Aggregation Bias in Price Indices for Multi-Family Rental Properties

*Randall S. Guttery**

Abstract. **This article examines aggregation bias in price index construction. Specifically, we test whether changes in values of 2- to 4-unit, multi-family rental housing properties vary systematically in the same market across property size. Moreover, we examine the time trend differences across locations within a geographic region for various sized multiplex properties, as well as investigate how size should be measured. Results suggest that absolute price changes are significantly different across property size, as determined by living area, and that the time trend does not differ across locations within a geographic region. Further research using this methodology is recommended for other property types.**

Introduction

Over the past thirty years, research on price index construction for single-family housing has expanded rapidly.[1] The purpose of such indices has been to provide real estate participants with a more efficient means of valuing property, studying mortgage default sensitivity and/or estimating housing demand. Price trends are important in estimating the probability of mortgagor default and prepayment patterns, and the predictive power of estimating housing demand is improved when prices can be measured with more confidence.

This article examines the extent of aggregation bias in constant-quality price index (CQI) construction. Specifically, we test whether *larger* multi-family rental properties appreciate or depreciate at a systematically different rate than *smaller* ones in the same market, as suggested by Clapp and Giaccotto (1995) for single-family dwellings.[2] Moreover, several approaches to classifying properties by size are offered. These price changes are measured via the hedonic approach by constructing CQIs for 2- to 4-unit, multi-family rental housing (hereafter multiplexes).[3] This study is an extension of Guttery and Sirmans (1995), who create *aggregate* multiplex constant-quality price indices for two geographic regions, Greater Manchester, Connecticut and Baton Rouge, Louisiana, from 1983 through 1988.

Guttery and Sirmans (1995) show that multiplex prices differed significantly among the two regions; Connecticut prices nearly doubled during the last three years of the

*Dept. of Finance, Insurance, Real Estate and Law, University of North Texas, Denton, TX 76203 or GUTTERY@UNT.EDU.
**Dept. of Finance Department, University of Connecticut, Storrs, CT 06269-2041.

sample period, while Baton Rouge prices declined over 55%. Moreover, prices generally were quite volatile from one quarter to the next. These finding complement those of Follain and Calhoun (1997), who create an index for multi-family properties with more than four units. What Guttery and Sirmans (1995) do not investigate, however, is whether CQIs for, say, duplexes change at a different rate than for triplexes or fourplexes, or whether CQIs for properties with less square feet of living area change at a different rate than for properties with more living area. Additionally, they do not test whether CQIs for various sized (*e.g.*, duplex vs. triplex/fourplex; small vs. large living area) multiplexes differ within each geographic region.

This article proceeds as follows. The second section provides a brief overview of possible sources of aggregation bias in price indices. The third section describes empirical models that can be used to create the CQIs for various sized multiplexes and offers the underlying hypotheses. The fourth section offers the methods of constructing price indices, while the fifth section discusses the data sets and offers descriptive statistics for the two samples. The sixth section provides the regression results, CQI values over the six-year period and recommendations as to which model(s) should be considered when classifying property by size. Section seven is the conclusion.

Sources of Aggregation Bias in Price Indices

The pooling of data often is necessary to have a sufficient number of observations, so that statistical inferences can be made with confidence. A trade-off of such aggregation may be the loss of explanatory power, however. For example, if a price index is being created for a geographic region without a large number of transactions, there may not be enough data points to test for time trend differences across locations or zones within the region. This is especially prevalent when dummy vectors are used as regressors; near singular matrices often result. Similarly, separate regressions for each location are not possible without sufficient data; by pooling the data, price changes by location cannot be measured. Therefore, the omitted ''location'' variables may have a significant effect on selling price, resulting in biased estimators.

Real estate participants such as agents, buyers, sellers and appraisers often contrast real estate by size. Examples include property being categorized by price or value, by living area, or by the number of units. While conceptually investors may agree upon what constitutes, say, a lower-valued property, relative to a mid-valued one, they likely will disagree when classifying properties at the margin. This article offers several approaches to codifying property by size.

Perhaps the most obvious method of categorizing multiplex properties is by sale price. An inherent problem with this approach, however, is that sale price will be both the dependent variable in the regression equation *and* used to formulate some of the independent variables in the form of dummy vectors for the various size classifications. Consequently, the coefficients of these explanatory variables will be highly significant because of collinearity, but at the expense of some regressors being self-contradicting (Arguea and Hsiao, 1993). If regressions are run for each classification, this is not a problem; however, an insufficient number of observations will result in pooling or aggregation of the data to obtain reliable results.

An alternative classification is by living area or leasable space. This is a legitimate method for all property types. The difficulty involved herein is determining the upper and lower bounds for each size category. While most real estate participants likely consider a duplex with 1200 total square feet of living area to be a small, lower-valued property, one with 1800 feet likely straddles the upper bound for small multiplexes and the lower bound for mid-valued ones. An approach to classifying the data by living area is offered in the next section.

A third method of classifying multiplexes is by the number of rental units in the dwelling. All else held equal, one would expect duplexes to sell for less than fourplexes, yet luxury duplexes may sell for more than lower-quality fourplexes. The present value of the duplexes' expected future income stream may exceed that of the fourplexes'. Nevertheless, classifying a property by its number of units is much simpler than doing so by its square feet of living area.

Once the preferred method of classifying the properties is chosen, a subclassification can be made, based upon the various locations within each geographic region. This will measure whether the time trend is different across the location zones. If the interaction terms between the regressors for each zone and the time trend are insignificantly different from zero, there is no need for locational interaction explanatory variables in the regression equation used to create the price indices.

