Alternative Approaches for Resale Housing Price Indexes

Erwin Diewert, Ning Huang, and Kate Burnett-Isaacs[1] May 1, 2017
Discussion Paper 17-03,
Vancouver School of Economics,
The University of British Columbia,
Vancouver, Canada, V6T 1L4.

Abstract

The use of hedonic regression models on the sales of detached housing units is widespread in the real estate literature. However, these models do not address the need to decompose the sale price into structure and land components. In the international System of National Accounts, it is necessary to obtain separate estimates for the price and quantity of housing structures and the land that these structures sit on. The builder's model accomplishes this decomposition but it has only been applied to Dutch and Japanese data. The paper will apply the builder's model to data on sales of detached houses in Richmond, British Columbia to test the robustness of the model. The property price indexes generated by the builder's model are also compared to the corresponding indexes generated by a traditional time product dummy hedonic regression model. The implied structure depreciation rates generated by both models are also compared. We find that if a sufficient number of housing characteristics are included in the hedonic regressions, the two approaches generate similar overall property price indexes and similar geometric depreciation rates. However, the two approaches do not generate similar land and structure subindexes.

Keywords

House price indexes, land and structure price indexes, hedonic regressions, net depreciation rates, System of National Accounts, the Builder's Model

JEL Classification Numbers

C2, C23, C43, E31, R21.

1. Introduction

In order to fill out the cells in a nation's balance sheets, it is necessary to decompose property value into land and structure components. The national balance sheets are used to construct measures of household wealth (which influences consumption decisions) and to construct estimates of the capital stock (by type of asset) that is used by the production sector. Information on the amount of land (and its price) is necessary in order to determine the productivity performance of the country. The problem is that information on the value of the *land* component of property value is almost completely lacking. Thus in this paper, we will attempt to partially fill this statistical gap by indicating how residential property values could be decomposed into land and structure components.

There are various hedonic regression models that could be used to obtain a decomposition of property value into land and structure components. We will use data on the sales of detached houses in a suburb of Vancouver, Canada to illustrate these methods. The data will be explained in Section 2. Most of our paper will focus on the use of the *builder's model* which will be explained in Section 3. The results of the hedonic regressions described in Section 3 will be used in Section 4 in order to construct constant quality overall property price indexes for the sales of detached houses as well as constant quality land and structure subindexes. In Section 5, we indicate how the results of Section 3 could be used to form approximate price indexes for the stock of houses (as opposed to the quarterly sales of houses).

The hedonic regression models that are utilized in Sections 3-5 of the paper make use of a geometric (or declining balance) model of structure depreciation. In Section 6, we switch to straight line depreciation and a generalization of straight line depreciation (piecewise linear depreciation). We rerun the models explained in Section 3 in Section 6, but using this alternative piecewise linear depreciation model. In the end, we find that under certain conditions, there is little difference in the indexes that are due to the different depreciation models.

In Section 7, we turn away from the builder's model for a hedonic property price regression and we look at the more *traditional property price hedonic regressions* which use the logarithm of the property's selling price as the dependent variable and enter the various characteristics of the property as independent variables in a linear regression with time dummy variables. We follow the example of McMillen (2003) and Shimizu, Nishimura and Watanabe (2010) and show how these traditional log price hedonic regressions can be manipulated to give estimates for a geometric depreciation rate for the structure. We compare this imputed depreciation rate with the geometric depreciation rates that we obtained in Section 3 (and we find that there is a close comparison). We also compare the overall property price index that is generated by the traditional log price hedonic regression approach with our comparable builder's model overall price indexes and again, we find that there is a close comparison. However, we show that the traditional log price hedonic regression approach does not generate reasonable land and structure price subindexes.

Section 8 concludes.

2. Data

The data for this study were obtained from the Multiple Listing Service for the city of Richmond, British Columbia, Canada.[2] Richmond is a suburb of the city of Vancouver and lies immediately to the south of Vancouver. There were a total of 13,199 observations on the sales of detached houses in the Richmond region for the period from January 2008 to November 2016.[3] However, not all of these observations were used in our hedonic regressions: before doing the estimation, we deleted the tails of the distributions of the dependent variable (the selling price of the property) and the explanatory characteristics. Including range outliers in the regression can distort the results to a considerable degree. Because the number of observations at the tails of each characteristic distribution is small, a hedonic regression surface cannot be reliably measured at these observations that have extreme values for the underlying property characteristics.

The definitions for the variables used in the regressions and their units of measurement are as follows:

V = selling price of the property in thousands of dollars;
L = size of the lot area measured in thousands of square feet;
S = structure area (total floor area) measured thousands of square feet;
A = age if the structure in years;
NBE = number of bedrooms;
NBA = Number of bathrooms.[4]

We now explain the details of our deletion process. We first removed 2 observations that were not freehold. The first main variable listed above that we undertake range deletions is the selling price variable V. Since the selling prices increase over years, we examine the histograms of selling prices quarter by quarter. We found that the distribution of selling prices for each year is right-skewed, with the mass of the distribution of the selling prices concentrated on the left part of the histograms. In order to avoid the problem that a few observations at the top end of the selling price are spread over a large range of prices, we dropped observations with selling prices above the top 3% of sales by year. We also deleted observations that are in the bottom 1% of sales by year. Through this process, 541 observations were dropped from the sample.

---

[2] We obtained the data from Raymond Chan, who obtained the data from the MLS@ system (Multiple Listing Service), a branch of the Canadian Real Estate Association, for research purposes at Simon Fraser University.
[3] The monthly sales data were aggregated into quarterly data. Our last "quarter" (Quarter 36) is missing the sales for the month of December.
[4] A half bathroom is regarded as a full bathroom in this paper.

At the next stage, we deleted outliers of the explanatory variables. Before deleting outliers, we dropped 472 observations that have missing values in age A, 6 observations with missing values in the floor space area S and 17 observations with zero bedrooms. To determine the range of the main explanatory variables, we examined the histograms of these variables. As mentioned earlier, the purpose of removing the outliers is to avoid having only a few observations at either the bottom end or top end of the distribution. After several rounds of trimming, the final dataset included 11,045 observations with the following characteristics:[5]

- The land plot area L is between 3000 and 12000 square feet;
- Floor space area S (also called living area or structure area) is between 1000 and 4800 square feet;
- The age A of the structure is less than or equal to 60 years;
- The structure has 1 to 6 bathrooms (NBA);
- The structure has 3 to 7 bedrooms (NBE);
- The structure has 1 to 3 kitchens;
- The structure has less than 4 covered parking spots;
- For the sales prices, we deleted the bottom 1% and approximately the top 3% of selling prices by year.

We did not use the kitchen or parking characteristics in our regressions; these variables were used to eliminate properties with an unusual number of kitchens or parking spots.

In addition to the above variables, we had information on which one of 6 postal code regions for Richmond was assigned to each property.[6]

Finally, we created a quarterly residential augmented construction cost index[7] for the Metropolitan Area of Vancouver which we denote as p_{St} for quarter $t = 1,...,36$. This index (in dollars per square foot of floor space area) is derived from Statistics Canada residential construction cost models for Metro Vancouver. This augmented construction cost index may be a bit low with respect to its levels but it probably captures the rates of change of augmented construction costs over our sample period.[8]

---

[5] Thus we deleted 16.3% of our observations. This seems to be a high number of deletions but we feel that accurate hedonic surfaces cannot be estimated when the number of observations is sparse at the edges of the hedonic surface.

[6] We grouped the properties based on the forward sortation area (FSA), which is a geographical area defined based on the first three characters in a Canadian postal code.

[7] It is an augmented construction cost index because it may include some element of contractor profit. However, for our purposes, this is appropriate: if one buys a newly constructed house, the purchaser certainly has to pay for the contractor's profit margin.

[8] It may be the case that this index understates the cost of house construction in the Greater Vancouver area. From the Altus Group (2015) Construction Cost Guide for 2015, we find the following range of house construction costs per square foot for the Vancouver area: Speculative Basic Quality: $100 - $165; Speculative Medium Quality: $165 - $225; Speculative High Quality: $225 - $350; Custom Built: $400 - $1,000. The Statistics Canada estimated Vancouver augmented construction cost may be on the low end of the quality spectrum. We will address this issue later in the paper; i.e., our Models 4-7 allow for an upward adjustment in the *level* of the Statistics Canada quarterly structure prices.

The basic descriptive statistics for the above variables are listed in Table 1 below.

Table 1: Descriptive Statistics for the Variables

Name	No. of Obs.	Mean	Std. Dev	Minimum	Maximum	Unit of Measurement
V	11045	$1134.1	473.61	$450	$3435	1000 dollars
A	11045	26.093	16.746	0	60	No. of Years
L	11045	6.5101	1.8720	3.003	12.000	1000 ft^2
S	11045	2.6196	0.7507	1.000	4.793	1000 ft^2
NBA	11045	4.3928	0.9789	3	7	Number
NBE	11045	3.5283	1.2768	1	6	Number

It can be seen that detached houses in Richmond sold for a considerable amount of money over our sample period; i.e., they sold for an average of $1,113,410.

3. The Builder's Model

The *builder's model* for valuing a detached dwelling unit postulates that the value of the property is the sum of two components: the value of the land which the structure sits on plus the value of the structure.

In order to justify the model, consider a property developer who builds a structure on a particular property. The total cost of the property after the structure is completed will be equal to the floor space area of the structure, say S square feet, times the building cost per square foot β_t during the quarter, plus the cost of the land, which will be equal to the cost per square foot α_t during the quarter times the area of the land site, say L square feet. Now think of a sample of properties of the same general type, which have prices or values V_{tn} in quarter t[9] and structure areas S_{tn} and land areas L_{tn} for n = 1,...,N(t) where N(t) is the number of observations in period t. Assume that these prices are equal to the sum of the land and structure costs plus error terms ε_{tn} which we assume are independently normally distributed with zero means and constant variances. This leads to the following *hedonic regression model* for period t where the α_t and β_t are the parameters to be estimated in the regression:[10]

(1) $V_{tn} = \alpha_t L_{tn} + \beta_t S_{tn} + \varepsilon_{tn}$;   t = 1,...,36; n = 1,...,N(t).

