Characteristics of a Sufficient Statistic to Measure City Housing Prices

Daniel Broxterman*
Florida State University Anthony M. Yezer†
George Washington University

Abstract

Accurately characterizing the levels of and changes in housing price in cities—whether measured by rent or asset value—is essential for various empirical analyses. However, this task is complicated by the heterogeneity of the housing stock, variation in accessibility, and the fact that housing is consumed jointly with neighborhood. This paper focuses on the issue of intra-city location, which economic theory suggests is systematically related to housing price. The key finding is that a sufficient statistic to describe both the level of and change in the average housing price requires that prices be measured as a function of distance from the city center and weighted by the fraction of the housing stock at each distance. While commonly used repeat sales and hedonic measures of price change are generally not weighted in this fashion, they 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)
†George Washington University, Washington, DC, USA (yezer@gwu.edu)

1 Introduction

The price indexes used by empirical researchers, market participants, and policy makers to characterize the cost of housing in cities are susceptible to measurement error. In addition to the stubborn index number problem of decomposing aggregate price and quantity changes for complex assets, tenure determines whether asset or rental values are observed and housing units are not distributed randomly by tenure across space. Consequently, assembling a sample of price changes that is representative of an entire metro market is challenging.

In this paper, we consider, based on economic theory, whether a sample of changes that is not fully randomized can sufficiently characterize the average rate of housing price change in a city. Conversely, if current techniques for measuring housing price change using hedonic or repeat sale indexes produce biased estimates, what modifications could be made to reduce or eliminate this bias.

There is substantial evidence that housing price movements, as well as their responses to specific shocks, vary significantly by location within a city. In the U.S., from the early-2000s housing boom until the COVID-19 pandemic, prices grew faster in areas closer to city centers, albeit with higher volatility. Notably, this spatial variation in appreciation rates was not mean reverting (Glaeser et al., 2012; Malone and Redfearn, 2018; Bogin et al., 2019a,b; Edlund et al., 2022). However, at the time of writing, the trend toward centralization has been disrupted by the market's response to the COVID-19 shock, which has shifted housing demand away from central cities to suburban areas with lower population density (Liu and Su, 2021; D'Lima et al., 2022).

The literature provides numerous other examples of housing price trends and shocks that vary spatially. For instance, a significant body of research shows that changes in transportation infrastructure (such as bus, highway, and rail facilities) affect the relationship between location and housing price appreciation in cities. Additionally, while Molloy and

Shan (2013) find that an increase in oil prices reduces construction near the city edge rather than lowering housing prices there, subsequent research by Larson and Zhao (2020) cannot reject an inverse relationship between oil prices and relative housing prices in the suburbs. Finally, Fischer et al. (2021) demonstrate that dynamic responses to monetary policy can vary dramatically across city housing markets.

The literature on weighted repeat sales price indexes has highlighted that disaggregation of the housing stock leads to differences in estimated rates of house price appreciation.[1] Additionally, the findings here align with those of Ambrose et al. (2015), who identify significant differences using a repeat rental index based on newly signed lease contracts, and Anenberg and Laufer (2017), who report differences with a repeat sales index based on contract rather than closing data.

This body of literature, however, does not rely on spatially randomized measures, so some may stem from the methods used to compute the housing price indexes. Consistent with assertions in these studies, we argue that even modest changes in the approach to measuring city housing price changes can significantly impact research results.

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.

[1] For early examples of diversity in patterns of appreciation within the housing stock, see Clapp and Giaccotto (1999) and Dreiman and Pennington-Cross (2004).

Laspeyres (1871) would answer the index number problem for measuring the average housing price and its rate of change, recognizing that price varies with distance from the city center, in a straightforward and familiar manner. Suppose that r_t^* is an index number indicating the average price per unit housing at time t, such that

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

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 $H_t = \sum_k h_t(k)$. The Laspeyres measure of the rate of change in housing price between $t-1$ and t is

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

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 is one assumption that can greatly simplify the computation of (2). If the percentage change in the price of housing is constant 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. The spatial distribution of housing may vary, but it is not essential to the computation of a sufficient statistic for measurement of the percentage change in housing cost.

3 Economic theory and sufficient housing price statistics

The previous section suggested that if the percentage change in housing price is not a function of distance, 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, 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, that varies with location, 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

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

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.[2] Housing

[2] 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

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.*

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

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

price differences as discussed below.