A final comment is in order. Sample size may dictate how many classifications one can examine. The greater the number of transactions and the longer the period over which the CQIs are created, the lower the error variance and the higher the accuracy of the price indices, respectively (Clapp and Giaccotto, 1995).

Empirical Models

To build the indices, hedonic price equations are constructed for two samples: Greater Manchester, Connecticut and Baton Rouge, Louisiana. The sale price of each multiplex property, the dependent variable, is regressed on traditional explanatory variables that capture market conditions, size, the date of sale, location and financing terms. Model 1, the basic price equation, is:

$$SALEPRICE_i = \beta_0 + \beta_1 DOM_i + \beta_2 LA_i + \beta_3 TREND_i + \beta_4 ZONE1_i + \ldots + \beta_8 ZONE5_i + \beta_9 FIN_i + \epsilon_i, \quad (1)$$

where $SALEPRICE_i$ is the sale price of property i; β_0 is the intercept term and β_k ($k = 1, 2, \ldots, 9$) are coefficients of the explanatory variables; *DOM* is the number of days the property was on the market; *LA* is the square feet of living area for each property; *TREND* is a quarterly time-trend variable equal to one for *83Q1*, two for *83Q2*, ..., twenty-four for *88Q4*;[4] *ZONE1* through *ZONE5* are location dummy variables equal to one if the property were located in that respective zone or geographic area, and zero otherwise; *FIN* is a financing dummy variable equal to one if the property was sold with below-market financing, and zero otherwise;[5] and ϵ is an error term.

Model 1's explanatory variables are hypothesized to have the following effects on selling price. *SALEPRICE* is expected to move inversely with *DOM* because an increase in the number of days the property is on the market should decrease the sale price, all else held equal. Conversely, *LA* is expected to be positive, as an increase in the number of square feet of leasable space is expected to increase sale price. *TREND* is hypothesized to be positive (negative) because the Greater Manchester (Baton Rouge) economy was improving (declining) over the sample period. *ZONE1* through *ZONE5* are expected to be negative because the base case zone was predicted, ex-ante, to contain the most highly valued properties. Finally, *FIN* is projected to be positive because favorable financing should provide the seller with an opportunity to increase the sale price.

Model 2, an extension of the basic price equation, is designed to measure whether the time trend is different across zones. This ''locational factor'' regression is:

$$SALEPRICE_i = \gamma_0 + \gamma_1 DOM_i + \gamma_2 LA_i + \gamma_3 TREND_i + \gamma_4 ZONE1_i + \ldots + \gamma_8 ZONE5_i + \gamma_9 \mathrm{FIN}_i + \gamma_{10} TRENDZONE1_i + \ldots + \gamma_{14} TRENDZONE5_i + \mu_i, \quad (2)$$

where $SALEPRICE_i$ is as stated in Equation 2; γ_0 is the intercept term and γ_k ($k = 1, 2, \ldots, 14$) are coefficients of the explanatory variables; *DOM*, *LA*, *TREND*, *ZONE1* through *ZONE5*, and *FIN* are as defined in Model 1; *TRENDZONE1* through *TRENDZONE5* are interaction terms between *TREND* and *ZONE1* through *ZONE5*, respectively; and μ is an error term. *TRENDZONE6*, an interaction term between *TREND* and *ZONE6*, is the base case to which *TRENDZONE1* through *TRENDZONE5* are compared, where *ZONE6* is a dummy variable equal to one if the property is located in Zone 6, and zero otherwise.

DOM, *LA*, *TREND*, *ZONE1* through *ZONE5*, and *FIN* are anticipated to have the same signs as in Model 1. The signs of the interaction terms' coefficients are anticipated to be negative (positive) for the Greater Manchester (Baton Rouge) sample because *TREND* is expected to be positive (negative) and *ZONE1* through *ZONE5* are expected to be negative. Should the interaction terms' coefficients be insignificant, then the time trend does not differ across zones.

Model 3, an extension of the basic price equation, is designed to measure whether the time trend is different for multiplexes with varying square feet of living area. How are the upper and lower bounds for each size category determined? This article proposes three classifications: (1) *SMALL* represents small properties whose square feet of living area is more than one standard deviation below the mean living area in the sample; (2) *MEDIUM* represents mid-sized multiplexes whose square feet of living area are within one standard deviation of the mean living area in the sample; and (3)

LARGE represents large dwellings whose square feet of living area are more than one standard deviation above the mean living area in the sample. This ''size factor'' regression is:

$$SALEPRICE_i = \Delta_0 + \Delta_1 DOM_i + \Delta_2 LA_i + \Delta_3 TREND_i + \Delta_4 ZONE1_i + ... + \Delta_8 ZONE5_i$$
$$+ \Delta_9 FIN_i + \Delta_{10} TRENDLA_{MIN,i} + \Delta_{11} TRENDLA_{MAX,i} + \upsilon_i, \quad (3)$$

where $SALEPRICE_i$ is as stated in Equation 1; Δ_0 is the intercept term and Δ_k (k = 1, 2, ..., 11) are coefficients of the explanatory variables; *DOM*, *LA*, *TREND*, *ZONE1* through *ZONE5*, and *FIN* are as defined in Model 1; $TRENDLA_{MIN}$ ($TRENDLA_{MAX}$) is an interaction term between *TREND* and *MINLA* (*MAXLA*), where *MINLA* (*MAXLA*) is a size dummy variable equal to one if the property's living area is more than one standard deviation below (above) the mean living area in the sample, and zero otherwise; and υ is an error term. $TRENDLA_{MID}$, an interaction term between *TREND* and *MIDLA*, is the base case to which $TRENDLA_{MIN}$ and $TRENDLA_{MAX}$ are compared, where *MIDLA* is a size dummy variable equal to one if the property's living area is within one standard deviation of the mean living area in the distribution, and zero otherwise. *DOM*, *LA*, *TREND*, *ZONE1* through *ZONE5*, and *FIN* are anticipated to have the same signs as in Model 1. The signs of the interaction terms' coefficients are difficult to predict a priori because their effects depend upon how the properties' sizes affect the slope of the price index.

