregate value of the single-family housing stock, yielding a transaction-value-weighted measure of house price dynamics. See S&P Dow Jones (2025) for details.

Taken together, these features underscore that standard repeat-sales indexes rely on observed transaction pairs and implicitly treat those transactions as representative of the broader housing stock. This assumption is particularly restrictive within metropolitan areas, where both transaction intensity and appreciation vary systematically across space.

Existing repeat-sales indexes implicitly treat observed transactions as representative of the metropolitan housing stock, despite systematic intra-city variation in transaction intensity and appreciation rates. To address this limitation, metro-level indexes can be constructed by first estimating repeat-sales subindexes for neighborhoods defined by distance from the city center and then aggregating these subindexes using weights proportional to

[4] If expected appreciation were homogeneous across properties within a city, such weighting could improve efficiency. However, if appreciation varies systematically within metropolitan areas, interval-based weighting may amplify disproportionate representation of high-turnover submarkets, increasing ex ante bias. Case and Shiller (1987), Sagi (2021), and others document that real estate returns are correlated with holding periods.

the housing stock at each distance. This approach aligns index construction with urban economic theory and provides a principled way to account for spatial heterogeneity in appreciation when measuring average city-level price change.

While repeat-sales indexes can, in principle, be estimated within geographic strata as recommended here, doing so raises feasibility and selection-bias concerns. Repeat-sales methods rely on a relatively small and non-random subset of properties that transact more than once, and further spatial disaggregation can quickly erode sample sizes and exacerbate resale selection bias, implying that aggregation of such submarket indexes to the metropolitan level would likely require estimation at a lower temporal frequency.[5]

3.2 Hedonic regression methods

With the exception of the United States, most OECD countries have demonstrated a preference for hedonic price indexes over repeat sales methods. Hedonic price indexes use regression techniques to estimate the implicit prices of the housing characteristics that contribute to a property's value. Relative to repeat-sales methods, this approach is more data-efficient because it considers both single- and repeat-sale properties. Provided that sufficiently detailed data on property attributes are available, hedonic models can account for changes in quality at the property level, addressing the fundamental limitation inherent in repeat-sales approaches.

The primary limitations of hedonic price indexes stem from their data and modeling requirements. Effective implementation requires comprehensive and consistently measured information on property characteristics, and empirical results may be sensitive to specification choices, including functional form and the selection of covariates.[6] This computational complexity tends to reduce transparency relative to simpler index construction methods, posing

[5] Although their analysis focuses on producing indexes for specific market segments rather than aggregating subindexes, Bogin et al. (2019a) provide a useful discussion of the tradeoff between temporal frequency and geographic granularity in repeat-sales estimation.

[6] Sirmans et al. (2005) and Owusu-Ansah (2011) provide helpful reviews and discussions on explanatory variables and functional forms.

challenges for routine production by statistical agencies and comparisons across countries.

The *Handbook on Residential Property Price Indices (RPPIs)* (Eurostat, 2013) reports that the hedonic literature distinguishes three approaches to constructing constant-quality property price indexes.[7] The time-dummy method estimates a pooled regression with period indicators that represent quality-adjusted price changes. The characteristic-pricing approach instead revalues a representative bundle of housing attributes using hedonic coefficients from different periods, producing Laspeyres-, Paasche-, or Fisher-type indexes depending on whether base-period characteristics, current-period characteristics, or a geometric mean of the two is held fixed. Lastly, the imputation approach operates at the level of individual properties rather than an average bundle, using hedonic coefficients to predict counterfactual prices across periods: imputing base-period prices for current-period properties corresponds to a Laspeyres-type index, current-period prices for base-period properties yields a Paasche-type index, and a symmetric combination produces a Fisher-type index.[8]

The RPPI Handbook contemplates stratification (also called mix adjustment) to reduce sample selection bias arising from compositional change in observed housing transactions.[9] Under this approach, properties are partitioned into relatively homogeneous strata—such as by location, dwelling type, or size—price change is measured within each stratum, and the resulting sub-indexes are aggregated using explicit weights. The arguments advanced here

[7] The *RPPI Handbook* was developed under the coordination of the Statistical Office of the European Union (Eurostat) to establish international standards and published jointly by Eurostat, the International Labor Organization (ILO), the International Monetary Fund (IMF), the Organization for Economic Co-operation and Development (OECD), the United Nations Economic Commission for Europe (UNECE), and the Inter-Secretariat Working Group on Price Statistics (IWGPS) at the World Bank.

[8] The time-dummy hedonic approach, while simple and widely used, relies on restrictive assumptions—most notably time-invariant characteristic prices—and offers limited scope for transparent and economically meaningful weighting. By contrast, Silver (2018) shows that, under consistent choices of functional form and aggregation, the characteristics/repricing and imputation approaches can yield numerically equivalent Laspeyres-, Paasche-, and Fisher-type indexes. On this basis, Silver argues that the practical choice between hedonic approaches is less about correctness per se and more about transparency, weighting, and how individual transactions enter the index, ultimately favoring weighted imputation, or the equivalent characteristics-based formulation, over time-dummy methods.

[9] For example, the text is direct about misspecification from pooling heterogeneous markets, e.g., "When using hedonic regression techniques to adjust for quality (mix) changes, stratification is highly recommended. It is very unlikely that a single hedonic model holds true for all market segments, hence separate regressions should be run for different types of properties, different locations, etc." (p. 55).

are theory-based rather than data-driven and suggest stratification based on location within the city. The reason for this conclusion is that rates of change in hedonic prices should not be uniform across locations. This same argument suggests that using time dummies from pooled hedonic equations will not produce equivalent results.

4 Urban spatial theory and sufficient housing price statistics

Section 2 suggested that if the percentage change in housing price is not a function of location, then measurement of housing price change does not require a spatially random measure of housing prices. In this case, it is not necessary to account for spatial differences in housing density or differences in location where housing price change is measured. Accordingly, the measurement of housing prices can be geographically concentrated. This section considers the relation between this assumption of constant rates of price change across space and standard urban economic theory. Specifically, under what circumstances is it reasonable to expect that the rate of change in house prices in cities is not a function of location?

In a neoclassical city, such as the classic monocentric model reviewed by Brueckner (1987), households achieve an iso-utility equilibrium in which utility of a composite commodity, c, whose price is normalized to unity and constant throughout the city, and housing, h, with price, r. Each of these values vary by location which is expressed as the distance from the CBD, k. Households must either commute to a city center where they earn income, y, or commute shorter distances where they face an urban wage gradient which is reduced by the amount of commuting cost saved. Therefore, the households' problem is to

$$\max U\left(c,h\right) \text{ subject to } y-\tau k=c+r(k)h, \tag{3}$$

where τ is commuting cost per unit distance.

The iso-utility of households implies that $\mathrm{d}r/\mathrm{d}k$ is determined by Muth's (1969) equation requiring that $\mathrm{d}r/\mathrm{d}k = -(\mathrm{d}\tau/\mathrm{d}k)/h$. The only general constraint on $\mathrm{d}\tau/\mathrm{d}k$ is the logical requirement that it is greater than zero, however the further simplification that it is constant will be maintained here.[10] Housing consumption, h, is determined by the iso-utility condition and earnings.

The essential question of determining a sufficient statistic for measuring house price change reduces to a simple question: What constraint does Muth's equation and the requirement that $(\mathrm{d}r/\mathrm{d}k)/r = \text{constant} = \theta$ place on the housing demand function?

Dividing through Muth's equation by r yields

$$(\mathrm{d}r/\mathrm{d}k)/r = -(\mathrm{d}\tau/\mathrm{d}k)/rh. \tag{4}$$

But this implies that for both Muth's equation and $(\mathrm{d}r/\mathrm{d}k)/r = \text{constant} = \theta$ to hold, then

$$(\mathrm{d}r/\mathrm{d}k)/r = -(\mathrm{d}\tau/\mathrm{d}k)/rh = \theta. \tag{5}$$

Thus, for the percentage change in the rental price of housing to be constant across all locations $k \leq k^*$ and be consistent with Muth's equation, the ratio of the marginal rise in transportation cost with distance to total housing expenditure must be constant.