The hedonic regression model defined by (1) applies to new structures. But it is likely that a model that is similar to (1) applies to older structures as well. Older structures will

---

[9] In the empirical work which follows, t will run from 1 to 36 where Quarter 1 is the first quarter of 2008 and Quarter 36 is the last quarter of 2016.

[10] Other papers that have suggested hedonic regression models that lead to additive decompositions of property values into land and structure components include Clapp (1980; 257-258), Bostic, Longhofer and Redfearn (2007; 184), Diewert (2008; 19-22) (2010) (2011), Francke and Vos (2004), Francke (2008; 167), Koev and Santos Silva (2008), de Haan and Diewert (2011), Rambaldi, McAllister, Collins and Fletcher (2010), Diewert, Haan and Hendriks (2011) (2015), Diewert and Shimizu (2015) (2016) (2016) and Burnett-Issacs, Huang and Diewert (2016).

be worth less than newer structures due to the depreciation of the structure. Assuming that we have information on the age of the structure n at time t, say A(t,n), and assuming a geometric (or declining balance) depreciation model, a more realistic hedonic regression model than that defined by (1) above is the following *basic builder's model*:

$$(2) \quad V_{tn} = \alpha_t L_{tn} + \beta_t(1 - \delta)^{A(t,n)}S_{tn} + \varepsilon_{tn} ; \qquad\qquad t = 1,...,36; \; n = 1,...,N(t)$$

where the parameter δ reflects the *net geometric depreciation rate* as the structure ages one additional period. Thus if the age of the structure is measured in years, we would expect an annual *net* depreciation rate to be around 2 to 3 percent per year.[11] Note that (2) is now a nonlinear regression model whereas (1) was a simple linear regression model. The period t constant quality price of land will be the estimated coefficient for the parameter α_t and the price of a unit of a newly built structure for period t will be the estimate for β_t. The period t quantity of land for property n is L_{tn} and the period t quantity of structure for property n, expressed in equivalent units of a new structure, is $(1 - \delta)^{A(t,n)}S_{tn}$ where S_{tn} is the floor space area of the structure for property n in period t.

Note that the above model is a *supply side model* as opposed to a *demand side model*.[12] Basically, we are assuming competitive suppliers of residential properties so that we are in Rosen's (1974; 44) Case (a), where the hedonic surface identifies the structure of supply. This assumption is justified for the case of newly built houses but it is less well justified for sales of properties with older structures where a demand side model may be more relevant.

There is a major practical problem with the hedonic regression model defined by (2): The multicollinearity problem. Experience has shown that it is usually not possible to estimate sensible land and structure prices in a hedonic regression like that defined by (2) due to the multicollinearity between lot size and structure size.[13] Thus in order to deal with the multicollinearity problem, we replace the parameter β_t in (2) by p_{St}, the period t Statistics Canada augmented construction cost price for houses in the Greater Vancouver area.[14] Thus **Model 1** is the following nonlinear hedonic regression model:

$$(3) \quad V_{tn} = \alpha_t L_{tn} + p_{St}(1 - \delta)^{A(t,n)}S_{tn} + \varepsilon_{tn} ; \qquad\qquad t = 1,...,36; \; n = 1,...,N(t)$$

This model has 36 quarterly land price parameters (the α_t) and one (net) geometric depreciation rate δ. The R^2 (between the observed values and the predicted values) was 0.7233 which is satisfactory for such a simple model. The estimated depreciation rate was

---

[11] This estimate of depreciation is regarded as a *net depreciation rate* because it is equal to a "true" gross structure depreciation rate less an average renovations appreciation rate. Since we do not have information on renovations and major repairs to a structure, our age variable will only pick up average gross depreciation less average real renovation expenditures.

[12] We will pursue a demand side model in Section 7 below.

[13] See Schwann (1998) and Diewert, de Haan and Hendriks (2011) (2015) on the multicollinearity problem.

[14] This formulation follows that of Diewert (2010) (2011), Diewert, Haan and Hendriks (2011) (2015), de Haan and Diewert (2011), Diewert and Shimizu (2015) (2016) (2017) and Burnett-Issacs, Huang and Diewert (2016) in assuming property value is the sum of land and structure components but movements in the price of structures are proportional to an exogenous structure price index.

2.85% per year. Land prices grew from $\alpha_1 = \$74.16$ per ft^2 in the first quarter of 2008 to $\alpha_{36} = \$214.00$ per ft^2 in the last quarter of 2016, a 2.89 fold increase over the sample period.[15]

In order to take into account possible neighbourhood effects on the price of land, we introduced (forward sortation) *postal code dummy variables*, $D_{PC,tn,j}$, into the hedonic regression (3). These 5 dummy variables are defined as follows: for $t = 1,...,36$; $n = 1,...,N(t)$; $j = 1,...,5$:[16]

(4) $D_{PC,tn,j} \equiv 1$ if observation n in period t is in Postal Code j of Richmond;
    $\equiv 0$ if observation n in period t is *not* in Postal Code j of Richmond.

We now modify the model defined by (3) to allow the *level* of land prices to differ across the 6 postal codes in Richmond. The new nonlinear regression model is the following one:

(5) $V_{tn} = \alpha_t(\sum_{j=1}^{6} \omega_j D_{PC,tn,j})L_{tn} + p_{St}(1 - \delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}$; $t = 1,...,36$; $n = 1,...,N(t)$.

Comparing the models defined by equations (3) and (5), it can be seen that we have added an additional 6 *neighbourhood relative land value parameters*, $\omega_1,...,\omega_6$, to the model defined by (3). However, looking at (5), it can be seen that the 36 land time parameters (the α_t) and the 6 location parameters (the ω_j) cannot all be identified. Thus we need to impose at least one identifying normalization on these parameters. We chose the following normalization:[17]

(6) $\omega_4 \equiv 1$.

Thus **Model 2** defined by equations (5) and (6) has 5 additional parameters compared to Model 1. Note that if we initially set all of the ω_j equal to unity, Model 2 collapses down to Model 1. We made use of this fact in running our sequence of nonlinear regressions. Our models are *nested* so that we can use the final parameter estimates from a previous model as starting parameter values in the next model's nonlinear regression.[18]

The final log likelihood (LL) for Model 2 was an improvement of 892.96 over the final LL for Model 1 (for adding 5 new neighbourhood parameters) which of course, is a highly significant increase. The R^2 increased to 0.7662 from the previous model R^2 of 0.7233. The new estimated depreciation rate turned out to be 0.0274 or 2.74% per year. The price of land increased 2.90 fold over the sample period.

---

[15] The Statistics Canada structure cost index increased 0.99 fold over the sample period.

[16] The number of observations over the sample period in the 6 postal code neighbourhoods are as follows: 791, 812, 2314, 3519, 3217 and 392..

[17] The fourth postal code region had the most observations over the sample period.

[18] In order to obtain sensible parameter estimates in our final (quite complex) nonlinear regression model, it is absolutely necessary to follow our procedure of sequentially estimating gradually more complex models, using the final coefficients from the previous model as starting values for the next model. We used Shazam to perform the nonlinear regressions; see White (2004).

In our next model, we introduce some nonlinearities into the pricing of the land area for each property. As mentioned above, the land plot areas in our sample of properties run from 3,000 to 12,000 ft^2. Up to this point, we have assumed that land plots in the same neighbourhood sell at a constant price per square foot of lot area. However, it is likely that there is some nonlinearity in this pricing schedule; for example, it is likely that large lots sell at an average price that is below the average price of medium sized lots. In order to capture this nonlinearity, we initially divided up our 11,045 observations into 9 groups of observations based on their lot size. The Group 1 properties had lots less than 4,000 ft^2, the Group 2 properties had lots greater than or equal to 4,000 ft^2 and less than 5,000 ft^2,..., the Group 8 properties had lots greater than or equal to 10,000 ft^2 and less than 11,000 ft^2 and the Group 9 properties had lots greater than or equal to 11,000 ft^2. However, there were very few observations in Groups 8 and 9 so we added these groups to Group 7.[19] For each observation n in period t, we define the 7 *land dummy variables*, $D_{L,tn,k}$, for k = 1,...,7 as follows:

(7) $D_{L,tn,k} \equiv 1$ if observation tn has land area that belongs to group k;
$\quad\quad \equiv 0$ if observation tn has land area that does not belong to group k.

These dummy variables are used in the definition of the following piecewise linear function of L_{tn}, $f_L(L_{tn})$, defined as follows:

(9) $f_L(L_{tn}) \equiv D_{L,tn,1}\lambda_1 L_{tn} + D_{L,tn,2}[\lambda_1 L_1 + \lambda_2(L_{tn}-L_1)] + D_{L,tn,3}[\lambda_1 L_1 + \lambda_2(L_2-L_1) + \lambda_3(L_{tn}-L_2)]$
$\quad\quad + ... + D_{L,tn,7}[\lambda_1 L_1 + \lambda_2(L_2-L_1) + ... + \lambda_6(L_6-L_5) + \lambda_7(L_{tn}-L_6)]$

where the λ_k are unknown parameters and $L_1 \equiv 4$, $L_2 \equiv 5$, $L_3 \equiv 6$, $L_4 \equiv 7$, $L_5 \equiv 8$ and $L_6 \equiv 9$. The function $f_L(L_{tn})$ defines a *relative valuation function for the land area of a house* as a function of the plot area.