Model 4, an extension of the basic price equation, is designed to measure whether the time trend is different for properties with two units vs. three or four units. This ''unit factor'' regression is:

$$SALEPRICE_i = \lambda_0 + \lambda_1 DOM_i + \lambda_2 LA_i + \lambda_3 TREND_i + \lambda_4 ZONE1_i + ...$$
$$+ \lambda_8 ZONE5_i + \lambda_9 FIN_i + \lambda_{10} TRENDUNIT_i + \omega_i, \quad (4)$$

where $SALEPRICE_i$ is as stated in Equation 1; λ_0 is the intercept term and λ_k (k = 1, 2, ..., 10) are coefficients of the explanatory variables; *DOM*, *LA*, *TREND*, *ZONE1* through *ZONE5*, and *FIN* are as defined in Model 1; *TRENDUNIT* is an interaction term between *TREND* and *TRIFOUR*, where *TRIFOUR* is a size dummy variable equal to one if the property is either a triplex or fourplex, and zero otherwise; and ω is an error term. *TREND2UNIT*, an interaction term between *TREND* and *DUPLEX*, is the base case to which *TRENDUNIT* is compared, where *DUPLEX* is a size dummy variable equal to one if the property is a duplex, and zero otherwise. *DOM*, *LA*, *TREND*, *ZONE1* through *ZONE5*, and *FIN* are anticipated to have the same signs as in Model 1. The sign of the interaction term's coefficient is difficult to predict a priori because its effect depends upon how the properties' number of units affects the slope of the price index.

The Indices

Once the price equations are estimated for the various models, a constant-quality price for each quarter of the sample period can be calculated for the two samples. Model 1's constant-quality price equals:

$$CQPRICE_Q = b_0 + b_1 AVGDOM + b_2 AVGLA + b_3 TREND, \tag{5}$$

where $CQPRICE_Q$ is the constant-quality price in quarter Q (Q = *83Q1*, *83Q2*, ..., *88Q4*); b_k ($k = 0, 1, 2, 3$) are the *estimated values* of the intercept and quarterly regression coefficient results of *DOM*, *LA* and *TREND* from Equation 1; *AVGDOM* is the average days on the market for the sample properties; *AVGLA* is the multiplexes' average square feet of living area; and *TREND* is as defined previously. Then, the constant-quality prices are transformed into raw index values, such that the raw index value for quarter Q equals the constant-quality price in quarter Q divided by the constant-quality price in the base quarter, *83Q1*. Finally, the raw index values are scaled by a factor of 100 to obtain the CQIs.

Should Model 2's interaction terms' coefficients be significant, the time trend differs across zones. In this case, the constant-quality price for each zone equals:

$$CQPRICE_{Q,j} = g_0 + g_1 AVGDOM_j + g_2 AVGLA_j + g_3 TREND + g_4 TRENDZONE_j, \tag{6}$$

where $CQPRICE_{Q,j}$ is the constant-quality price in quarter Q (Q = *83Q1*, *83Q2*, ..., *88Q4*) for zone j ($j = 1, 2, ..., 6$); g_k ($k = 0, 1, ..., 4$) are the *estimated values* of the intercept and quarterly regression coefficient results of *DOM*, *LA*, *TREND* and $TRENDZONE_j$ from Equation 2; $AVGDOM_j$ is the average days on the market for the sample properties in zone j; $AVGLA_j$ is the average living area for properties in zone j; and, *TREND* and $TRENDZONE_j$ for zone j are as defined previously. If the interaction terms' coefficients are insignificant, Equation 6's corresponding CQI need not be calculated.

Should Model 3's interaction terms' coefficients be significant, classifying the properties by living area is an appropriate method of creating CQIs for various sized properties. In this case, the constant-quality price for each size classification equals:

$$CQPRICE_{Q,j} = d_0 + d_1 AVGDOM_j + d_2 AVGLA_j + d_3 TREND + d_4 TRENDLA_j, \tag{7}$$

where $CQPRICE_{Q,j}$ is the constant-quality price in quarter Q (Q = *83Q1*, *83Q2*, ..., *88Q4*) for properties with living area of size j (j = *SMALL*, *MEDIUM*, *LARGE*); d_k ($k = 0, 1, ..., 4$) are the *estimated values* of the intercept and quarterly regression coefficient results of *DOM*, *LA*, *TREND* and $TRENDLA_j$ from Equation 3; $AVGDOM_j$ is the average days on the market for the sample properties that are size j; $AVGLA_j$ is the average living area for properties that are size j; and *TREND* and $TRENDLA_j$ for size j are as defined previously. If the interaction terms' coefficients are insignificant, Equation 7's corresponding CQI need not be calculated.