Proposition 1. *If commuting cost per mile is constant and the iso-utility condition of households holds, constancy of the elasticity of housing price change with distance from the city center implies that the household utility function must be quasi-linear.*

Clearly if $\mathrm{d}\tau/\mathrm{d}k$ is constant, then total expenditure on housing must be constant. If the price of the composite commodity does not vary systematically with location, as is commonly

[10] The units of $\mathrm{d}\tau/\mathrm{d}k$ must be chosen so that they are consistent with the measurement of rental price. This issue has been discussed extensively in the literature. The simplification that this measure is not a function of location is adopted here because if that is not true, it would further complicate the measure of housing price differences as discussed below.

assumed, this implies that the utility function of households takes a specific quasi-linear form,

$$U(c,h) = \lambda \ln h + \rho c. \tag{6}$$

Maximization of (6) as constrained in (3) produces a housing demand function in which

$$r(k)h(k) = \lambda/\rho, \tag{7}$$

which satisfies the criterion in (5) that both the marginal increase in commuting cost with distance and total housing expenditure do not vary with location. Thus, Proposition (1) is proved. The consequences of more complex commuting cost functions for the possibility that housing price is iso-elastic require simulation of a numerical simulation model. Larson et al. (2022) demonstrate endogenous congestion makes the iso-elastic assumption even more problematic than the case where transportation cost is exogenous and constant.

The difficulty with this result is that there is no evidence that preferences for housing are quasi-linear. There is debate in the empirical literature regarding where the income elasticity of demand for a primary residence lies on the interval from some positive value to unity, but agreement that the zero income elasticity of quasi-linear utility is inconsistent with all evidence.

Proposition 2. *The assumption of iso-utility that is not quasi-linear implies that a sufficient statistic for constructing a Laspeyres index of price change in a city requires that rates of price change be measured as a function of distance from the city center and weighted by the fraction of housing at each distance interval.*

This proposition follows directly from Proposition (1), the empirical evidence on demand for a primary residence, and the definition of a Laspeyres index from Equation (1) and so no proof is given.[11]

[11] Related theoretical work reinforces these implications. Broxterman et al. (2025) show that housing supply elasticity varies systematically with intra-urban location, linking observed price responses to underlying

5 An illustration of bias from ignoring spatial heterogeneity

As a preliminary empirical example, we use the American Housing Survey Metropolitan Sample (AHS-MS) to illustrate the potential effects of failing to consider spatial variation in constructing an index of city housing price changes. Both rental and asset prices are examined. Rental prices are important for measuring cost of living and value of current use, whereas asset prices also incorporate expectations and option value. The AHS-MS is well suited for this purpose for several reasons. First, it surveys households in both renter- and owner-units and classifies housing by central-city and suburban location. This allows comparison of price indexes by tenure status and centrality where differences are likely substantial. In addition, the AHS-MS is a panel dataset, so that price changes can be calculated based on the same units in the same locations over time. As a result, changes in median values and rents can be interpreted as Laspeyres-like to the extent that the underlying set of sampled housing units is held constant over time.

The central-city versus suburban location indicator within metropolitan areas is not available in the public-use microdata files of the AHS-MS. Accordingly, we rely on published summary tabulations, in which median values and rents are reported in interval form. To compute compound annual growth rates, we assign each reported interval its midpoint. This approach introduces approximation error, and the resulting growth rates will not exactly match those constructed from the underlying micro-data. For the present illustrative purpose, however, the midpoint approximation provides a consistent basis for comparing growth across central-city and suburban locations.

Using AHS-MS data, we first measure changes in city house prices by calculating the annual growth rate in the median estimate of value and rent (both reported in ranges) based on all sample units in the metro area regardless of location. The results of this exercise

spatial structure.

are displayed in the first two columns of Table 1 for 25 US cities for the years 2015 to 2019. This period was characterized by rapid growth in housing prices. The median value of owner-occupied homes grew at a compound annual rate ranging from 1.4% to 13.6%, with an average of 7.1%. For median rent, the range was 1.6% to 7.9%, with an average of 4.8%.

Like most large-scale surveys, AHS-MS is a stratified sample. The nine categories employed in the sampling strategy pertain mostly to tenure and structure type, but not to location within the metro. As a result, there is no reason to believe that the stratified samples are spatially random, because the sample rate for each stratum is constant while the fraction of units varies by location. In other words, the Census Bureau does not cross-stratify by location, e.g., (renter vs. owner) $\times$ (central city vs. suburbs). As explained in Section 2, variation in the spatial distribution of housing is only consequential if the percentage change in the price of housing is not constant across locations. Accordingly, we next examine growth rates of asset and rental prices in the AHS-MS separately for units located in the central city and suburbs.

Based on the literature cited in Section 1, 2015 to 2019 was a period during which centrality had become a more sought-after amenity. Therefore, our first hypothesis is that the growth rates of rental and asset prices displayed in Table 1 will differ when location is divided into central city and suburb. As expected, columns three through six show substantial differences in rates of change between these two broad areas of cities. While the suburbs did outperform in some metros, growth rates of value and rent in the central city were 137 and 15 basis points (BPS) higher on average, respectively, consistent with the pre-COVID urban resurgence narrative.

The estimates in Table 1 demonstrate that rates of appreciation are not spatially invariant within cities. Taken to the logical extreme, if either central city or suburban samples were used alone to estimate the average rate of change for the entire city, the bias would be substantial. Of course, to the extent that sampling in the AHS-MS approaches the spatially randomized ideal, a weighted average of spatially varying changes should approximate the mean change rate for the entire city.

We next consider two alternative ways of characterizing housing price change at the metro level using annual growth rates in median value and rent from the AHS-MS. In the first construction, we simply average the rates for the central city and suburbs. These naïve indexes, which appear in the first two columns of Table 2, indicate that the median value of owner-occupied homes grew at an average of 7.8% and rents by 5.4% in the 25 cities covered. These results can be compared with the alternative index in columns three and four that weights the growth rates by all units, renter and owner, in each location as suggested by Proposition 2, which implies that a sufficient statistic for a Laspeyres index requires weighting location-specific price changes by the fraction of housing at each distance interval.

The alternative index provides a useful reference point because it is consistent with well-known price index formulations. Differences between the alternative and simple average indexes appear in columns five and six of Table 2. On average, growth rates based on the alternative (sufficient statistic) index are 43 and 4 BPS lower, respectively, than their simple average analogs. (Relative to the metrowide growth rates in Table 1, values for the alternative indexes are 29 and 53 BPS higher.) For particular cities, the differences are much more substantial. For example, house price growth is 495 BPS lower in Atlanta and rent growth is 205 BPS lower in Pittsburgh relative to the simple average of the central city and suburbs rates for those metros.

The median-based growth rates reported in Tables 1 and 2 are not directly comparable to stratified or mix-adjustment methods described in Eurostat (2013). Instead, they should be interpreted as within-sample comparisons constructed using uniform definitions and procedures. This exercise illustrates that when appreciation rates vary systematically across locations, failure to weight location-specific growth rates by the spatial distribution of housing—as required by by Proposition (2) for a Laspeyres-consistent index—can generate biased measures of citywide housing price growth. The differences shown in Table 2 underscore that the bias is not merely theoretical but quantitatively meaningful. More broadly, the results demonstrate that spatial heterogeneity is not a secondary refinement to index construction but a necessary component of any measure intended to sufficinetly characterize aggregate price change within a city.