The new nonlinear regression model is the following one:

(10) $V_{tn} = \alpha_t(\sum_{j=1}^{6} \omega_j D_{PC,tn,j})f_L(L_{tn}) + p_{St}(1 - \delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}$; $\quad\quad$ t = 1,...,36; n = 1,...,N(t).

Comparing the models defined by equations (5) and (10), it can be seen that we have added an additional 7 *land plot size parameters*, $\lambda_1,...,\lambda_7$, to the model defined by (5). However, looking at (10), it can be seen that the 36 land time parameters (the α_t), the 6 postal code parameters (the ω_j) and the 7 land plot size parameters (the λ_k) cannot all be identified. Thus we impose the following identification normalizations on the parameters for **Model 3** defined by (10) and (11):[20]

(11) $\omega_4 \equiv 1$; $\lambda_5 \equiv 1$.

---

[19] The sample probabilities of an observation falling in the 7 final groups were: 0.109, 0.164, 0.109, 0.136, 0.270, 0.126 and 0.086.
[20] The fifth land size group had the most observations over the sample period.

Note that if we set all of the λ_k equal to unity, Model 3 collapses down to Model 2. The final log likelihood for Model 3 was an improvement of 1761.68 over the final LL for Model 2 (for adding 6 new lot size parameters) which is a highly significant increase. The R^2 increased to 0.8283 from the previous model R^2 of 0.7662. The new estimated depreciation rate turned out to be 0.0405 or 4.05% per year. The price of land increased 2.69 fold over the sample period. Thus the new model has given rise to somewhat different land prices and the new depreciation rate is considerably higher than the depreciation rates in Models 1 and 2. The sequence of marginal land prices generated by this model are as follows: $\lambda_1 = 1.1350$, $\lambda_2 = 0.6915$, $\lambda_3 = 0.1198$, $\lambda_4 = 0.2963$, $\lambda_5 = 1.0000$ (this value was imposed), $\lambda_6 = 0.3871$ and $\lambda_7 = 0.0086$. Thus the marginal land prices as functions of lot size are not monotonic but for very large land plots, the marginal price of an extra square foot of land is very low.

Our next model is similar to Model 3 except we now allow the marginal price of adding an extra amount of structure to vary as the size of the structure increases. It is likely that the quality of the structure increases as the size of the structure increases. In order to capture this nonlinearity, we divided up our 11,045 observations into 5 groups of observations based on their structure size. The Group 1 properties had structures with floor space area less than 2,000 ft^2, the Group 2 properties had structure areas greater than or equal to 2,000 ft^2 and less than 2,500 ft^2, the Group 3 properties had structure areas greater than or equal to 2,500 ft^2 and less than 3,000 ft^2, the Group 4 properties had structure areas greater than or equal to 3,000 ft^2 and less than 3,500 ft^2 and the Group 5 properties had structure areas greater than or equal to 3,500 ft^2.[21] For each observation n in period t, we define the 5 *structure dummy variables*, $D_{S,tn,m}$, for $m = 1,...,5$ as follows:

(12) $D_{S,tn,m} \equiv 1$ if observation tn has structure area that belongs to structure group m;
$\equiv 0$ if observation tn has structure area that does not belong to group m.

These dummy variables are used in the definition of the following piecewise linear function of S_{tn}, $g_S(S_{tn})$, defined as follows:

(13) $g_S(S_{tn}) \equiv D_{S,tn,1}\mu_1 S_{tn} + D_{S,tn,2}[\mu_1 S_1 + \mu_2(S_{tn}-S_1)] + D_{S,tn,3}[\mu_1 S_1 + \mu_2(S_2-S_1) + \mu_3(S_{tn}-S_2)]$
$+ D_{S,tn,4}[\mu_1 S_1 + \mu_2(S_2-S_1) + \mu_3(S_3-S_2) + \mu_4(S_{tn}-S_3)]$
$+ D_{S,tn,5}[\mu_1 S_1 + \mu_2(S_2-S_1) + \mu_3(S_3-S_2) + \mu_4(S_4-S_3) + \mu_5(S_{tn}-S_4)]$.

where the μ_m are unknown parameters and $S_1 \equiv 2$, $S_2 \equiv 2.5$, $S_3 \equiv 3$ and $S_4 \equiv 3.5$. The function $g_S(S_{tn})$ defines a *relative valuation function for the structure area of a house* as a function of the structure area.

The new nonlinear regression model is the following **Model 4**:

(14) $V_{tn} = \alpha_t(\sum_{j=1}^{6} \omega_j D_{PC,tn,j}) f_L(L_{tn}) + p_{St}(1-\delta)^{A(t,n)} g_S(S_{tn}) + \varepsilon_{tn}$; $t = 1,...,36$; $n = 1,...,N(t)$.

---

[21] The sample probabilities of an observation falling in the 5 structure groups were: 0.207, 0.308, 0.191, 0.134 and 0.160.

Comparing the models defined by equations (10) and (14), it can be seen that we have added an additional 5 *structure floor space parameters*, $\mu_1,...,\mu_5$, to the model defined by (10).[22] Again, we add the normalizations (11) in order to identify all of the parameters in the model. There are a total of 53 unknown parameters in Model 4. Note that if we set all of the μ_m equal to unity, Model 4 collapses down to Model 3.

The final log likelihood for Model 4 was an improvement of 934.91 over the final LL for Model 3 (for adding 5 new structure size parameters) which is a highly significant increase. The R^2 increased to 0.8520 from the previous model R^2 of 0.8283. The new estimated depreciation rate turned out to be 0.0298 or 2.98% per year. Using this model, the price of land increased 3.34 fold over the sample period. The sequence of marginal land prices generated by this model are as follows: $\lambda_1 = 1.3042$, $\lambda_2 = 0.7729$, $\lambda_3 = 0.2613$, $\lambda_4 = 0.3536$, $\lambda_5 = 1.0000$ (this value was imposed), $\lambda_6 = 0.6410$ and $\lambda_7 = 0.1268$. These marginal land prices are much larger than the corresponding prices in the previous model and they appear to be more reasonable. The marginal structure floor space prices estimated by Model 4 were as follows: $\mu_1 = 1.4318$, $\mu_2 = 1.1697$, $\mu_3 = 1.5871$, $\mu_4 = 2.4859$, $\mu_5 = 0.6222$. Thus there is a considerable amount of variability in both the marginal land and structure price parameters.

In Model 5, we allowed the geometric depreciation rates to differ after each 10 year interval. We divided up our 11,045 observations into 6 groups of observations based on the age of the structure at the time of the sale. The Group 1 properties had structures with structure age less than 10 years, the Group 2 properties had structure ages greater than or equal to 10 years but less than 20 years, ..., the Group 6 properties had structure ages greater than or equal to 50 years.[23] For each observation n in period t, we define the 6 *age dummy variables*, $D_{A,tn,i}$, for i = 1,...,6 as follows:

(15) $D_{A,tn,i} \equiv 1$ if observation tn has structure age that belongs to age group i;
$\equiv 0$ if observation tn has structure age that does not belong to age group i.

These age dummy variables are used in the definition of the following function of the age variable, $g_A(A_{tn})$, defined as follows:[24]

(16) $g_A(A_{tn}) \equiv D_{A,tn,1}(1-\delta_1)^{A(t,n)} + D_{A,tn,2}(1-\delta_1)^{10}(1-\delta_2)^{A(t,n)-10}$
$+ D_{A,tn,3}(1-\delta_1)^{10}(1-\delta_2)^{10}(1-\delta_3)^{A(t,n)-20} + D_{A,tn,4}(1-\delta_1)^{10}(1-\delta_2)^{10}(1-\delta_3)^{10}(1-\delta_4)^{A(t,n)-30}$
$+ D_{A,tn,5}(1-\delta_1)^{10}...(1-\delta_4)^{10}(1-\delta_5)^{A(t,n)-40} + D_{A,tn,6}(1-\delta_1)^{10}...(1-\delta_5)^{10}(1-\delta_6)^{A(t,n)-50}$.

Thus the annual geometric depreciation rates are allowed to change at the end of each decade that the structure survives.

---

[22] Note that we did not normalize one of the parameters μ_1-μ_5. This lack of a normalization means that we are allowing our estimated *levels* of structure imputed values to be totally independent of the earlier imputed structure values which were tied to the Statistics Canada estimated augmented construction cost index.

[23] The sample probabilities of an observation falling in the 6 age groups were: 0.224, 0.146, 0.164, 0.225, 0.149 and 0.093.

[24] A_{tn} is the same as A(t,n).

The new **Model 5** nonlinear regression model is the following one:

$$(17) \quad V_{tn} = \alpha_t \left(\sum_{j=1}^{6} \omega_j D_{PC,tn,j} \right) f_L(L_{tn}) + p_{St} g_S(S_{tn}) g_A(A_{tn}) + \varepsilon_{tn} ; \qquad t = 1,...,36; \, n = 1,...,N(t).$$

Comparing the models defined by equations (14) and (17), it can be seen that we now have 6 *depreciation rates*, $\delta_1,...,\delta_6$, in place of the single depreciation rate δ which appeared in Model 4. Again, we add the normalizations (11) in order to identify all of the parameters in the model. There are a total of 58 unknown parameters in Model 5. Note that if we initially set all of the δ_i equal to the δ which appeared in Model 4, then Model 5 collapses down to Model 4.