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 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 iso-utility assumption regarding urban households 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.

4 Illustration of the effect of distance on measures of housing price change

In this section, we use the the American Housing Survey Metropolitan Sample (AHS-MS) to illustrate, in a preliminary fashion, the potential effects of failing to consider

location 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. For purposes of illustration, the AHS-MS is particularly convenient for multiple reasons. To begin with, it surveys households in both renter and owner units and it divides housing spatially into central-city and suburban locations. In addition, the AHS-MS is longitudinal so that price changes can be calculated based on the same units in the same locations over time.

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 (reported in ranges) for all sample units in the metro area regardless of location. The results of this initial exercise are displayed in the first two columns of Table 1 for 25 US cities for the years 2015 to 2019. This was clearly a period of 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%.

What are the potential deficiencies of using the metrowide values in Table 1 to characterize housing price changes in the 25 cities covered? 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. Unfortunately, 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, we know that 2015 to 2019 was a period

during which centrality had become a more sought-after amenity in many large cities. Therefore, our first hypothesis is that the growth the rates of rental and asset prices displayed in Table 1 will be different 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 bps higher on average, respectively, consistent with the pre-COVID urban resurgence narrative. For our purposes, the estimates in Table 1 demonstrate that rates of appreciation are clearly 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 measurement error 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 simple 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. Taken at face value, these naive indexes, which appear in the first two columns of Table 2, would 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 a Laspeyres 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 in the previous section.

The Laspeyres index provides a useful reference point because it is consistent with other well known price index formulations. Differences between the Laspeyres and simple average indexes are in columns five and six of Table 2. On average, growth rates based on the Laspeyres (sufficient statistic) index are 43 and 4 bps lowers, respectively, in comparison with their simple average analogs. (Relative to the metrowide growth rates in Table 1, values

for the Laspeyres 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 using a Laspeyres index relative to a simple average of the central city and suburbs rates for those metros.

The empirical section in the final version of the paper will compare four different measures of house price change compiled from FHFA HPI estimates at the census tract level. First, one in which there is no weighting and city appreciation rates are based on a simple repeat sales measure. Second, one with weighting by census tract but the weighting is identical for all tracts—this gives a pure location effect. Third, weighting by location and unit size. Fourth the recommended weighting that includes all units (rental and owner) size adjusted, and locationally disaggregated.

5 Conclusions

Precise measurement of housing price changes, both rental and asset prices, is important for a number of research and practical purposes. Over time, numerous modifications to commonly used hedonic and repeat sale index methods have been suggested. To advance this effort, we address the issue of intra-city location which, based on empirical research and economic theory, is systematically related to housing price.

Previous papers have demonstrated that price changes are not uniform within cities. In particular changes in commuting cost, especially those associated with new transportation facilities or gasoline prices, have non-uniform effects on price change. This is the first paper to explore the implications of basic urban economic theory for spatial patterns of housing price for the construction of city-level housing price indexes. The results are that there are theoretical reasons, in addition to the empirical results, to expect that rates of appreciation

vary systematically with distance from the city center.

The implications of this paper for construction of a Laspeyers index are clear. Given that appreciation rates are not uniform, researchers must consider the spatial distribution of data used to create a housing price index, whether from survey responses or transactions. In particular, observations of price change should be weighted according to the fraction of housing at each distance from the city center in order for the index to sufficiently characterize the average rate of change for that city. Repeat sales and hedonic measures of price change are generally not weighted in this fashion. Fortunately, implementing such measures, particularly in repeat sales indexes, should not be difficult 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.

References

Ambrose, B. W., Coulson, N. E., and Yoshida, J. (2015). The repeat rent index. *Review of Economics and Statistics*, 97(5):939–950.

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

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.

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

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.

Dreiman, M. H. and Pennington-Cross, A. (2004). Alternative methods of increasing the precision of weighted repeat sales house prices indices. *The Journal of Real Estate Finance and Economics*, 28(4):299–317.

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

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.

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

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.

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.

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.

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: Alternative index constructions: 2015–2019

	Simple Avg		Laspeyres		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 Laspeyres index weights the growth rates by the total number of units, renter and owner, in each location. Differences between the indexes are calculated as Laspyres minus simple average.