6 The sufficiency of the weighted repeat-sales index

The primary empirical illustration in this paper relies on the most widely used method for measuring city-level housing asset price changes in the US: a weighted repeat-sales index estimated from transaction prices. Baum-Snow and Han (2024) document that repeat sales-based appreciation rates vary substantially across census tracts within the same city. They further demonstrate that estimates of housing price change and supply elasticity differ depending on whether they are based on housing units or interior space—findings consistent with the propositions advanced in this paper. Standard city-level index construction (e.g. FHFA or Case-Shiller) using repeat-sales methods aggregates all available transaction pairs for the same housing unit and weights each pair's price change by the time between transactions. This method, though ubiquitous, is *ex ante* unlikely to be proportional to any relevant housing stock quantity measure, and thus to yield a Laspeyres index. Because Proposition 2 requires weights proportional to the spatial distribution of housing quantities, standard repeat-sales aggregation need not satisfy the sufficiency condition implied by spatial equilibrium.

To examine the implications of alternative aggregation schemes more directly, we also draw on tract-level estimates of housing price appreciation from Contat and Larson (2024), which provide annual price changes for 63,122 census tracts across 581 core-based statistical areas from 1989 to 2021. These measures are designed to provide prices for exhaustive, mutually-exclusive submarkets within each city, making them well suited for aggregation based on spatial characteristics such as distance from the city center. Because these measures pertain exclusively to housing asset prices and do not include rental price changes, our empirical illustration focuses on how asset price appreciation responds to alternative aggregation schemes. This framework allows us to test whether different aggregation rules satisfy the sufficiency condition in Proposition 2.

In the neoclassical radial monocentric city model considered in Section 4, commuting distance to the city center is sufficient to characterize differences across housing submarkets. In real-world cities, however, substantial heterogeneity exists in neighborhood attributes even within narrow distance bands. Accordingly, we allow for a more general disaggregation at the tract level. The implicit assumption is that housing and households are homogeneous within tracts.

Dimensionality becomes an additional concern when attempting to create a Laspeyres index, in practice. In the neoclassical city, the quantity of housing is given by a single attribute, h. In real-world applications, however, housing is typically observed in one of three ways: (1) In Leontief terms, with housing consumption represented simply as a unit; (2) along the intensive margin of the housing structure, as square feet; or (3) in economic terms, as the quantity of housing services generated by use of the unit, including property characteristics—such as interior space, lot size, and structure age—as well as location-based amenities, including school quality and proximity to other urban services. Approaches (1) and (2) can be approximated using Census data on the number of units and size of units (via number of rooms). Approach (3), while seemingly more complex due to the high dimensionality of characteristics required for hedonic estimation, can be implemented using an alternative formulation of the Laspeyres index. The choice of quantity measures determines whether aggregation weights correspond to the housing stock concept required for sufficiency.

Balk (2012) shows that a Laspeyres (or Lowe) index can be constructed in a manner equivalent to the base-year quantity formulation using base-year value shares,

$$\frac{\Delta r_t^*}{r_{t-1}} = \sum_k \frac{r_t(k)}{r_{t-1}(k)} v_{t-1}(k), \tag{8}$$

where $v_{t-1}(k) = r_{t-1}(k)h_{t-1}(k)/\sum_{\kappa} r_{t-1}(\kappa)h_{t-1}(\kappa)$. This formulation is advantageous in the study of the housing price because of the multidimensional nature of housing services, which cannot be fully captured even in hedonic models because of unobservables. The primary drawback, common to all uses of housing asset prices in place of flow housing prices, is that asset prices may be influenced by factors unrelated to contemporaneous rents. Crucially, sufficiency depends on base-period quantity weights—not transaction frequencies.

Census data provide information on the number of rental and owner-occupied units in each tract, as well as average unit size (in rooms) and reported value. Consequently, results for asset price change weighted by total number of units versus those weighted by only owner or only renter aggregations can be evaluated and compared with Laspeyres-type aggregations weighted by measures of units, unit size, or value.

6.1 Submarket aggregation

The general form of the city housing asset price computation is

$$\%\Delta P_{jt} = \sum_{i} W_{ijt} \%\Delta P_{it} \quad , \tag{9}$$

where i is an index of Census tract, $\%\Delta P_{it}$ is the weighted repeat sales estimate of the change in asset price for tract i in period t, and W_{ijt} is type j tract-based weight used to aggregate repeat sales estimates across the city. Formally, $W_{ijt} = J_{it-1}/\sum_i J_{it-1}$ where J is a measure of the tract weight based on one of three factors, which depend on the target index characterization: total housing units, total number of rooms, or total asset value in tract i. The baseline for comparison is an index that treats all tracts equally, with $W_{ijt} = 1/I$, where I is the number of tracts in the city.[12] Under Proposition 2, only those weighting schemes in which J_{it-1} corresponds to the spatial distribution of housing quantities are consistent

[12] Weights constructed using the tract's share of transactions within the city would be the ideal comparison to reflect a standard, city-level index computation, but transactions by tract are unavailable for the present analysis. Also note that we do not consider the effects of weighting transactions based on holding period as is standard (see, e.g., Case and Shiller, 1987; Calhoun, 1996).

with the sufficient-statistic result. Equal weighting generally violates this condition unless appreciation is spatially homogeneous.

6.2 Illustration of index differences

The effects of location and alternative measures of housing services are illustrated using examples from Boston and Houston. These cities represent opposite ends of the spectrum among large metropolitan areas in terms of rental share and land-use regulation. In Boston, nearly 65% of the housing stock is renter-occupied, compared to about 40% in Houston. Boston also enforces more restrictive planning policies, while Houston is known for its minimal regulatory environment. Given these contrasts, we expect not only differences in appreciation rates by location and housing service measure within each city, but also distinct spatial and temporal patterns between them.

Figure 1 presents tract-level estimates of price appreciation for the full 1990-2020 period, as well as for two subperiods corresponding to the housing boom (2001-2006) and bust (2006-2011). The maps display census tracts whose centroids fall within a 15-mile radius of each city's City Hall. As expected, both cities exhibit substantial within-city variation in appreciation rates across census tracts across each time interval.

The temporal dynamics, however, differ markedly across the two cities. In Boston, the pattern is consistent with mean reversion: tracts that experienced the highest appreciation during the boom tended to see the steepest declines during the bust. This suggests that price increases in those areas outpaced fundamentals and were subsequently corrected. In contrast, Houston shows considerable persistence: tracts that appreciated more during the boom continued to outperform during the bust. This pattern is consistent with relatively elastic housing supply and more stable, fundamentals-driven price dynamics across neighborhoods. In both cases, the presence of systematic spatial gradients implies that aggregation weights correlated with tract characteristics will affect measured citywide appreciation, as predicted by the sufficiency condition.

Figure 2 displays the share weights used to aggregate census tracts into city-level indexes for the two sample cities. These weights—based on the number of housing units, number of rooms, or total housing value—are derived from various Decennial Census American Community Survey datasets.[13] For both Boston and Houston, the spatial distributions of unit and room shares, shown in panels (a) and (b), are broadly similar, suggesting that appreciation rates weighted by these measures are likely to yield comparable results. In contrast, the value-based weights in panel (c) exhibit a markedly different spatial pattern, indicating that indexes based on value weighting may diverge substantially from those based on physical housing characteristics.

Figure 3 presents estimates of annual and cumulative house price changes under the three alternative weighting schemes. An equally weighted index—constructed by averaging appreciation rates across tracts without regard to housing characteristics or value—is included for comparison.

Panels (a) and (c) show that annual appreciation rates are similar across weighting methods. However, this similarity does not extend to cumulative changes shown in panels (b) and (d): even small annual differences compound over time when appreciation patterns are persistent. The cumulative rate of appreciation varies materially with the choice of weighting scheme. Relative to equal weighting or weighting by number of units, room-weighted indexes tend to show lower cumulative appreciation, suggesting that larger (typically newer) units appreciated more slowly over the sample period. Value-weighted indexes exhibit greater volatility and diverge more sharply across cities, reflecting spatial differences in price levels. These differences arise precisely because the alternative weights place different mass on submarkets with systematically different appreciation rates, violating the aggregation restrictions required for sufficiency.