Should Model 4's interaction term's coefficient be significant, classifying the properties by their number of rental units (*i.e.*, 2, 3 or 4) is an appropriate method of creating CQIs for various sized properties. In this case, the constant-quality price for each size classification equals:

$$CQPRICE_{Q,j} = l_0 + l_1 AVGDOM_j + l_2 AVGLA_j + l_3 TREND + l_4 TRENDUNIT_j, \tag{8}$$

where $CQPRICE_{Q,j}$ is the constant-quality price in quarter Q (Q = *83Q1*, *83Q2*, ..., *88Q4*) for property type j (j = duplex, triplex/fourplex); l_k (k = 0, 1, ..., 4) are the *estimated values* of the intercept and quarterly regression coefficient results of *DOM*, *LA*, *TREND* and *TRENDUNIT* from Equation 4; $AVGDOM_j$ is the average days on the market for the sample properties that are type j; $AVGLA_j$ is the average living area for properties that are type j; and *TREND* and $TRENDUNIT_j$ for size j are as defined previously. If *TRENDUNIT* is insignificant, Equation 8's corresponding CQI need not be calculated.

Data and General Analysis

The Multiple Listing Service (MLS) provided data for both samples from its *Quarterly Comparable Sales Books*. The Greater Manchester (Baton Rouge) sample consists of 528 (382) duplexes, triplexes and fourplexes that sold over the period January 1983 through December 1988. Sales are excluded from the sample if any relevant information is omitted from the MLS Comp Books, or if the properties are located outside of the six zones in each geographic region. Exhibit 1 reports the samples' mean values and standard deviations for each variable in Model 1, the basic regression.

The sale prices of Greater Manchester properties range from $35,000 to $267,000, with a mean of $128,325, while Baton Rouge properties range from $8,800 to $218,000, with a mean of $88,661. Days on the market, a proxy for market conditions, average 66 for the Greater Manchester sample, but 107 for Baton Rouge. The mean living area is 2560 square feet in Greater Manchester and 3564 square feet in Baton Rouge.

Greater Manchester's economy fared quite well over the sample period, while Baton Rouge's did not. For example, Greater Manchester's civilian labor force unemployment rate averaged 3.4%, the population grew 6.7%, average house prices

Exhibit 1
Model 1 Descriptive Statistics for Small Multi-Family Residential Properties in Greater Manchester and Baton Rouge

	Greater Manchester		Baton Rouge	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
SALEPRICE	128325	46600	88661	48640
DOM	66	67	107	126
LA	2560	608	3564	1010
TREND	16.86	5.28	13.72	8.42
ZONE1	0.02	na	0.09	na
ZONE2	0.02	na	0.11	na
ZONE3	0.25	na	0.08	na
ZONE4	0.04	na	0.30	na
ZONE5	0.03	na	0.19	na
FIN	0.05	na	0.17	na

increased 123.3% and inflation-adjusted per capita personal income rose 30.1%. On the other hand, Baton Rouge's civilian labor force unemployment rate averaged 8.4%, the population declined 1.2%, average house prices dropped 13.0%, and inflation-adjusted per capita personal income rose only 0.2%.[6] These statistics lend support as to why Greater Manchester's multiplex prices are expected to rise in value, while Baton Rouge's are predicted to decline.

Exhibit 2 reports selected descriptive statistics by classification for multiplexes in the two geographic regions. As in Model 3, Panel A separates the data by living area for small, mid-sized and large properties, reporting each category's number of transactions, mean square feet of living area and mean days on the market, as well as corresponding standard deviations, maximums and minimums. For Connecticut, there are 93 small properties, 351 mid-sized ones and 84 large multiplexes. The mean living area is 1655 square feet for small dwellings, 2578 square feet for the mid-sized and 3485 square feet for large properties. Interestingly, the mean days on the market is inversely related to property size. Small properties averaged 84 days on the market, while the mid-sized ones averaged 65 days and large multiplexes sold in 48 days. These results are in contrast to the single-family dwelling market, where large, higher-valued properties are generally on the market for longer periods than small, lower-valued houses.

For Baton Rouge, there are 79 small properties, 259 mid-sized ones and 44 large multiplexes. The mean living area is 1989 square feet for small multiplexes, 3797 square feet for mid-sized ones and 5020 square feet for large properties. Small properties averaged 127 days on the market, while mid-sized ones averaged 99 days and large ones sold in 117 days.

As in Model 4, Panel B separates the data by the number of units (*i.e.*, duplex vs. triplex/fourplex), reporting the two categories' number of transactions, mean square feet of living area and mean days on the market, as well as corresponding standard deviations, maximums and minimums. For Connecticut, there are 352 duplexes and 176 triplexes/fourplexes. The mean living area for duplexes is 2460 square feet, which is only 297 square feet smaller than the mean living area of 2757 square feet for three- or four-unit dwellings. As well, duplexes averaged 63 days on the market, while triplexes/fourplexes averaged 71 days.

For Baton Rouge, there are 120 duplexes and 262 triplexes/fourplexes. Interestingly, the percentage of duplexes vs. triplexes/fourplexes in Greater Manchester is exactly the opposite in Baton Rouge; two-thirds (one-third) of Greater Manchester's (Baton Rouge's) sample properties are (is) duplexes, while one-third (two-thirds) is (are) three- to four-unit dwellings. The mean living area is 2985 square feet for duplexes, but 3829 square feet for three- or four-unit dwellings. As well, duplexes averaged 110 days on the market, while triplexes/fourplexes averaged 106 days.