The final log likelihood for Model 5 was a modest improvement of 24.36 over the final LL for Model 4 (for adding 5 new depreciation rate parameters). The R^2 increased to 0.8526 from the previous model R^2 of 0.8520. Using this model, the price of land increased 3.34 fold over the sample period, the same increase as occurred in Model 4. The new decade by decade estimated depreciation rates turned out to be as follows: $\delta_1 = 0.0351$, $\delta_2 = 0.0229$, $\delta_3 = 0.0267$,[25] $\delta_4 = 0.0462$, $\delta_5 = 0.0220$ and $\delta_6 = -0.0239$. Thus properties with structures which are over 50 years old tend to have a negative depreciation rate; i.e., the value of the structure tends to *increase* by 2.39% per year.[26] The marginal land and structure prices did not change much from the estimates from Model 4. The conclusion we can draw from Models 4 and 5 is that there is some evidence that the (net) geometric depreciation rate for detached houses in Richmond is around 2 to 3 percent per year for the first 30 years of the structure life and then for the next 10 years the rate increases to something around 4.6% per year. The rate then decreases to around 2.2% per year for houses of age 40-50 and finally, the depreciation rate for houses that survive more than 50 years becomes negative with an appreciation rate of approximately 2.4% per year. However, the improvement in the descriptive power of this more general depreciation model is fairly limited; a single geometric depreciation rate of around 3% per year describes the data almost as well.

For our final two models, we introduce bathroom and bedroom variables into the hedonic regressions. The number of bathrooms in our sample ranged from 1 to 6 bathrooms. We assume that the number of bathrooms in the structure affects the quality of the structure. Thus define the following 6 dummy variables, $D_{Ba,tn,i}$: for $t = 1,...,36$; $n = 1,...,N(t)$; $i = 1,...,6$:[27]

$(18) \quad D_{Ba,tn,i} \equiv 1$ if observation n in period t is a house with i bathrooms;
$\qquad \equiv 0$ if observation n in period t does *not* have i bathrooms.

---

[25] Remember that these depreciation rates are net depreciation rates. As houses approach their middle age, renovations become important and thus a decline in the net depreciation rate is plausible.

[26] This phenomenon has been observed in the literature before; i.e., older heritage houses that have been extensively renovated may increase in value over time rather than depreciate as they age.

[27] The sample probabilities of an observation falling in the 6 number of bathroom groups were: 0.029, 0.151, 0.423, 0.165, 0.122 and 0.110. Thus 3 bathroom houses were by far the most common type of house and so we set $\eta_3 = 1$ when normalizing one of the bathroom coefficients in the restrictions (21).

We use the bathroom dummy variables defined above in order to define the following bathroom quality adjustment function, $g_{Ba}(NBA_{tn})$

(19) $g_{Ba}(NBA_{tn}) \equiv (\Sigma_{i=1}^{6} \eta_i D_{Ba,tn,i}) NBA_{tn}$.

The new **Model 6** nonlinear regression model is the following one:

(20) $V_{tn} = \alpha_t (\Sigma_{j=1}^{6} \omega_j D_{PC,tn,j}) f_L(L_{tn}) + p_{St} g_S(S_{tn}) g_A(A_{tn}) g_{Ba}(NBA_{tn}) + \varepsilon_{tn}$;
$$t = 1,...,36; \ n = 1,...,N(t).$$

Comparing the models defined by equations (17) and (20), it can be seen that we now have added 6 new bathroom parameters, $\eta_1,...,\eta_6$, to Model 5. Not all of these new parameters can be identified so we impose the identifying normalizations (21) on the parameters in (20):

(21) $\omega_4 \equiv 1; \ \lambda_5 \equiv 1; \ \eta_3 \equiv 1$.

There are a total of 63 unknown parameters in Model 6. Note that if we initially set all of the η_i equal to 1, then Model 6 collapses down to Model 5.

The final log likelihood for Model 6 was an improvement of 38.08 over the final LL for Model 5 (for adding 5 new bathroom parameters). The R^2 increased to 0.8536 from the previous model R^2 of 0.8526. Using this model, the price of land again increased 3.34 fold over the sample period, the same increase as occurred in Models 4 and 5. The new bedroom parameters turned out to be as follows: $\eta_1 = 1.1021$, $\eta_2 = 0.9368$, $\eta_3 = 1.000$ (imposed), $\eta_4 = 0.9526$, $\eta_5 = 0.8946$ and $\eta_6 = 0.9631$. These estimates are reasonable. The new decade by decade estimated depreciation rates turned out to be as follows: $\delta_1 = 0.0339$, $\delta_2 = 0.0242$, $\delta_3 = 0.0304$, $\delta_4 = 0.0445$, $\delta_5 = 0.0202$ and $\delta_6 = -0.0237$. These rates are similar to the depreciation rates generated by the previous model. The marginal land and structure prices did not change much from the previous estimates from Models 4 and 5.

Finally, we introduce the number of bedrooms into the hedonic regressions. The number of bedrooms in our sample ranged from 3 to 7 bedrooms. We assume that the number of bedrooms in the structure affects the quality of the structure. Thus define the following 5 dummy variables, $D_{Be,tn,i}$, for $t = 1,...,36$; $n = 1,...,N(t)$; $i = 3,4,...,7$:[28]

(22) $D_{Be,tn,i} \equiv 1$ if observation n in period t is a house with i bedrooms;
$\equiv 0$ if observation n in period t does *not* have i bedrooms.

We use the bedroom dummy variables defined above in order to define the following bedroom quality adjustment function, $g_{Be}(NBE_{tn})$:

---

[28] The sample probabilities of an observation falling in the 5 number of bedroom groups were: 0.210, 0.318, 0.358, 0.096, 0.018. Thus 5 bedroom houses were the most common type of house and so we set $\theta_5 = 1$ when normalizing one of the bedroom coefficients in the restrictions (25).

(23) $g_{Be}(NBE_{tn}) \equiv (\Sigma_{i=3}^{7} \theta_i D_{Ba,tn,i})NBE_{tn}$.

The new **Model 7** nonlinear regression model is the following one:

(24) $V_{tn} = \alpha_t(\Sigma_{j=1}^{6} \omega_j D_{PC,tn,j})f_L(L_{tn}) + p_{St}g_S(S_{tn})g_A(A_{tn})g_{Ba}(NBA_{tn})g_{Be}(NBE_{tn}) + \varepsilon_{tn}$;
$$t = 1,...,36; n = 1,...,N(t).$$

Comparing the models defined by equations (20) and (24), it can be seen that we now have added 5 new bedroom parameters, $\theta_3,...,\theta_7$, to Model 6. Not all of these new parameters can be identified so we impose the identifying normalizations (25) on the parameters in (24):

(25) $\omega_4 \equiv 1; \lambda_5 \equiv 1; \eta_3 \equiv 1; \theta_5 = 1$.

There are a total of 67 unknown parameters in Model 7. Note that if we set all of the θ_i equal to 1, then Model 7 collapses down to Model 6.

The final log likelihood for Model 7 was a big improvement of 180.70 over the final LL for Model 6 (for adding 4 new bathroom parameters). The R^2 increased to 0.8583 from the previous model R^2 of 0.8536. Using this model, the price of land increased 3.50 fold which is an increase over the 3.34 fold increase that occurred in Models 4-6. The estimated coefficients for Model 7 are listed below in Table 2.[29] The new decade by decade estimated depreciation rates turned out to be as follows (the corresponding rates from Model 6 are listed in brackets): $\delta_1 = 0.0298$ (0.0339), $\delta_2 = 0.0202$ (0.0242), $\delta_3 = 0.0301$ (0.0304), $\delta_4 = 0.0345$ (0.0445), $\delta_5 = 0.0132$ (0.0202) and $\delta_6 = -0.0180$ (−0.0237). These new depreciation rates are broadly similar to the depreciation rates generated by the previous two models but there are some substantial differences. The problem is that the standard errors on these parameter estimates are fairly large and hence the depreciation parameters are not estimated with great precision. The marginal land and structure prices did not change much from the previous estimates from Models 4 and 5.

Table 2: Estimated Coefficients for Model 7

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
α_1	62.30	16.68	α_{24}	117.27	18.46	λ_7	0.1274	2.52
α_2	63.56	16.29	α_{25}	122.10	18.49	μ_1	1.4427	27.12
α_3	59.64	14.04	α_{26}	128.08	18.29	μ_2	1.3993	7.87
α_4	50.48	11.75	α_{27}	130.00	18.38	μ_3	1.5694	8.61
α_5	52.37	15.49	α_{28}	130.99	18.63	μ_4	2.3765	11.91
α_6	63.78	17.16	α_{29}	142.46	18.74	μ_5	0.8117	6.65
α_7	68.54	17.38	α_{30}	153.56	18.71	δ_1	0.0298	19.38
α_8	75.25	17.61	α_{31}	169.88	18.79	δ_2	0.0202	10.15
α_9	88.73	18.13	α_{32}	182.24	18.80	δ_3	0.0301	10.22
α_{10}	92.02	18.26	α_{33}	229.69	18.79	δ_4	0.0345	6.90

---

[29] Standard errors for each estimated coefficient can be obtained by dividing the estimate by the corresponding listed t statistic.

α_{11}	90.42	18.13	α_{34}	245.21	18.80	δ_5	0.0132	1.97
α_{12}	94.19	18.19	α_{35}	236.03	18.65	δ_6	−0.0180	−1.27
α_{13}	125.40	18.63	α_{36}	217.87	18.10	η_1	1.1223	11.61
α_{14}	122.37	18.64	ω_1	0.6392	69.92	η_2	0.9473	26.11
α_{15}	116.76	18.43	ω_2	0.9880	129.47	η_4	1.0077	62.35
α_{16}	118.47	18.34	ω_3	0.8643	169.34	η_5	0.9837	52.81
α_{17}	126.91	18.67	ω_5	0.9334	185.06	η_6	1.0572	47.39
α_{18}	120.76	18.32	ω_6	0.5508	41.20	θ_3	1.1048	40.35
α_{19}	114.56	18.13	λ_1	1.2651	17.79	θ_4	1.0835	88.41
α_{20}	107.87	17.34	λ_2	0.7524	8.47	θ_6	0.8603	72.92
α_{21}	110.06	17.74	λ_3	0.3241	4.19	θ_7	0.7334	30.49
α_{22}	113.78	18.32	λ_4	0.3237	4.76			
α_{23}	114.19	18.32	λ_6	0.6115	6.51			

In the following section, we will use the results from Models 4-7 to construct overall property price indexes and the corresponding land price indexes in order to determine how similar they are.