[13] 1990, 2000, and 2020 tract definitions are converted to 2010 definitions using crosswalk files available from the National Historical Geographic Information System at `https://www.nhgis.org/geographic-crosswalks`. For inter-census years, straight-line imputation is used for imputation.

6.3 Index differences by city type

Differences in index values across cities are highly associated with city attributes. Table 3 shows estimates of six models related to three different Laspeyres formulations, each with two different samples. The alternative formulations include city-level aggregations of tract-level indices using housing units, rooms, and value, and these are compared to an index created using uniform tract weights.

Across all three Laspeyres formulations, covariates indicative of large, growing cities are most highly associated with larger negative index gaps. Why is city size so impactful? Recall that index differences arise when (1) submarket appreciation rates differ and (2) aggregation weights are correlated with those differences (Malone and Redfearn, 2022; Contat and Larson, 2024). These are the two necessary conditions embedded in Proposition 2.

Cities with relatively elastic and homogeneous housing stocks and uncongested commutes tend to have small index value differences (Bogin et al., 2019a). This tends to occur in small and medium-sized cities. By contrast, large and supply-inelastic cities exhibited negative appreciation gradients, with center-city housing appreciating faster than suburban housing. In such cities, weighting schemes that do not reflect the spatial distribution of housing quantities systematically mis-measure aggregate appreciation.

In the case where appreciation gradients do exist, they indicate a correlation between high-valued tracts and appreciation, and higher average city value. Because new housing is more prevalent in suburban locations where appreciation is lower, unit-weighted indices produce less city-level appreciation, echoing results in the previous section using AHS data. New homes also tend to be larger, so when considering the intensive margin of housing, index differences become even more apparent. Overall, these attributes combine to give negative gaps between a uniform-weighted aggregation and each of the three Laspeyres formulations.[14]

These results underscore the difficulty of constructing long-run repeat-sales indexes in cities with substantial spatial heterogeneity. Differences in appreciation rates across locations

[14] For regressions of individual covariates on index gaps, see Tables A.1, A.2, and A.3.

matter for aggregate measurement, and the choice of weighting scheme is consequential.

Consistent with Proposition **??**, a city-level price index constitutes a sufficient statistic only when appreciation is aggregated using weights proportional to the spatial distribution of housing quantities. In the presence of persistent spatial gradients, naive aggregation schemes—such as uniform tract weighting or transaction-based weighting—systematically depart from this condition. In the United States over the past three decades, such departures would have understated appreciation in large or supply-inelastic cities.

7 Conclusions

Precise measurement of housing price changes, including both rental and asset prices, is critical for a wide range of research and practical applications. Over time, numerous refinements to commonly used hedonic and repeat sales index methods have been proposed. This research advances that effort by incorporating intra-city location and unit characteristics, which, according to both economic theory and empirical evidence, are systematically related to housing price.

Previous studies have shown that price changes are not uniform within cities. In particular, shifts in commuting cost—especially those associated with changes in gasoline prices or the introduction of new transportation infrastructure—have been shown to generate non-uniform appreciation patterns. This is the first paper to explore how basic urban economic theory can inform the construction of city-level housing price indexes. The results are that there are theoretical reasons, in addition to the empirical evidence, suggesting that rates of appreciation vary systematically with distance from the city center and that weighting observations by housing units produces different results than considering housing services.

The implications for constructing Laspeyres-type indexes are clear. Given that appreciation rates vary across space, researchers must account for the spatial distribution of data used to create a housing price index. Whether that data come from survey responses or transactions, observations should be weighted by the fraction of housing services at each distance from the city center in order for the index to sufficiently characterize the average rate of change for that city.

Conventional repeat sales and hedonic measures of price change are generally not weighted in this fashion. Fortunately, implementing such measures, particularly in repeat sales indexes, is straightforward and may result in meaningful changes to observed growth rates, as we illustrate using data for 25 US cities from the Metropolitan Sample of the American Housing Survey and repeat sales price estimates for Boston and Houston.

References

Ahlfeldt, G. M., Heblich, S., and Seidel, T. (2023). Micro-geographic property price and rent indices. *Regional Science and Urban Economics*, 98:103836.

Alonso, W. (1964). *Location and land use: Toward a general theory of land rent*. Harvard university press.

Ambrose, B. W., Coulson, N. E., and Yoshida, J. (2023). Housing rents and inflation rates. *Journal of Money, Credit and Banking*, 55(4):975–992.

Anenberg, E. and Laufer, S. (2017). A more timely house price index. *Review of Economics and Statistics*, 99(4):722–734.

Bailey, M. J., Muth, R. F., and Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, 58(304):933–942.

Balk, B. M. (2012). *Price and quantity index numbers: models for measuring aggregate change and difference*. Cambridge University Press.

Baum-Snow, N. and Han, L. (2024). The microgeography of housing supply. *Journal of Political Economy*, 132(6):1897–1946.

Bogin, A., Doerner, W., and Larson, W. (2019a). Local house price dynamics: New indices and stylized facts. *Real Estate Economics*, 47(2):365–398.

Bogin, A. N., Doerner, W. M., and Larson, W. D. (2019b). Local house price paths: accelerations, declines, and recoveries. *The Journal of Real Estate Finance and Economics*, 58(2):201–222.

Broxterman, D., Liu, Y., and Yezer, A. M. (2025). Why we still don't know much about housing supply elasticity. *Available at SSRN 3380629*.

Brueckner, J. K. (1987). The structure of urban equilibria: A unified treatment of the muth-mills model. In *Handbook of regional and urban economics*, volume 2, pages 821–845. Elsevier.

Calhoun, C. A. (1996). Ofheo house price indexes: Hpi technical description. *Office of Federal Housing Enterprise Oversight*, 20552:1–15.

Case, K. E. and Shiller, R. J. (1987). Prices of single family homes since 1970: New indexes for four cities.

Clapham, E., Englund, P., Quigley, J. M., and Redfearn, C. L. (2006). Revisiting the past and settling the score: index revision for house price derivatives. *Real Estate Economics*, 34(2):275–302.

Clapp, J. M., Bardos, K. S., and Wong, S. K. (2012). Empirical estimation of the option premium for residential redevelopment. *Regional Science and Urban Economics*, 42(1-2):240–256.

Clapp, J. M. and Giaccotto, C. (1999). Revisions in repeat-sales price indexes: here today, gone tomorrow? *Real Estate Economics*, 27(1):79–104.

Contat, J. and Larson, W. D. (2024). A flexible method of housing price index construction using repeat-sales aggregates. *Real Estate Economics*, 52(6):1551–1583.

Deng, Y. and Quigley, J. M. (2008). Index revision, house price risk, and the market for house price derivatives. *The Journal of Real Estate Finance and Economics*, 37(3):191–209.

Diewert, W. E. (2012). Consumer price statistics in the uk. *Manuscript*.

D'Lima, W., Lopez, L. A., and Pradhan, A. (2022). Covid-19 and housing market effects: Evidence from us shutdown orders. *Real Estate Economics*, 50(2):303–339.

Edlund, L., Machado, C., and Sviatschi, M. (2022). Gentrification and the rising returns to skill. *Economica*, 89(354):258–292.

Eurostat (2013). *Handbook on Residential Property Price Indices*. OECD Publishing.

Fischer, M. M., Huber, F., Pfarrhofer, M., and Staufer-Steinnocher, P. (2021). The dynamic impact of monetary policy on regional housing prices in the united states. *Real Estate Economics*, 49(4):1039–1068.

Gedal, M. and Ellen, I. G. (2018). Valuing urban land: Comparing the use of teardown and vacant land sales. *Regional Science and Urban Economics*, 70:190–203.