Exhibit 2
Descriptive Statistics by Classification for Small Multi-Family Residential Properties in Greater Manch
Baton Rouge

	Greater Manchester, CT					Baton Rouge, LA			
	Trans.[a]	Mean	Std. Err.	Min.	Max.	Trans.[a]	Mean	Std. Err.	M
Panel A: Model 3 The Size Factor Regression									
MINLA	93	1655	221	1020	1952	79	1989	406	9
MIDLA	351	2578	332	1953	3168	259	3797	502	25
MAXLA	84	3485	205	3169	4062	44	5020	424	45
MINDOM	na	84	85	1	409	na	127	142	
MIDDOM	na	65	65	1	483	na	99	122	
MAXDOM	na	48	46	2	270	na	117	117	
Panel B: Model 4 The Unit Factor Regression									
2-Unit *LA*	352	2460	598	1020	4054	120	2985	1133	9
3/4-Unit *LA*	176	2757	580	1458	4062	262	3829	825	13
2-Unit *DOM*	na	63	66	1	483	na	110	120	
3/4-Unit *DOM*	na	71	69	1	291	na	106	129	

[a]Number of transactions.

Regression Results and Constant-Quality Index Values

Model 1: The Basic Regression

Exhibit 3, Panel A (Panel B) reports Greater Manchester's (Baton Rouge's) regression results for all four models. The basic regression model shows that for the Greater Manchester sample, *DOM*, *LA* and *TREND* are significant at the 1% level and of the hypothesized signs, and *FIN* is insignificant. The adjusted R^2 is .71, and the *F*-Statistic

Exhibit 3
Regression Results for Creating Constant-Quality Indices in Greater Manchester and Baton Rouge

	Model 1	Model 2	Model 3	Model 4
Panel A: Greater Manchester, CT				
Constant	13318.2	−2908.5	36683.4	14437.0
	(2.1)*	(−0.3)	(4.1)**	(2.3)*
DOM	−44.7	−47.8	−40.8	−45.7
	(−2.7)**	(−2.9)**	(−2.5)*	(−2.8)**
LA	24.7	24.1	15.0	24.3
	(12.3)**	(11.9)**	(4.4)**	(11.6)**
TREND	3878.1	4841.7	3998.9	3858.0
	(14.9)**	(10.4)**	(15.4)**	(14.7)**
ZONE1	−12565.0	−57033.8	−8358.5	−12985.9
	(−1.5)	(−0.6)	(−1.0)	(−1.5)
ZONE2	−28532.4	2192.5	−29570.6	−28479.9
	(−3.7)**	(0.1)	(−3.9)**	(−3.7)**
ZONE3	−30731.3	−8867.0	−31394.8	−31094.5
	(−10.6)**	(−0.9)	(−10.9)**	(−10.6)**
ZONE4	−27280.0	−6182.1	−28677.5	−27701.7
	(−4.8)**	(−0.4)	(−5.0)**	(−4.8)**
ZONE5	−32290.2	3614.1	−33752.5	−32288.2
	(−4.5)**	(0.1)	(−4.8)**	(−4.5)**
FIN	−4198.5	−4294.1	−6205.5	−4661.0
	(−0.8)	(−0.8)	(−1.2)	(−0.9)
TRENDZONE1		2717.1		
		(0.5)		
TRENDZONE2		−2276.5		
		(−1.7)		
TRENDZONE3		−1263.0		
		(−2.3)*		
TRENDZONE4		−1185.6		
		(−1.4)		
TRENDZONE5		−2159.0		
		(−1.4)		
$TRENDLA_{MIN}$			−958.7	
			(−3.4)**	
$TRENDLA_{MAX}$			503.2	
			(2.4)*	
TRENDUNIT				105.2
				(0.8)

Exhibit 3 (continued)
Regression Results for Creating Constant-Quality Indices in Greater Manchester and Baton Rouge

	Model 1	Model 2	Model 3	Model 4
Panel B: Baton Rouge, LA				
Constant	66221.2	75780.6	31354.8	67168.8
	(11.6)**	(11.5)**	(4.7)**	(11.7)**
DOM	11.5	6.4	9.2	11.5
	(1.2)	(0.7)	(1.0)	(1.2)
LA	25.3	24.6	35.6	25.1
	(19.5)**	(20.0)**	(21.1)**	(19.3)**
TREND	−4601.2	−5053.5	−4656.3	−4791.2
	(−28.9)**	(−18.0)**	(−30.9)**	(−23.5)**
ZONE1	−13225.8	−56913.0	−16042.7	−13113.8
	(−2.8)**	(−4.6)**	(−3.6)**	(−2.8)**
ZONE2	−7996.1	−16717.2	−9563.9	−7836.5
	(−1.8)	(−1.9)	(−2.4)*	(−1.8)
ZONE3	−25084.6	−62508.3	−21030.9	−24951.6
	(−4.8)**	(−7.6)**	(−4.4)**	(−4.8)**
ZONE4	−2715.1	−4422.5	−5852.4	−2934.0
	(−0.8)	(−0.7)	(−1.9)	(−0.9)
ZONE5	−7887.6	−13303.6	−11717.6	−7881.3
	(−2.1)*	(−1.8)	(−3.4)**	(−2.1)*
FIN	2526.6	3041.1	3062.6	2451.0
	(0.7)	(0.9)	(1.0)	(0.7)
TRENDZONE1		2366.1		
		(3.8)**		
TRENDZONE2		591.6		
		(1.1)		
TRENDZONE3		4776.2		
		(5.9)**		
TRENDZONE4		22.8		
		(0.1)		
TRENDZONE5		382.5		
		(0.9)		
$TRENDLA_{MIN}$			1633.6	
			(6.2)**	
$TRENDLA_{MAX}$			−1610.7	
			(−7.2)**	
TRENDUNIT				245.6
				(1.5)