4. Sales Price Indexes for the Geometric Depreciation Models

Recall equations (14) in the previous section which regressed the sale price of property n in quarter t, V_{tn}, on terms associated with the land and structure values of the property. Once the parameters for Model 4 have been determined, the *predicted constant quality amount of land* for property n sold during period t, L_{tn}^{*}, can be defined as follows:

$$(26)\ L_{tn}^{*} \equiv (\textstyle\sum_{j=1}^{6} \omega_j D_{PC,tn,j}) f_L(L_{tn}) ; \qquad\qquad t = 1,...,36; n = 1,...,N(t).$$

The corresponding quarter t *constant quality land price* is the estimated coefficient, α_t, for t = 1,...,36. Similarly, the *predicted constant quality amount of structure* for property n sold during period t, S_{tn}^{*}, can be defined as follows:

$$(27)\ S_{tn}^{*} \equiv (1 - \delta)^{A(t,n)} g_S(S_{tn}) ; \qquad\qquad t = 1,...,36; n = 1,...,N(t).$$

The corresponding quarter t *constant quality structure price* is the quarter t Statistics Canada price index p_{St} for t = 1,...,36.

In order to form an overall quarter t property price index, it is necessary to aggregate the individual property amounts of constant quality land and structure into quarterly aggregates, say L_t^{*} and S_t^{*}. This task can be accomplished by simple addition. The corresponding quarter t aggregate prices will be α_t and p_{St} respectively. It is convenient to rescale these price indexes so that they equal 1 in quarter 1. Denote these normalized price indexes by P_{Lt} and P_{St} for quarter t. Normalizing the price indexes means that the corresponding quarter t quantity aggregates need to be normalized in the opposite direction. Thus for Model 4, we end up with the following definitions for the *price and quantity of aggregate constant quality land and structures* for each quarter t = 1,...,36:

$$(28)\ P_{4Lt} \equiv \alpha_t/\alpha_1 ;$$
$$(29)\ L_{4t}^{*} \equiv \alpha_1 \textstyle\sum_{n=1}^{N(t)} (\textstyle\sum_{j=1}^{6} \omega_j D_{PC,tn,j}) f_L(L_{tn}) ;$$

(30) $P_{4St} \equiv p_{St}/p_{S1}$;

(31) $S_{4t}^{*} \equiv p_{S1}\Sigma_{n=1}^{N(t)} (1 - \delta)^{A(t,n)} g_S(S_{tn})$.

The aggregate quarterly land and structure price indexes generated by Model 4, P_{4Lt} and P_{4St}, are listed in Table 4 below. Equations (28)-(31) generate aggregate land and structure prices and quantities for the 36 quarters in our sample. We used these land and structure aggregates to form Fisher (1922) chained property price indexes. These Model 4 *aggregate property price indexes* are listed as P_{4t} in Table 3.

A similar procedure can be used in order to construct land and structure subindexes for Models 5, 6 and 7, P_{5Lt}- P_{7Lt} and P_{5St}- P_{7St}, along with the corresponding overall property price indexes, which we denote by P_{5t}, P_{6t} and P_{7t} in Table 3. Note that definitions (28) and (30)[30] are used to define the land and structure price indexes for each model but the counterparts to definitions (31) will differ for each model. For example, the counterpart definitions to (31) for Model 7 are the following definitions for t = 1,...,36:

(32) $S_{7t}^{*} \equiv p_{S1}\Sigma_{n=1}^{N(t)} g_S(S_{tn})g_A(A_{tn})g_{Ba}(NBA_{tn})g_{Be}(NBE_{tn})$.

Table 3: Property Price Indexes P_4-P_7, Land Price Indexes P_{4L}-P_{7L}, Structure Price Index P_S and Mean and Median Price Indexes for Sales of Properties in Richmond Q1 2008 to Q4 2016

t	P_{4t}	P_{5t}	P_{6t}	P_{7t}	P_{4Lt}	P_{5Lt}	P_{6Lt}	P_{7Lt}	P_{St}	P_{Mean}	P_{Median}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0082	1.0086	1.0052	1.0103	1.0146	1.0154	1.0086	1.0203	1.0020	1.0360	1.0280
3	0.9867	0.9874	0.9854	0.9805	0.9737	0.9750	0.9708	0.9574	0.9990	1.0100	0.9800
4	0.8934	0.8933	0.8970	0.8945	0.8203	0.8195	0.8267	0.8103	0.9610	0.9870	0.9430
5	0.8751	0.8750	0.8747	0.8729	0.8507	0.8503	0.8490	0.8407	0.9006	0.9870	0.9620
6	0.9283	0.9258	0.9254	0.9236	1.0199	1.0157	1.0153	1.0238	0.8509	0.9550	0.9520
7	0.9751	0.9734	0.9707	0.9703	1.0921	1.0900	1.0852	1.1002	0.8743	1.0340	0.9960
8	1.0334	1.0339	1.0316	1.0357	1.1792	1.1819	1.1782	1.2080	0.9064	1.1080	1.0780
9	1.1499	1.1489	1.1478	1.1468	1.3913	1.3919	1.3918	1.4243	0.9318	1.2140	1.1620
10	1.1797	1.1787	1.1773	1.1794	1.4356	1.4365	1.4357	1.4772	0.9474	1.1760	1.1770
11	1.1671	1.1668	1.1669	1.1649	1.4153	1.4174	1.4197	1.4514	0.9425	1.2030	1.1700
12	1.1929	1.1914	1.1917	1.1892	1.4735	1.4735	1.4765	1.5120	0.9357	1.2680	1.2450
13	1.4356	1.4336	1.4315	1.4266	1.9441	1.9455	1.9458	2.0130	0.9367	1.4330	1.3900
14	1.4093	1.4101	1.4090	1.4059	1.8893	1.8959	1.8975	1.9642	0.9425	1.3800	1.3480
15	1.3675	1.3663	1.3657	1.3625	1.8095	1.8118	1.8141	1.8743	0.9406	1.4090	1.3870
16	1.3813	1.3810	1.3785	1.3721	1.8423	1.8465	1.8454	1.9018	0.9337	1.3630	1.3450
17	1.4470	1.4447	1.4415	1.4350	1.9714	1.9722	1.9706	2.0372	0.9298	1.4160	1.3800
18	1.4001	1.3970	1.3955	1.3880	1.8811	1.8801	1.8810	1.9384	0.9279	1.3940	1.3160
19	1.3553	1.3530	1.3483	1.3431	1.7904	1.7904	1.7843	1.8390	0.9279	1.4310	1.3330
20	1.3019	1.2987	1.3003	1.2930	1.6849	1.6823	1.6890	1.7315	0.9240	1.3790	1.3900
21	1.3183	1.3164	1.3135	1.3057	1.7242	1.7246	1.7222	1.7668	0.9181	1.3390	1.2620
22	1.3441	1.3438	1.3399	1.3307	1.7798	1.7836	1.7796	1.8264	0.9142	1.3870	1.3090
23	1.3457	1.3446	1.3405	1.3321	1.7860	1.7883	1.7840	1.8330	0.9113	1.4130	1.3210
24	1.3584	1.3574	1.3569	1.3501	1.8198	1.8224	1.8257	1.8824	0.9045	1.4680	1.4470
25	1.3913	1.3910	1.3883	1.3830	1.8904	1.8949	1.8940	1.9599	0.9016	1.4870	1.4580
26	1.4351	1.4345	1.4329	1.4228	1.9856	1.9898	1.9917	2.0559	0.8938	1.4390	1.3900
27	1.4485	1.4468	1.4449	1.4359	2.0142	2.0167	2.0179	2.0867	0.8918	1.5340	1.5310
28	1.4607	1.4586	1.4559	1.4458	2.0339	2.0355	2.0353	2.1027	0.8967	1.5420	1.4930

---

[30] The structure price indexes P_{4St}-P_{7St} are all equal and are listed as P_{St} in Table 3.

29	1.5441	1.5414	1.5390	1.5290	2.2024	2.2036	2.2048	2.2868	0.8947	1.6420	1.5630
30	1.6302	1.6279	1.6245	1.6169	2.3631	2.3657	2.3657	2.4650	0.9064	1.6820	1.6740
31	1.7575	1.7550	1.7543	1.7439	2.6046	2.6082	2.6147	2.7270	0.9172	1.8290	1.8440
32	1.8580	1.8554	1.8524	1.8420	2.7921	2.7963	2.7988	2.9253	0.9289	1.9550	1.9360
33	2.2296	2.2277	2.2248	2.2110	3.4966	3.5053	3.5110	3.6870	0.9493	2.2800	2.2700
34	2.3628	2.3623	2.3563	2.3401	3.7348	3.7470	3.7477	3.9361	0.9717	2.3800	2.4110
35	2.3075	2.3048	2.2992	2.2802	3.6106	3.6181	3.6186	3.7888	0.9883	2.3320	2.3490
36	2.1694	2.1651	2.1595	2.1449	3.3394	3.3422	3.3412	3.4972	0.9873	2.1470	2.1210

It can be seen that there is little difference in the land and overall property price indexes generated by Models 4-6. For Model 7, the land price index P_{7L} ends up 4.7% higher than the Model 6 land price index P_{6L} in Quarter 36 but the difference between the overall indexes P_6 and P_7 is less than 1% in Quarter 36. For comparison purposes, the Mean and Median price indexes (normalizations of the mean and median selling price of houses in each quarter) are also listed in Table 3. These indexes generally capture the trend of the hedonic property price indexes but they are much more volatile than the hedonic indexes.