Glaeser, E. L., Gottlieb, J. D., and Tobio, K. (2012). Housing booms and city centers. *American Economic Review*, 102(3):127–133.

Glaeser, E. L. and Gyourko, J. (2005). Urban decline and durable housing. *Journal of political economy*, 113(2):345–375.

Guren, A. M., McKay, A., Nakamura, E., and Steinsson, J. (2021). Housing wealth effects: The long view. *The Review of Economic Studies*, 88(2):669–707.

Jevons, W. S. (1865). On the variation of prices and the value of the currency since 1782. *Journal of the Statistical Society of London*, 28(2):294–320.

Keynes, J. M. (1930). *The pure theory of money*. Macmillen.

Larson, W., Yezer, A., and Zhao, W. (2022). Urban planning policies and the cost of living in large cities. *Regional Science and Urban Economics*, 96:103802.

Larson, W. D. and Zhao, W. (2020). Oil prices and urban housing demand. *Real Estate Economics*, 48(3):808–849.

Laspeyres, E. (1871). Die berechnung einer mittleren waarenpreissteigerung. *Jahrbücher für Nationalökonomie und Statistik*.

Liu, S. and Su, Y. (2021). The impact of the covid-19 pandemic on the demand for density: Evidence from the us housing market. *Economics letters*, 207:110010.

Liu, Y. (2018). Estimating the elasticity of supply of housing space rather than units. *Regional Science and Urban Economics*, 68:1–10.

Malone, T. and Redfearn, C. L. (2018). Shocks & ossification: The durable hierarchy of neighborhoods in us metropolitan areas from 1970 to 2010. *Regional Science and Urban Economics*, 69:94–121.

Malone, T. and Redfearn, C. L. (2022). To measure globally, aggregate locally: Urban land, submarkets, and biased estimates of the housing stock. *Real Estate Economics*, 50(3):656–671.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 57(2):197–210.

Molloy, R. and Shan, H. (2013). The effect of gasoline prices on household location. *Review of Economics and Statistics*, 95(4):1212–1221.

Muth, R. F. (1969). *Cities and housing: The spatial pattern of urban residential land use.* University of Chicago Press.

Nagaraja, C., Brown, L., and Wachter, S. (2014). Repeat sales house price index methodology. *Journal of Real Estate Literature*, 22(1):23–46.

Owusu-Ansah, A. (2011). A review of hedonic pricing models in housing research. *Journal of International Real Estate and Construction Studies*, 1(1):19.

Rosenthal, S. S. and Helsley, R. W. (1994). Redevelopment and the urban land price gradient. *Journal of Urban Economics*, 35(2):182–200.

Sagi, J. S. (2021). Asset-level risk and return in real estate investments. *The Review of Financial Studies*, 34(8):3647–3694.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

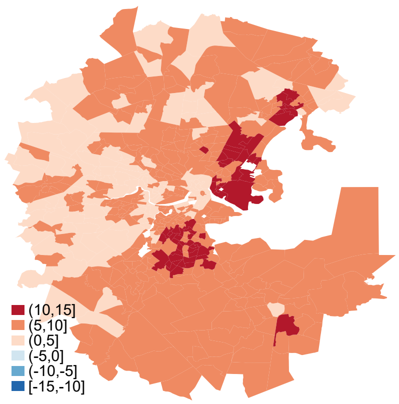
Seagraves, C. and Gatzlaff, D. H. (2025). Examining house price movements within the metropolitan market. *Available at SSRN 5296393*.

Silver, M. S. (2018). How to measure hedonic property price indexes better. *Eurostat Review on National Accounts and Macroeconomic Indicators (EURONA)*, (1):35–66.

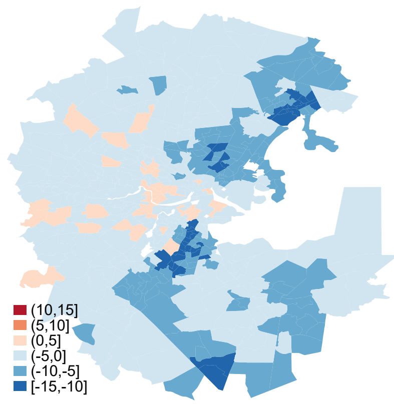
Sirmans, S., Macpherson, D., and Zietz, E. (2005). The composition of hedonic pricing models. *Journal of real estate literature*, 13(1):1–44.

S&P Dow Jones (2025). S&p cotality case-shiller home price indices methodology. Technical report, S&P Dow Jones.

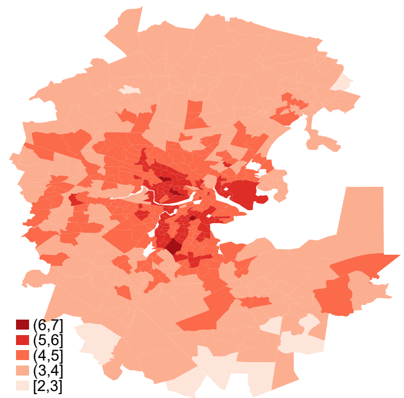
Figure 1: Within-city appreciation differences (annual average)



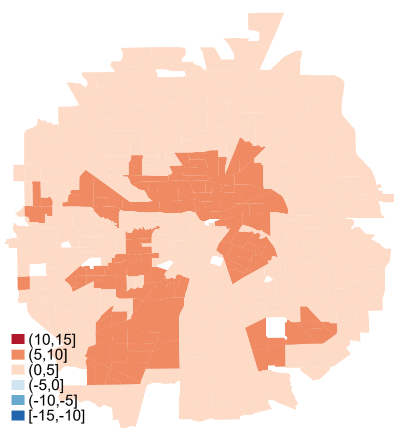
(a) Boston, MA, 2001–2006



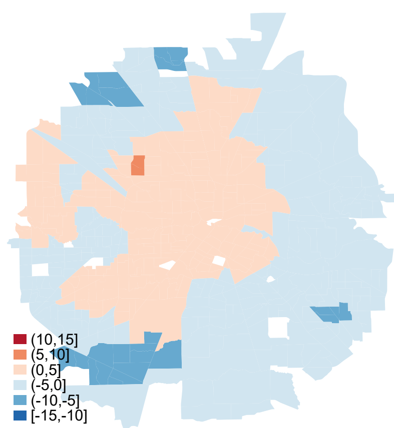
(b) Boston, MA, 2006–2011



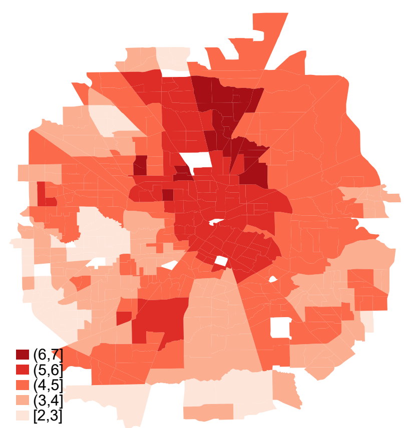
(c) Boston, MA, 1990–2020



(d) Houston, TX, 2001–2006



(e) Houston, TX, 2006–2011

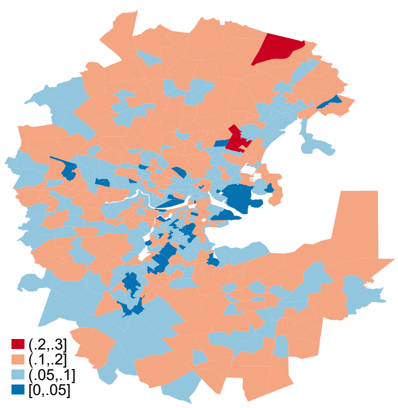


(f) Houston, TX, 1990–2020

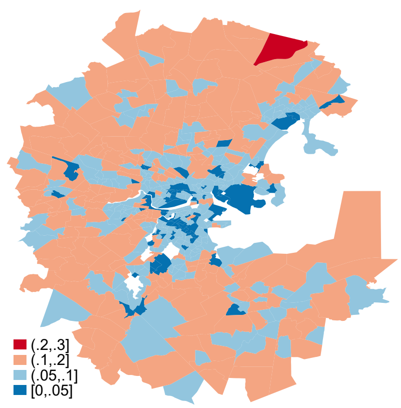
Notes: Maps show the annual geometric average nominal appreciation rate for Census tracts (2010 definitions). The maps display census tracts whose centroids lie within 15 miles of the centroid of the tract containing the respective city's City Hall.
Sources: Contat and Larson (2024); Authors' analysis.