Note: The *t*-Statistics are in parenthesis.
*Significance at the 5% level.
**Significance at the 1% level.

of 147.53 is significant at the 1% level. The other three models' results are similar. There is a $3,878 upward price trend per quarter,[7] such that the constant-quality price in *83Q1*, $73,600, would appreciate over 121% to $162,794 by *88Q4*. All zones' multiplexes sold for about $30,000 less than those in the base location, with the exception of those in Zone 1, where average prices were about \$12,600 less. *ZONE2* through *ZONE5* are significant at the 1% level and *ZONE1* is insignificant.

For the Baton Rouge sample, *LA* and *TREND* are significant at the 1% level and of the hypothesized signs, and *DOM* is insignificant. The adjusted R^2 is .78, and the *F*-Statistic of 150.00 is significant at the 1% level. The other three models' results are similar. There is a \$4,601 downward price trend per quarter, such that the constant-quality price in *83Q1*, \$157,654, would depreciate over 67% to \$51,831 by *88Q4*. All zones' multiplexes sold for less than those in the base location and most zone variables are significant.

Model 2: The "Locational Factor" Regression

Model 2 in Exhibit 3 provides a test of whether the time trend is different across locations. For the Greater Manchester sample, the interaction terms between *TREND* and all but one of the five zones are insignificant, indicating that there were no significant differences across most zones. For the Baton Rouge sample, the interaction terms between *TREND* and three of the five zones are insignificant. The time trend is not statistically different across most locations. The basic result for Greater Manchester (Baton Rouge) is that the market appears to be appreciating (depreciating) at about the same rate across locations.

Model 3: The "Size Factor" Regression

Are there variations in price changes with variations in square feet of living area? Exhibit 2 reports the three classifications' descriptive statistics for both geographic regions. For the Greater Manchester sample, Exhibit 3 shows that there is a \$3,999 upward price trend per quarter for a mid-sized property, such that its constant-quality price in *83Q1*, \$72,798, would appreciate over 126% to \$164,775 by *88Q4*. $TRENDLA_{MIN}$'s coefficient of −958.74, significant at the 1% level, is added to *TREND* to obtain the quarterly net trend adjustment of \$3,040 for small multiplexes. Similarly, $TRENDLA_{MAX}$'s coefficient of 503.24, significant at the 2% level, obtains the quarterly net trend adjustment of \$4,502 for large properties.

For the Baton Rouge sample, there is a \$4,656 downward price trend per quarter for a mid-sized property, such that its constant-quality price in *83Q1*, \$167,538, would depreciate nearly 64% to \$60,450 by *88Q4*. $TRENDLA_{MIN}$'s coefficient of 1633.64, significant at the 1% level, obtains the quarterly net trend adjustment of −\$3,022 for small properties. Similarly, $TRENDLA_{MAX}$'s coefficient of −1610.74, significant at the 1% level, obtains the quarterly net trend adjustment of −\$6,267 for large ones. Because the time trend in both geographic regions differs significantly for the three size classifications based on living area, these properties are categorized accordingly, in order to create the CQIs. The basic result is that size, as determined by living area, is an important consideration. Further research using this methodology is recommended for other property types, such as single-family dwellings.

Exhibit 4
Constant-Quality Prices and Standardized Indices for Various Sized Multi-Family Rental Properties in Greater Manchester, CT from 1983 through 1988

	Small *LA*		Mid-Sized *LA*		Large *LA*	
Quarter	Price ($)	Index	Price ($)	Index	Price ($)	Index
83Q1	58,117	100.0	72,798	100.0	87,111	100.0
83Q2	61,157	105.2	76,797	105.5	91,613	105.2
83Q3	64,197	110.5	80,796	111.0	96,115	110.3
83Q4	67,237	115.7	84,795	116.5	100,617	115.5
84Q1	70,277	120.9	88,794	122.0	105,119	120.7
84Q2	73,317	126.2	92,793	127.5	109,621	125.8
84Q3	76,357	131.4	96,792	133.0	114,123	131.0
84Q4	79,397	136.6	100,791	138.5	118,625	136.2
85Q1	82,437	141.9	104,790	144.0	123,127	141.3
85Q2	85,477	147.1	108,789	149.4	127,629	146.5
85Q3	88,517	152.3	112,788	154.9	132,131	151.7
85Q4	91,557	157.5	116,787	160.4	136,633	156.9
86Q1	94,597	162.8	120,786	165.9	141,135	162.0
86Q2	97,637	168.0	124,785	171.4	145,637	167.2
86Q3	100,677	173.2	128,784	176.9	150,139	172.4
86Q4	103,717	178.5	132,783	182.4	154,641	177.5
87Q1	106,757	183.7	136,782	187.9	159,143	182.7
87Q2	109,797	188.9	140,781	193.4	163,645	187.9
87Q3	112,837	194.2	144,780	198.9	168,147	193.0
87Q4	115,877	199.4	148,779	204.4	172,649	198.2
88Q1	118,917	204.6	152,778	209.9	177,151	203.4
88Q2	121,957	209.9	156,777	215.4	181,653	208.5
88Q3	124,997	215.1	160,776	220.9	186,155	213.7
88Q4	128,037	220.3	164,775	226.4	190,657	218.9

Model 4: The "Unit Factor" Regression

Model 4 in Exhibit 3 reports the results of categorizing multiplexes by the number of units. Because of a relatively small number of properties with four (three) units in Greater Manchester (Baton Rouge), however, triplexes and fourplexes are aggregated. Neither sample's time trend is significantly different for multiplexes with two units vs. those with three or four units. These results suggest that multiplexes not be classified by the number of units because the price trend is the same across the number of units.