Our conclusion here is that it appears that Model 4, which has only a single geometric depreciation rate and does not make use of the bathroom and bedroom variables, generates land and overall property price indexes which adequately approximate our subsequent hedonic regression Models which require additional information on housing characteristics. This is an important result for statistical agencies in that it is typically difficult to get information on housing characteristics. Thus information on property location, the floor space area of the structure, the age of the structure and the lot size can be sufficient to generate price indexes that are reasonably accurate.

The price indexes listed in Table 3 pertain to the *sales* of houses in Richmond. In the following section, we attempt to estimate price indexes that apply to the *stock* of houses in Richmond.

5. Approximate Price Indexes for Housing Stocks

House price indexes are used for different purposes. For many macroeconomic purposes, price indexes for the *sales* of houses are not so important; what is more important is the construction of price indexes for the *stock* of houses.[31] In this section, we show how our sales price indexes could be used to value residential housing stocks.

The first step is to form estimates for the total stock of residential housing in an area. In this section, we will show how a very rough approximation to the total stock of housing in Richmond can be formed. Once we have estimates for the total stock of constant quality land and structures at the end of our sample period, we can apply our quarterly price indexes to these stock estimates and get approximate total property price indexes for Richmond over our sample period.

---

[31] Price and quantity information on the land and structure components of the stock of rental housing is important for the measurement of the productivity of the rental residential property industry. Information on the price of the stock of Owner Occupied Housing is important for the construction of a Consumer Price Index and for the valuation of the real consumption services that can be attributed to OOH.

Recall definitions (29) and (31) in the previous section which defined total constant quality land and structures sold in quarter t for Model 4, L_{4t}^* and S_{4t}^* respectively. We aggregate these quarterly quantities into annual quantities sold (by simple addition)[32] and we obtain 9 annual constant quality amounts of land and structures sold, L_{4y}^{**} and S_{4y}^{**}, for years y = 1,2,...,9.[33] A very rough approximation to the total stock of quality adjusted land that exists in Richmond at the end of our sample period can be obtained by adding up these annual amounts of quality adjusted land that sold over the 9 years in our sample; i.e., define L_4^{**} as follows:

$$(32)\ L_4^{**} \equiv \Sigma_{y=1}^9\ L_{4y}^{**}.$$

The same procedure could be applied to our annual estimates for quality adjusted structures that were sold during the 9 years but simple addition would overstate the stock at the end of the sample period because the properties sold at the beginning of the sample period would encounter structure depreciation for 9 years. Thus the amount of quality adjusted structure that sold over the 9 years, adjusted for annual depreciation, S_4^{**}, is defined as follows:

$$(33)\ S_4^{**} \equiv S_{4,9}^{**} + (1-\delta)S_{4,8}^{**} + (1-\delta)^2 S_{4,7}^{**} + ... + (1-\delta)^8 S_{4,1}^{**}.$$

The approximate stocks of quality adjusted land and structures defined by (32) and (33) can be used to form the following Model 4 *approximate price index for the Richmond housing stock* for quarter t, P_{Lowe4t}:[34]

$$(34)\ P_{Lowe4t} \equiv [P_{4Lt}L_4^{**} + P_{St}S_4^{**}]/[P_{4L1}L_4^{**} + P_{S1}S_4^{**}]\ ; \qquad\qquad t = 1,...,36$$

where the land and structure price indexes P_{4Lt} and P_{St} are listed above in Table 3. The overall property price index defined by (34) is called a Lowe (1823) Index since it is a basket type index that uses constant quantity weights throughout the sample period.

The same procedure can be followed using the results of Models 5, 6 and 7 in order to construct the approximate housing stock indexes, P_{Lowe5t}, P_{Lowe7t} and P_{Lowe7t}. Of course, the quarterly L_{4t}^* defined by (29) and S_{4t}^* defined by (31) need to be replaced by their Model 5, 6 and 7 counterparts in carrying out the above approximations to the constant quality stocks of land and structures.[35] The Lowe indexes for Models 4-7 are listed in Table 4.

---

[32] If we had information on the age of structures by their age in quarters, it would not be necessary to do this aggregation. But because our information on the age of the structure is in years and because we estimate annual depreciation rates, we aggregated the quarterly data into annual data.

[33] The annual ratios of L_{4y}^{**} to S_{4y}^{**} were 0.968, 0.918, 1.028, 1.165, 1.018, 0.958, 1.001, 1.035 and 1.146 while the corresponding ratios of L_{7y}^{**} to S_{7y}^{**} were 0.826, 0.785, 0.873, 0.974, 0.851, 0.805, 0.841, 0.869 and 0.947. Thus the alternative models do change the estimated land to structure ratios.

[34] The ratio of L_4^{**} to S_4^{**} was 1.146 while the corresponding ratio of L_7^{**} to S_7^{**} was 0.965.

[35] Models 5, 6 and 7 had multiple depreciation rates rather than the single depreciation rate that was generated by Model 4. Hence the aggregation equation (33) is no longer valid for Models 5-7. However, we continued to use equation (33) to aggregate the annual stocks into the end of period constant quality stock of structures using the single depreciation rate $\delta \equiv 0.03$.

Table 4: Approximate Property Price Indexes P$_{Lowe4}$-P$_{Lowe7}$ for the Housing Stock in Richmond Q1 2008 to Q4 2016

t	P$_{Lowe4t}$	P$_{Lowe5t}$	P$_{Lowe6t}$	P$_{Lowe7t}$
1	1.0000	1.0000	1.0000	1.0000
2	1.0087	1.0091	1.0055	1.0110
3	0.9855	0.9862	0.9841	0.9786
4	0.8858	0.8857	0.8899	0.8870
5	0.8739	0.8738	0.8733	0.8712
6	0.9412	0.9385	0.9379	0.9358
7	0.9906	0.9890	0.9859	0.9852
8	1.0521	1.0529	1.0502	1.0545
9	1.1772	1.1764	1.1751	1.1736
10	1.2081	1.2075	1.2057	1.2075
11	1.1950	1.1950	1.1950	1.1924
12	1.2229	1.2216	1.2218	1.2187
13	1.4747	1.4731	1.4705	1.4652
14	1.4481	1.4494	1.4478	1.4443
15	1.4046	1.4038	1.4027	1.3991
16	1.4190	1.4191	1.4160	1.4091
17	1.4860	1.4841	1.4804	1.4737
18	1.4370	1.4342	1.4322	1.4241
19	1.3885	1.3865	1.3810	1.3753
20	1.3303	1.3272	1.3287	1.3206
21	1.3486	1.3469	1.3435	1.3349
22	1.3765	1.3765	1.3721	1.3622
23	1.3784	1.3776	1.3730	1.3640
24	1.3933	1.3926	1.3918	1.3847
25	1.4297	1.4297	1.4266	1.4213
26	1.4768	1.4766	1.4747	1.4645
27	1.4912	1.4899	1.4876	1.4786
28	1.5040	1.5023	1.4991	1.4889
29	1.5931	1.5907	1.5878	1.5784
30	1.6843	1.6824	1.6785	1.6719
31	1.8183	1.8163	1.8152	1.8060
32	1.9239	1.9218	1.9181	1.9093
33	2.3097	2.3084	2.3046	2.2938
34	2.4473	2.4475	2.4404	2.4275
35	2.3887	2.3867	2.3799	2.3636
36	2.2434	2.2395	2.2327	2.2199

From Table 4, it can be seen that the Model 4-7 Lowe indexes approximate each other quite closely. It can also be seen that the Lowe indexes P$_{Lowe4t}$-P$_{Lowe7t}$ are approximately 3.5% above their sales property price counterparts, P$_{4t}$-P$_{7t}$, at the end of the sample period when t = 36. This is to be expected: our approximate stock price indexes have a higher land to structure ratio because the stock indexes better reflect the composition of the structure stock by depreciating the structures for the early years in our sample of sold properties. A more accurate measure of the Richmond housing stock would produce an even higher property inflation rate for the stock measure.

The form of depreciation that we used in Models 5-7 gave rise to rather complex nonlinear regression models which could lead to a lack of convergence in applications to other data sets. Thus in the following section, we replace our geometric depreciation

models with piecewise linear depreciation models which are not as nonlinear. These alternative models also provide a robustness check on our earlier computations.

6. Sales Price Indexes for Piecewise Linear Depreciation Models

In this section, we re-estimate Models 1-7 which were explained in Section 3. The geometric depreciation rates which were used in Section 3 are replaced by either straight line depreciation for Models 1-4 and by piecewise linear depreciation for Models 5-7.

Recall that Models 1-4 were defined by equations (3), (5), (10) and (14) in Section 3.[36] For the new Models 1-4, replace the term $(1 - \delta)^{A(t,n)}$ which appears in these equations by the term $(1 - \delta A_{tn})$ so that δ is now interpreted as the straight line depreciation rate which was previously a geometric or declining balance depreciation rate.