Figure 2: Shares used in city index calculation (2010 values)

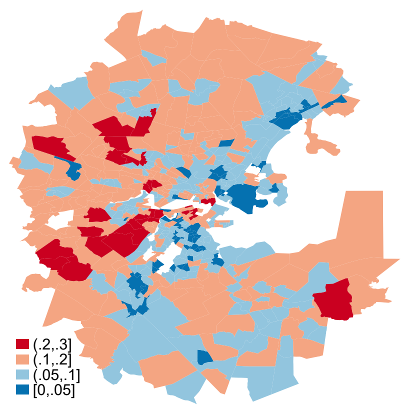
(a) Boston, MA, Housing units



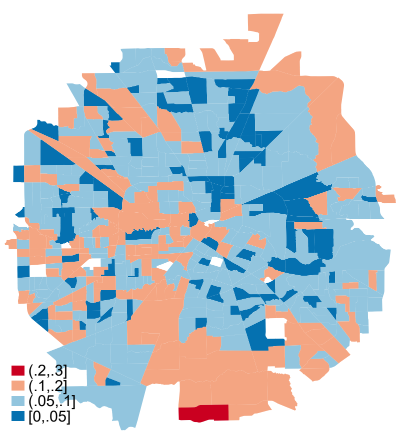
(b) Boston, MA, Rooms



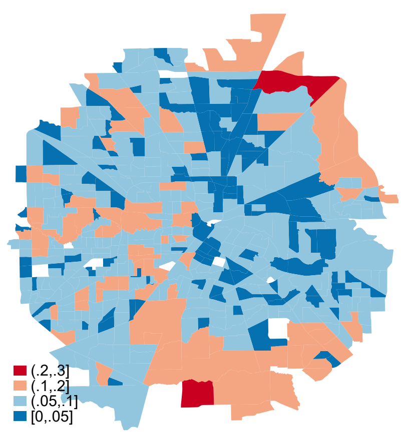
(c) Boston, MA, Value



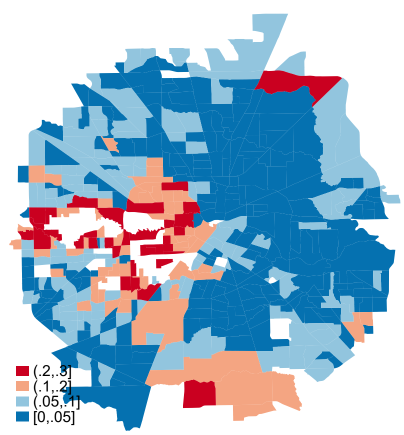
(d) Houston, TX, Housing units



(e) Houston, TX, Rooms



(f) Houston, TX, Value

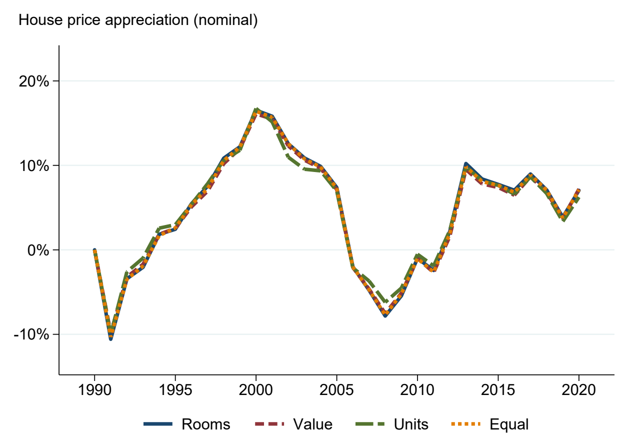


Notes: Maps show the shares used in CBSA-level house price index construction, by Census tract (2010 definitions). Tracts shown have centroids within 15 miles of the centroid of the Census tract containing the respective city's City Hall.

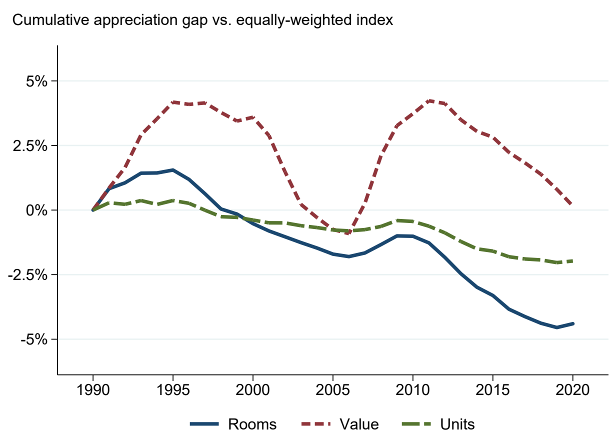
Sources: Census/ACS; Authors' analysis.

Figure 3: Appreciation index differences

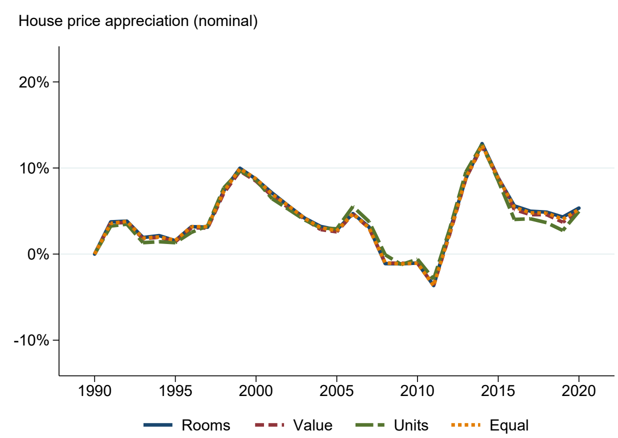
(a) Boston, MA, Annual appreciation



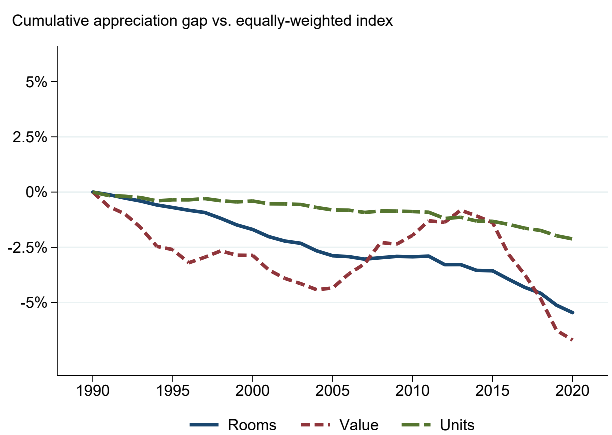
(b) Boston, MA, Cumulative difference



(c) Houston, TX, Annual appreciation



(d) Houston, TX, Cumulative difference



Notes: The figures present alternative city-level house price indices. Each index type (shown in the legend) is constructed by aggregating Census tract (2010 definitions)-level house price appreciation rates using time-varying tract-level weights. The equally-weighted index assigns each tract the same weight.

Sources: Contat and Larson (2024); Census/ACS; Authors' analysis.