Constant-Quality Index Values

The empirical results suggest that multiplexes be classified by square feet of living area, rather than by sale price or the number of units. As well, the time trend is not significantly different across zones within a geographic region. Therefore, Equation 7 is used to create the price of an average, constant-quality multiplex property over the

Exhibit 5
Constant-Quality Prices and Standardized Indices for Various Sized Multi-Family Rental Properties in Baton Rouge, LA from 1983 through 1988

	Small *LA*		Mid-Sized *LA*		Large *LA*	
Quarter	Price ($)	Index	Price ($)	Index	Price ($)	Index
83Q1	103,382	100.0	167,538	100.0	211,274	100.0
83Q2	100,360	97.1	162,882	97.2	205,007	97.3
83Q3	97,338	94.2	158,226	94.4	198,740	94.1
83Q4	94,316	91.2	153,570	91.7	192,473	91.1
84Q1	91,294	88.3	148,914	88.9	186,206	88.1
84Q2	88,272	85.4	144,258	86.1	179,939	85.2
84Q3	85,250	82.5	139,602	83.3	173,672	82.2
84Q4	82,228	79.5	134,946	80.6	167,405	79.2
85Q1	79,206	76.6	130,290	77.8	161,138	76.3
85Q2	76,184	73.7	125,634	75.0	154,871	73.3
85Q3	73,162	70.8	120,978	72.2	148,604	70.3
85Q4	70,140	67.9	116,322	69.4	142,337	67.4
86Q1	67,118	64.9	111,666	66.7	136,070	64.4
86Q2	64,096	62.0	107,010	63.9	129,803	61.4
86Q3	61,074	59.1	102,354	61.1	123,536	58.5
86Q4	58,052	56.2	97,698	58.3	117,269	55.5
87Q1	55,030	53.2	93,042	55.5	111,002	52.5
87Q2	52,008	50.3	88,386	52.8	104,735	49.6
87Q3	48,986	47.4	83,730	50.0	98,468	46.6
87Q4	45,964	44.5	79,074	47.2	92,201	43.6
88Q1	42,942	41.5	74,418	44.4	85,934	40.7
88Q2	39,920	38.6	69,762	41.6	79,667	37.7
88Q3	36,898	35.7	65,106	38.9	73,400	34.7
88Q4	33,876	32.8	60,450	36.1	67,133	31.8

sample period. Exhibit 4 (5) reports these prices and indices for the Greater Manchester (Baton Rouge) sample.

The constant-quality price of a small multiplex in the first quarter of 1983 equals $58,117 for Greater Manchester and $103,382 for Baton Rouge, but by the fourth quarter of 1988, Greater Manchester's property increases over 120% to $128,037, while Baton Rouge's decreases over 67% to $33,876. The constant-quality price of a mid-sized property in the first quarter of 1983 equals $72,798 for Greater Manchester and $167,538 for Baton Rouge, but by the fourth quarter of 1988, Greater Manchester's property increases over 126% to $164,775, while Baton Rouge's decreases nearly 64% to $60,450. Finally, the constant-quality price of a large multiplex in the first quarter of 1983 equals $87,111 for Greater Manchester and $211,274 for Baton Rouge, but by the fourth quarter of 1988, Greater Manchester's property increases over 118% to $190,657, while Baton Rouge's decreases nearly 69% to $67,133.

The Greater Manchester constant-quality indices for all three size classifications increased approximately 122% over the six-year sample period. This is virtually identical to local single-family dwellings' appreciation rate of 123%. A small, multi-family rental property valued at $100,000 in *83Q1* would be worth about $222,000 by *88Q4*. The Baton Rouge CQIs for all three classifications decreased approximately 67%, unlike the single-family dwellings' depreciation rate of only 13%. This disparity raises an interesting question as to whether single-family dwellings are better insulated against declining markets than residential investment property. A small, multi-family rental property valued at $100,000 in *83Q1* would be worth only about $33,000 by *88Q4*. Overall, a constant-quality multiplex property would be worth nearly seven times more in Greater Manchester than in Baton Rouge over the sample period.

Conclusion

Constant-quality price indices have been created by numerous researchers for single-family residential properties, but the small multi-family rental housing market (*i.e.*, properties with two to four units) has been largely ignored. Furthermore, CQIs based on property size have not been constructed, and the time trend differences across zones within a geographic region have not been tested for various sized multiplex properties. This article creates CQIs for such properties in Greater Manchester, Connecticut and Baton Rouge, Louisiana from January 1983 through December 1988.

Specifically, this article explores the effects of aggregation bias on price index construction. Property size is based on both the square feet of living area and the number of rental units. Measuring size by living area, the CQIs suggest small income-producing residential real estate values more than doubled in Greater Manchester, but depreciated by two-thirds in Baton Rouge. The two regions' multiplex prices were quite volatile, especially from 1985 through 1988.