Recall that Models 5-7 in Section 3 were defined by equations (17), (20) and (24). The same depreciation rate function $g_A(A_{tn})$ defined by (16) was used in each of these geometric depreciation models. For the piecewise linear depreciation Models 5-7 that are estimated in this section, replace the old depreciation function $g_A(A_{tn})$ by the new depreciation function defined as follows (where A_{tn} is the age of the structure on property n sold during period t and the age dummy variables $D_{A,tn,i}$ are defined by (15)):

$$(35)\ g_A(A_{tn}) \equiv D_{A,tn,1}(1 - \delta_1 A_{tn}) + D_{A,tn,2}(1 - 10\delta_1 - \delta_2(A_{tn} - 10))$$
$$+ D_{A,tn,3}(1 - 10\delta_1 - 10\delta_2 - \delta_3(A_{tn} - 20)) + D_{A,tn,4}(1 - 10\delta_1 - 10\delta_2 - 10\delta_3 - \delta_4(A_{tn} - 30))$$
$$+ D_{A,tn,5}(1 - 10\delta_1 - 10\delta_2 - 10\delta_3 - 10\delta_4 - \delta_5(A_{tn} - 40))$$
$$+ D_{A,tn,6}(1 - 10\delta_1 - 10\delta_2 - 10\delta_3 - 10\delta_4 - 10\delta_5 - \delta_6(A_{tn} - 50)).$$

The remaining parts of Models 1-7 remain unchanged. In Table 5 below, we list the R^2 and improvement in the log likelihood (ΔLL) for each new model over the previous model. For Models 1-4, we list the estimated straight line depreciation rate δ_1 for each model and for Models 5-7, we list the 6 decade by decade linear depreciation rates δ_1-δ_6 that were estimated by these models. These depreciation rates of course are somewhat different from our previously estimated geometric depreciation rates. However, for both the geometric depreciation Model 7 and the piecewise linear depreciation Model 7, the depreciation rate δ_6 becomes an appreciation rate for structures in the 50-60 year old range.

Table 5: Results for the Piecewise Linear Depreciation Models 1-7

	R^2	ΔLL	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6
Model 1	0.7494		0.0247					
Model 2	0.7914	955.60	0.0242					
Model 3	0.8429	1861.43	0.0279					
Model 4	0.8476	170.45	0.0210					
Model 5	0.8525	181.44	0.0301	0.0140	0.0141	0.0150	0.0052	−0.0054
Model 6	0.8535	38.43	0.0292	0.0150	0.0154	0.0140	0.0047	−0.0053

---

[36] The parameter normalizations that were used in Section 3 are also used in this section.

| Model 7 | 0.8583 | 182.55 | 0.0261 | 0.0134 | 0.0166 | 0.0123 | 0.0039 | −0.0052 |

Model 2 added 5 additional neighbourhood parameters to Model 1, Model 3 added 6 marginal price parameters for the size of the land plot, Model 4 added 5 marginal price parameters for the size of the structure, Model 5 added 5 depreciation rate parameters, Model 6 added 5 extra bathroom parameters and Model 7 added 4 extra bedroom parameters. Looking at the increases in the log likelihood that are listed in Table 5, it can be seen that the additional variables do add a considerable amount of explanatory power. The estimated coefficients for the new Model 7 are listed in Table 6. It can be seen that with the exception of the depreciation rates, the estimated coefficients for the new piecewise linear Model 7 are similar to the estimated coefficients listed in Table 3 for the geometric depreciation Model 7.

Table 6: Estimated Coefficients for Model 7 Using Piecewise Linear Depreciation

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
α_1	62.12	17.07	α_{24}	117.17	18.39	λ_7	0.1284	2.57
α_2	63.36	16.70	α_{25}	122.00	18.36	μ_1	1.4405	28.39
α_3	59.38	13.84	α_{26}	128.00	18.59	μ_2	1.3871	7.95
α_4	50.32	11.62	α_{27}	129.90	18.32	μ_3	1.5689	8.65
α_5	52.21	16.44	α_{28}	130.90	18.68	μ_4	2.3665	12.22
α_6	63.65	17.32	α_{29}	142.39	18.63	μ_5	0.8046	6.79
α_7	68.42	17.71	α_{30}	153.49	18.53	δ_1	0.0261	22.16
α_8	75.13	17.55	α_{31}	169.83	18.36	δ_2	0.0134	10.01
α_9	88.62	17.89	α_{32}	182.21	18.47	δ_3	0.0166	11.80
α_{10}	91.88	18.33	α_{33}	229.75	18.48	δ_4	0.0123	7.52
α_{11}	90.28	18.12	α_{34}	245.28	18.45	δ_5	0.0039	2.18
α_{12}	94.06	18.34	α_{35}	236.14	18.32	δ_6	-0.0052	-1.36
α_{13}	125.33	18.54	α_{36}	217.91	17.72	η_1	1.1244	11.95
α_{14}	122.23	18.60	ω_1	0.6384	69.17	η_2	0.9483	25.98
α_{15}	116.67	18.55	ω_2	0.9883	128.92	η_4	1.0079	60.23
α_{16}	118.36	18.57	ω_3	0.8641	166.89	η_5	0.9844	51.76
α_{17}	126.83	18.10	ω_5	0.9335	185.39	η_6	1.0583	46.62
α_{18}	120.68	18.29	ω_6	0.5499	41.06	θ_3	1.1056	40.22
α_{19}	114.49	17.60	λ_1	1.2629	17.40	θ_4	1.0838	90.07
α_{20}	107.77	17.44	λ_2	0.7536	8.64	θ_6	0.8602	73.16
α_{21}	110.00	18.20	λ_3	0.3192	4.15	θ_7	0.7337	29.76
α_{22}	113.64	18.38	λ_4	0.3226	4.73			
α_{23}	114.08	18.39	λ_6	0.6102	6.57			

Although the estimated depreciation rates for our new Model 7 differ considerably from the geometric depreciation rates that we estimated for Model 7 in Section 3, the age functions that these two models generate are somewhat similar. If structure depreciation is geometric at the rate of 3% per year, the quality adjusted fraction of a structure that is A years old is $g_1(A) \equiv (1 − 0.03)^A$. This function of age A is the lowest curve on Chart 1 below. Using the 6 decade by decade geometric depreciation rates that are listed in Table 2 and the 6 piecewise linear depreciation rates that are listed in Table 6 for Model 7, we can generate similar age functions which are denoted by $g_2(A)$ and $g_3(A)$ on Chart 1. It can be seen that all three age functions approximate each other closely for the first decade of age but they diverge in subsequent decades. The two geometric age functions, $g_1(A)$

and $g_2(A)$, are reasonably close to each other but the piecewise linear age function $g_3(A)$ lies considerably above the two other age functions. However, as we shall see, since most of the sales of houses in Richmond are for relatively young structures, the new and the old Model 7 generate much the same sales price indexes.[37]

Chart 1: Alternative Age Functions
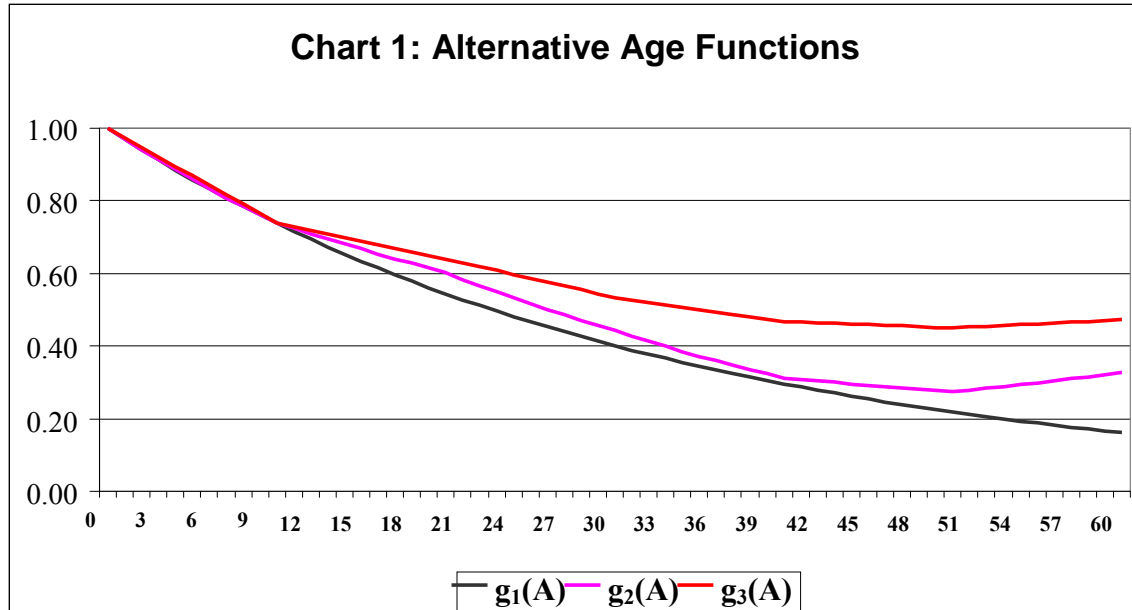
We use the same procedures that were used in Section 4 in order to calculate the land and property price indexes that are implied by Models 4-7. The overall sales property price indexes for quarter t for the new Models 4-7 are denoted by P_{4Nt}-P_{7Nt} and the corresponding quality adjusted land prices are denoted by P_{4LNt}-P_{7LNt}. These indexes are listed in Table 7. The quality adjusted structure price indexes for the new Models 4-7 are all equal to the indexes P_{St} listed in Table 3 in Section 3.

Table 7: New Property Price Indexes P_{4N}-P_{7N} and Land Price Indexes P_{4LN}-P_{7LN} for Sales of Properties in Richmond Q1 2008 to Q4 2016

| t	P_{4Nt}	P_{5Nt}	P_{6Nt}	P_{7Nt}	P_{4LNt}	P_{5LNt}	P_{6LNt}	P_{7LNt}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0058	1.0082	1.0049	1.0101	1.0084	1.0148	1.0079	1.0199
3	0.9840	0.9868	0.9848	0.9799	0.9736	0.9735	0.9693	0.9558
4	0.8925	0.8937	0.8974	0.8947	0.8436	0.8192	0.8265	0.8100
5	0.8769	0.8750	0.8747	0.8729	0.8627	0.8499	0.8487	0.8404
6	0.9273	0.9258	0.9254	0.9236	0.9882	1.0167	1.0164	1.0246
7	0.9752	0.9734	0.9708	0.9704	1.0537	1.0915	1.0868	1.1014
8	1.0350	1.0338	1.0315	1.0357	1.1334	1.1834	1.1798	1.2093

[37] The final LL for the Model 7 geometric model of depreciation was −72920.50 while the final LL for the Model 7 piecewise linear model of depreciation was slightly lower at −72923.68 so the geometric model is a slightly preferred model. But the depreciation rates for older houses are not determined very precisely so the issue of which depreciation model best fits the data is open.