Table 1: Compound annual growth rates in median values (2015–2019)

	Metrowide		Central City		Suburbs		Difference	
	Value	Rent	Value	Rent	Value	Rent	Value	Rent
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Atlanta	13.6	5.7	24.5	7.8	11.3	5.7	13.3	2.0
Boston	8.2	6.2	11.7	4.5	8.2	8.8	3.5	-4.3
Chicago	5.7	3.8	9.7	4.1	4.7	3.4	5.1	0.8
Cincinnati	4.7	5.7	7.0	7.5	5.4	8.3	1.6	-0.9
Cleveland	3.2	2.0	0.0	1.8	4.1	6.9	-4.1	-5.1
Dallas	11.0	6.1	10.1	6.1	10.1	6.7	0.0	-0.6
Denver	9.2	7.5	12.5	4.8	10.4	6.2	2.1	-1.4
Detroit	7.5	2.7	10.7	3.5	7.0	2.7	3.6	0.7
Houston	9.3	3.9	9.5	4.3	8.1	5.1	1.4	-0.8
Kansas City	6.2	3.4	5.6	3.4	5.2	2.7	0.3	0.7
Los Angeles	5.5	6.7	6.2	7.4	6.8	6.2	-0.6	1.2
Memphis	9.2	3.9	5.1	3.9	6.8	3.6	-1.7	0.3
Miami	6.9	4.5	8.8	5.6	7.5	4.3	1.3	1.3
Milwaukee	5.0	2.7	5.1	1.9	6.1	5.4	-1.0	-3.6
New Orleans	4.1	1.6	5.6	1.6	5.7	3.2	-0.1	-1.6
New York City	4.3	4.1	4.7	3.8	4.0	4.5	0.6	-0.7
Philadelphia	5.3	4.1	7.5	5.6	4.7	4.2	2.9	1.4
Phoenix	11.1	7.0	11.5	7.5	10.7	4.6	0.8	3.0
Pittsburgh	5.3	2.0	13.6	7.7	5.1	2.1	8.5	5.6
Portland	9.4	7.9	9.8	10.1	9.3	7.5	0.5	2.6
Raleigh	4.7	4.9	1.5	9.3	8.0	6.1	-6.5	3.2
Riverside	6.8	6.8	7.5	8.7	6.6	5.7	0.8	3.0
San Francisco	8.7	6.6	9.3	7.0	9.5	8.5	-0.2	-1.5
Seattle	10.7	7.2	13.6	6.8	10.0	7.4	3.7	-0.6
Washington DC	1.4	2.7	1.0	1.6	2.7	2.7	-1.6	-1.1
Average	7.1	4.8	8.5	5.5	7.1	5.3	1.4	0.1
Deviation	2.8	1.9	4.9	2.4	2.3	1.9	3.8	2.4

Values in percents. Authors' calculations using reported median values from the American Housing Survey Metropolitan Sample. Approximately 2,000 housing units are interviewed for each metro area. Differences in growth rates are calculated as central city minus suburbs.

Table 2: Compound annual growth rates in median values (2015–2019)

	Simple Avg		Alternative		Difference	
	Value	Rent	Value	Rent	Value	Rent
	[1]	[2]	[3]	[4]	[5]	[6]
Atlanta	17.91	6.75	12.96	5.99	-4.95	-0.75
Boston	9.92	6.64	9.03	7.72	-0.89	1.09
Chicago	7.20	3.77	6.65	3.69	-0.55	-0.08
Cincinnati	6.20	7.90	5.69	8.18	-0.51	0.28
Cleveland	2.03	4.32	3.05	5.60	1.02	1.28
Dallas	10.13	6.37	10.13	6.36	0.00	-0.01
Denver	11.41	5.53	11.23	5.65	-0.18	0.12
Detroit	8.85	3.10	8.25	2.98	-0.60	-0.12
Houston	8.83	4.74	8.71	4.82	-0.13	0.07
Kansas City	5.40	3.07	5.37	3.01	-0.03	-0.06
Los Angeles	6.50	6.77	6.46	6.85	-0.04	0.08
Memphis	5.99	3.79	5.97	3.79	-0.02	0.00
Miami	8.12	4.93	7.78	4.59	-0.34	-0.34
Milwaukee	5.58	3.65	5.62	3.80	0.04	0.15
New Orleans	5.67	2.40	5.67	2.41	0.00	0.00
New York City	4.35	4.13	4.32	4.16	-0.03	0.03
Philadelphia	6.09	4.86	5.45	4.54	-0.64	-0.31
Phoenix	11.11	6.05	11.19	6.37	0.08	0.32
Pittsburgh	9.38	4.92	6.29	2.87	-3.09	-2.05
Portland	9.57	8.81	9.54	8.64	-0.03	-0.17
Raleigh	4.73	7.71	4.88	7.64	0.15	-0.07
Riverside	7.05	7.24	6.87	6.61	-0.17	-0.63
San Francisco	9.44	7.78	9.43	7.72	-0.01	-0.06
Seattle	11.79	7.10	11.46	7.16	-0.33	0.06
Washington DC	1.85	2.14	2.24	2.40	0.40	0.25
Average	7.80	5.38	7.37	5.34	-0.43	-0.04
Deviation	3.30	1.80	2.70	1.90	1.15	0.58

Values in percents. The simple average index is the mean of the central city and suburbs growth rates from Table 1. The alternative index weights the growth rates by the total number of units, renter and owner, in each location. Differences between the indexes are calculated as Laspeyres minus simple average.

Table 3: Index differences across cities, vs uniform weights

Dependent variable: $P_{2020}^{Laspeyres}/P_{2020}^{Uniform} - 1$

Model	(1)	(2)	(3)	(4)	(5)	(6)
Laspeyres formulation	Units		Rooms		Value	
Housing units (log, 2019)	-0.689***	-0.342***	-1.189***	-0.693***	-0.946	-0.427*
	(0.167)	(0.079)	(0.222)	(0.113)	(0.516)	(0.201)
Housing supply elasticity	-0.156	-0.080	-0.610*	-0.180*	-1.323**	-0.509**
	(0.251)	(0.055)	(0.285)	(0.070)	(0.466)	(0.155)
Structure age (mean)	-0.010	0.037	0.032	0.064**	0.160	0.133**
	(0.040)	(0.021)	(0.043)	(0.024)	(0.085)	(0.044)
Urban decline	1.712*		2.034		2.341	
	(0.822)		(1.114)		(2.729)	
Housing value (mean, log, 2019)	-0.252	-0.546*	-1.855*	-1.337***	-3.285**	-1.668**
	(0.604)	(0.249)	(0.726)	(0.335)	(1.237)	(0.595)
Household income (mean, log, 2019)	3.544**	1.361	4.548**	1.573	8.131*	3.306*
	(1.321)	(0.785)	(1.469)	(0.888)	(3.108)	(1.520)
Constant	-24.530	-3.290	-8.947	8.152	-31.129	-8.365
	(12.927)	(6.993)	(14.264)	(7.760)	(24.918)	(13.565)
N	72	245	72	245	72	245
R^2	0.208	0.113	0.458	0.297	0.277	0.121

Notes: *, **, and *** indicate $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively. Heteroskedasticity-consistent standard errors in parentheses. The dependent variable is the 30-year accumulated gap in the house price index, calculated as $P_{2020}^{Laspeyres}/P_{2020}^{Uniform} - 1$. Housing units (log), housing value (log), income (log), and structure age are CBSA averages from the 5-year American Community Survey in 2019. Urban decline is the share of housing units below replacement costs in 1990 from Glaeser and Gyourko (2005); supply elasticity is from Saiz (2010).

Sources: Contat and Larson (2024); Glaeser and Gyourko (2005); Saiz (2010); Guren et al. (2021); Baum-Snow and Han (2024); Census, American Community Survey; Authors' analysis.