Regression results for both samples suggest that property size should be determined by square feet of living area, rather than by the number of units, and that the time trend does not differ across zones within a geographic region. Furthermore, the CQIs suggest multiplexes appear to appreciate or depreciate (in relative terms) at about the same rate, regardless of size. These findings are strengthened because virtually identical conclusions are reached for the sample properties in two geographic regions that were affected by very different economic conditions. This study adds to the meager evidence on price movements among regions, property types, property size and zones within a geographic region.

Notes

[1]A partial list of references include Bailey, Muth and Nourse (1963), Mark and Goldberg (1984), Case and Shiller (1987, 1989), Clapp and Giaccotto (1990, 1992), Pollakowski and Wachter (1990), Abraham and Schauman (1991), Case, Pollakowski and Wachter (1991), Case and Quigley (1991), Haurin and Hendershott (1991), Haurin, Hendershott and Kim (1991), Shiller (1991), Gatzlaff and Haurin (1993) and Jud and Seaks (1994).

[2]Follain and Calhoun (1997) suggest that "no widely available index of the price of multi-family rental housing properties exists." Though there are four multi-family price indices—the

Department of Commerce Index, NACREIT Apartment Index, National Real Estate Index (NREI) and Freddie Mac Repeat Sales Index—each has limitations.

[3]Hedonic prices are defined by Rosen (1974) as the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specified amounts of characteristics associated with them.

[4]Arguea and Hsiao (1993) suggest that in competitive markets, the hedonic pricing function should be linear.

[5]Both Greater Manchester and Baton Rouge are subdivided into six zones: the zone with the highest income and/or lowest unemployment rate is the base case zone to which other zones are compared. Below-market financing is defined as properties that were owner-financed, sold on assumption or exchanged.

[6]Sources: Bureau of Labor Statistics, Bureau of the Census, the University of Connecticut's Center for Real Estate and Urban Economic Studies and Louisiana State University's Real Estate Research Institute.

[7]When alternative functional forms were run (*e.g.*, the log of sale price), the index values for both samples were virtually unchanged. Therefore, the unlogged cases are reported for ease of interpretation. The linear trend is utilized because it appears to capture market conditions well.

References

Arguea, N. M. and C. Hsiao, Econometric Issues of Estimating Hedonic Price Functions, *Journal of Econometrics*, 1993, 56, 243–67.

Abraham, J. M. and W. S. Schauman, Measuring House Price Inflation, *Secondary Mortgage Markets*, 1991, Winter, 8–12.

Bailey, M. J., R. F. Muth and H. O. Nourse, A Regression Method for Real Estate Price Index Construction, *Journal of the American Statistical Association*, 1963, 58, 933–42.

Case, B., H. O. Pollakowski and S. M. Wachter, On Choosing Among House Price Index Methodologies, *Journal of the American Real Estate and Urban Economics Association*, 1991, 19, 286–07.

Case, B. and J. M. Quigley, Dynamics of Real Estate Prices, *Review of Economics and Statistics*, 1991, 73, 50–8.

Case, K. E. and R. J. Shiller, Prices of Single-Family Homes Since 1970: New Indexes for Four Cities, *New England Economic Review*, 1987, September/October, 45–56.

——, The Efficiency of the Market for Single-Family Homes, *American Economic Review*, 1989, 79, 125–37.

Clapp, J. M. and C. Giaccotto, Estimating Price Indices for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods, *Journal of the American Statistical Association*, 1990, 87, 300–06.

——, Estimating Price Trends for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods, *Journal of Real Estate Finance and Economics*, 1992, 5, 357–74.

——, Which Real Estate Price Index is Right for Mortgage Backed Securities? *Real Estate Finance*, 1995, 11:2, 41–7.

Follain, J. R. and C. A. Calhoun, Constructing Indexes of the Price of Multifamily Properties Using the 1991 Residential Survey of Finance, *Journal of Real Estate Finance and Economics*, 1997, 14:1,2, 235–55.

Gatzlaff, D. H. and D. R. Haurin, Sample Selection and Biases in Local House Value Indices, paper presented at the American Real Estate and Urban Economics Association meeting, January 1993.

Guttery, R. S. and C. F. Sirmans, Creating a Constant-Quality Index for Small, Multi-Family Rental Housing, *Journal of Real Estate Research*, 1995, 10, 557–67.

Haurin, D. R. and P. H. Hendershott, Housing Price Indexes: Issues and Results, *Journal of the American Real Estate and Urban Economics Association*, 1991, 19, 259–69.

Haurin, D. R., P. H. Hendershott and D. Kim, Local House Price Indexes, 1982–1991, *Journal of the American Real Estate and Urban Economics Association*, 1991, 19, 451–72.

Jud, G. D. and T. G. Seaks, Sample Selection Bias in Estimating Housing Sales Prices, *Journal of Real Estate Research*, 1994, 9, 289–98.

Mark, J. H. and M. A. Goldberg, Alternative Housing Price Indices: An Evaluation, *Journal of the American Real Estate and Urban Economics Association*, 1984, 12, 31–49.

Pollakowski, H. O. and S. M. Wachter, The Effects of Land-Use Constraints on Housing Prices, *Land Economics*, 1990, 66, 315–24.

Rosen, S., Hedonic Prices and Implicit Markets: Product Differentiation in Perfect Competition, *Journal of Political Economy*, 1974, 82, 34–55.

Shiller, R. J., Arithmetic Repeat Sales Price Estimators, *Journal of Housing Economics*, 1991, 1, 110–26.

The authors thank Joseph D. Albert and G. Don Jud, editors, as well as two anonymous referees. We are grateful to the Greater Baton Rouge Association of Realtors and the Willimantic Board of Realtors for providing data.