Table A.1: Index differences across cities, housing unit vs uniform weights, single correlate

	Dependent variable: 30 year cumulative index gap (percent)								
Small sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Hunits	Value	Income	Str. Age	Elasticity	Decline	Guren γ	Units γ	Space γ
Variable in column	-0.458**	-0.699*	-0.275	0.049	0.227	1.565**	-0.776**	1.265	1.404
	(0.145)	(0.347)	(1.211)	(0.033)	(0.165)	(0.559)	(0.281)	(1.183)	(1.185)
Constant	6.033**	8.199	2.467	-1.068*	-0.882*	-0.896***	0.337	-0.691*	-1.077
	(2.082)	(4.307)	(12.617)	(0.484)	(0.376)	(0.242)	(0.332)	(0.275)	(0.546)
N	76	76	76	80	75	80	80	80	80
R^2	0.096	0.040	0.001	0.027	0.023	0.071	0.054	0.013	0.016
Large sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Hunits	Value	Income	Str. Age	Elasticity	Decline	Guren γ	Units γ	Space γ
Variable in column	-0.325***	-0.594**	-0.641	0.059**	0.105*	1.658**	-0.366	0.035	0.080
	(0.068)	(0.200)	(0.590)	(0.019)	(0.047)	(0.585)	(0.206)	(0.615)	(0.618)
Constant	4.134***	7.121**	6.510	-0.863**	-0.358*	-0.989***	0.215	-0.108	-0.142
	(0.887)	(2.427)	(6.089)	(0.266)	(0.165)	(0.264)	(0.179)	(0.205)	(0.355)
N	245	245	245	249	249	75	247	244	244
R^2	0.069	0.035	0.004	0.045	0.012	0.079	0.015	0.000	0.000

Notes: *, **, and *** indicate $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively. Heteroskedasticity-consistent standard errors in parentheses. The dependent variable is the 30-year accumulated gap in the house price index, calculated as $P_{2020}^{hunits}/P_{2020}^{Uniform} - 1$. Housing units (log), housing value (log), income (log), and structure age are CBSA averages from the 5-year American Community Survey in 2019. Urban decline is the share of housing units below replacement costs in 1990 from Glaeser and Gyourko (2005), supply elasticity is from Saiz (2010), Guren γ is the sensitivity parameter in Guren et al. (2021), and the new units and space γs are CBSA mean supply elasticities from Baum-Snow and Han (2024) using the gamma01a formulation.

Sources: Contat and Larson (2024); Glaeser and Gyourko (2005); Saiz (2010); Census, American Community Survey; Authors' analysis.

Table A.2: Index differences across cities, rooms vs uniform weights, single correlate

	Dependent variable: 30 year cumulative index gap (percent)								
Small sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Hunits	Value	Income	Str. Age	Elasticity	Decline	Guren γ	Units γ	Space γ
Variable in column	-1.076***	-2.271***	-3.644*	0.145***	0.362	3.091***	-1.033*	2.459	2.719
	(0.215)	(0.522)	(1.654)	(0.042)	(0.231)	(0.830)	(0.428)	(1.923)	(1.941)
Constant	13.757***	26.577***	36.431*	-3.265***	-2.069***	-2.264***	-0.292	-1.850***	-2.593**
	(3.072)	(6.462)	(17.222)	(0.539)	(0.527)	(0.343)	(0.466)	(0.483)	(0.946)
N	76	76	76	80	75	80	80	80	80
R^2	0.263	0.212	0.063	0.115	0.029	0.137	0.047	0.025	0.029
Large sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Hunits	Value	Income	Str. Age	Elasticity	Decline	Guren γ	Units γ	Space γ
Variable in column	-0.754***	-1.617***	-2.910***	0.114***	0.254***	3.327***	-0.551*	1.222	1.379
	(0.092)	(0.293)	(0.801)	(0.025)	(0.067)	(0.863)	(0.250)	(0.942)	(0.959)
Constant	9.249***	19.092***	29.415***	-2.036***	-1.193***	-2.471***	-0.083	-0.947**	-1.346*
	(1.206)	(3.553)	(8.256)	(0.332)	(0.229)	(0.367)	(0.222)	(0.325)	(0.565)
N	245	245	245	249	249	75	247	244	244
R^2	0.203	0.143	0.049	0.091	0.039	0.161	0.018	0.008	0.011

Notes: *, **, and *** indicate $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively. Heteroskedasticity-consistent standard errors in parentheses. The dependent variable is the 30-year accumulated gap in the house price index, calculated as $P_{2020}^{rooms}/P_{2020}^{Uniform} - 1$. Housing units (log), housing value (log), income (log), and structure age are CBSA averages from the 5-year American Community Survey in 2019. Urban decline is the share of housing units below replacement costs in 1990 from Glaeser and Gyourko (2005), supply elasticity is from Saiz (2010), Guren γ is the sensitivity parameter in Guren et al. (2021), and the new units and space γs are CBSA mean supply elasticities from Baum-Snow and Han (2024) using the gamma01a formulation. Small sample and large sample include all CBSA's with a G&G (2005) decline value and Saiz (2010) elasticity value, respectively.
Sources: Contat and Larson (2024); Glaeser and Gyourko (2005); Saiz (2010); Guren et al. (2021); Baum-Snow and Han (2024); Census, American Community Survey; Authors' analysis.

Table A.3: Index differences across cities, value vs uniform weights, single correlate

Dependent variable: 30 year cumulative index gap (percent)

Small sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Hunits	Value	Income	Str. Age	Elasticity	Decline	Guren γ	Units γ	Space γ
Variable in column	-0.625	-2.029	-0.561	0.295***	-0.243	4.494*	0.272	0.737	0.770
	(0.452)	(1.110)	(2.932)	(0.075)	(0.387)	(2.138)	(0.938)	(3.503)	(3.487)
Constant	8.600	24.762	5.648	-4.157***	0.324	-1.543**	-0.321	-0.249	-0.450
	(6.496)	(13.889)	(30.655)	(1.015)	(0.715)	(0.580)	(0.847)	(0.681)	(1.503)
N	76	76	76	80	75	80	80	80	80
R^2	0.030	0.056	0.000	0.159	0.004	0.097	0.001	0.001	0.001

Large sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Hunits	Value	Income	Str. Age	Elasticity	Decline	Guren γ	Units γ	Space γ
Variable in column	-0.279	-0.872	0.332	0.177***	-0.146	4.776*	0.946*	-0.340	-0.401
	(0.159)	(0.474)	(1.226)	(0.041)	(0.099)	(2.194)	(0.454)	(1.380)	(1.377)
Constant	3.872	10.841	-3.173	-2.042***	0.657	-1.787**	-0.493	0.374	0.495
	(2.059)	(5.805)	(12.667)	(0.548)	(0.340)	(0.622)	(0.368)	(0.444)	(0.778)
N	245	245	245	249	249	75	247	244	244
R^2	0.010	0.016	0.000	0.082	0.005	0.108	0.020	0.000	0.000

Notes: *, **, and *** indicate $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively. Heteroskedasticity-consistent standard errors in parentheses. The dependent variable is the 30-year accumulated gap in the house price index, calculated as $P_{2020}^{rooms}/P_{2020}^{Uniform} - 1$. Housing units (log), housing value (log), income (log), and structure age are CBSA averages from the 5-year American Community Survey in 2019. Urban decline is the share of housing units below replacement costs in 1990 from Glaeser and Gyourko (2005), supply elasticity is from Saiz (2010), Guren γ is the sensitivity parameter in Guren et al. (2021), and the new units and space γs are CBSA mean supply elasticities from Baum-Snow and Han (2024) using the gamma01a formulation. Small sample and large sample include all CBSA's with a G&G (2005) decline value and Saiz (2010) elasticity value, respectively.

Sources: Contat and Larson (2024); Glaeser and Gyourko (2005); Saiz (2010); Guren et al. (2021); Baum-Snow and Han (2024); Census, American Community Survey; Authors' analysis.