9	1.1534	1.1488	1.1477	1.1468	1.3153	1.3946	1.3947	1.4265
10	1.1814	1.1783	1.1768	1.1791	1.3515	1.4386	1.4381	1.4790
11	1.1689	1.1665	1.1666	1.1647	1.3341	1.4198	1.4223	1.4533
12	1.1933	1.1910	1.1914	1.1890	1.3786	1.4761	1.4795	1.5142
13	1.4438	1.4333	1.4313	1.4263	1.7767	1.9514	1.9521	2.0175
14	1.4144	1.4092	1.4082	1.4052	1.7282	1.9001	1.9024	1.9676
15	1.3722	1.3658	1.3652	1.3622	1.6617	1.8163	1.8191	1.8780
16	1.3871	1.3804	1.3780	1.3716	1.6892	1.8513	1.8505	1.9054
17	1.4541	1.4443	1.4412	1.4346	1.7961	1.9783	1.9770	2.0417
18	1.4127	1.3967	1.3953	1.3877	1.7308	1.8856	1.8871	1.9427
19	1.3629	1.3529	1.3483	1.3429	1.6483	1.7958	1.7901	1.8431
20	1.3099	1.2984	1.3002	1.2928	1.5619	1.6866	1.6938	1.7349
21	1.3265	1.3164	1.3135	1.3056	1.5935	1.7298	1.7277	1.7707
22	1.3508	1.3430	1.3392	1.3301	1.6358	1.7878	1.7842	1.8294
23	1.3568	1.3440	1.3401	1.3316	1.6477	1.7928	1.7890	1.8364
24	1.3693	1.3569	1.3564	1.3496	1.6737	1.8273	1.8311	1.8861
25	1.4063	1.3904	1.3878	1.3824	1.7375	1.9002	1.8999	1.9638
26	1.4479	1.4340	1.4325	1.4223	1.8101	1.9960	1.9985	2.0604
27	1.4628	1.4462	1.4443	1.4351	1.8355	2.0227	2.0243	2.0910
28	1.4774	1.4580	1.4554	1.4451	1.8565	2.0417	2.0420	2.1072
29	1.5674	1.5406	1.5383	1.5283	2.0036	2.2106	2.2125	2.2921
30	1.6477	1.6269	1.6236	1.6160	2.1270	2.3734	2.3740	2.4708
31	1.7765	1.7540	1.7534	1.7429	2.3282	2.6174	2.6246	2.7339
32	1.8817	1.8544	1.8515	1.8410	2.4905	2.8067	2.8099	2.9332
33	2.2623	2.2265	2.2238	2.2098	3.0810	3.5200	3.5269	3.6984
34	2.3942	2.3610	2.3550	2.3387	3.2768	3.7630	3.7647	3.9483
35	2.3363	2.3042	2.2986	2.2793	3.1753	3.6345	3.6360	3.8012
36	2.2070	2.1642	2.1586	2.1438	2.9691	3.3561	3.3559	3.5077

Comparing the indexes listed in Tables 3 and 7, it can be seen that while P_{4N} differs somewhat from P_4 (and P_{4LN} differs considerably from P_{4L}), the overall property price indexes P_5-P_7 are very close to their new counterparts P_{5N}-P_{7N} and the land price indexes P_{5L}-P_{7L} are very close to their new counterparts P_{5LN}-P_{7LN}. This close correspondence is explained by the fact that once multiple depreciation rates are introduced into the age function for structures, the geometric and piecewise linear models of depreciation approximate each other sufficiently so that Models 5-7 generate similar results for both types of depreciation.

Chart 2 plots the land price indexes P_{4L}, P_{6L}, P_{7L} for the geometric models and compares these indexes with the land price indexes P_{4LN}, P_{6LN}, P_{7LN} for the piecewise linear models of depreciation.[38] The lowest land price series, P_{4LN}, which uses a single straight line depreciation rate, generates indexes that are rather different from the other models. The land price indexes P_{4L}, P_{6L} (based on geometric depreciation) and the land price index P_{6LN} (based on piecewise linear depreciation rates) are also too close to distinguish on Chart 2. The highest land price indexes are P_{7L} and P_{7LN} and they are too close to each other to distinguish on Chart 2.

---

[38] P_{5L} is too close to P_{6L} and P_{5LN} is too close to P_{6LN} to be distinguished on a chart.

Chart 2: Land Price Indexes Using Geometric and Piecewise Linear Depreciation Models
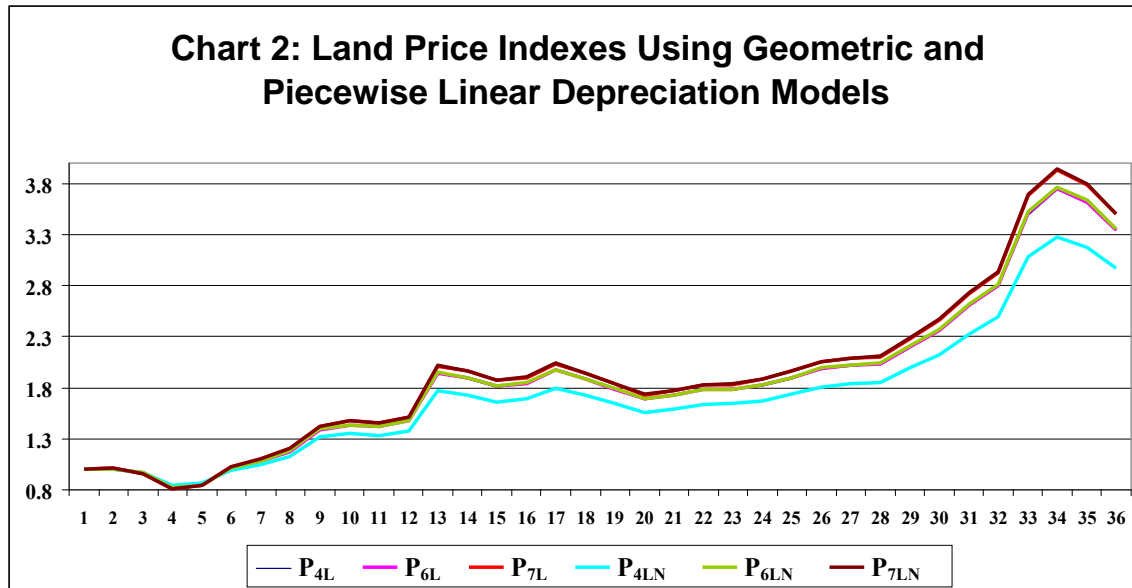
Our conclusion at this point is that if we use multiple depreciation rates, the price indexes that the geometric and generalized straight line depreciation generate are essentially identical. A more tentative conclusion is that the geometric model of depreciation with a single rate (Model 4) generates an acceptable land price index and hence also generates an overall sales property price index that approximates subsequent property price indexes based on models that require information on more characteristics.

It should be noted that the depreciation rates that we have estimated in this paper do not account for all of the structure depreciation that occurs in the housing market. Our data on sales of properties do not include structures which have been demolished before they reach the end of their useful life. Hulten and Wykoff (1981) indicate one method for adjusting for this additional demolition depreciation and Diewert and Shimizu (2017) suggest an alternative method which requires information on the age of structures when they are demolished.[39] Another issue which requires additional analysis is associated with the treatment of renovation expenditures: should they be immediately expensed in a consumer price index or should they be capitalized and depreciated separately over time? Looking at our estimated age functions that appear on Chart 1, it can be seen that renovation expenditures are probably significant determinants of structure value for older houses. This suggests that these expenditures should be capitalized.

In the following section, we turn our attention to "traditional" house price hedonic regressions and ask whether these regressions can generate property price indexes and depreciation rates that are comparable to the estimates generated by the builder's model.

---

[39] In their study of commercial properties in Tokyo, Diewert and Shimizu (2017) found that the geometric depreciation rate generated by the builder's model was 2.6% per year and demolition depreciation amounted to an additional 1.2% per year contribution to annual depreciation. Thus it is important to account for demolition when constructing estimates of depreciation for the national accounts or for tax policy purposes.

7. Estimating Structure Depreciation Rates from Traditional Log Price Hedonic Regression Models

A way of rationalizing the traditional log price time dummy hedonic regression model for properties with varying amounts of land area L and constant quality structure area S^* is that the utility that these properties yield to consumers is proportional to the Cobb-Douglas utility function $L^\alpha S^{*\beta}$ where α and β are positive parameters (which do not necessarily sum to one).[40] Initially, we assume that the constant quality structure area S^* is equal to the floor space area of the structure, S, times an age adjustment, $(1-\delta)^A$, where A is the age of the structure in years and δ is a positive depreciation rate that is less than 1. Thus S^* is related to S as follows:

$$(36) \quad S^* \equiv S(1-\delta)^A.$$

In any given time period t, we assume that the nominal value of a property, V_t, with the amount of land L and the amount of quality adjusted structure S^* is equal to the following expression:

$$(37) \quad \begin{aligned} V_t &= p_t L^\alpha S^{*\beta} \\ &= p_t L^\alpha [S(1-\delta)^A]^\beta \qquad \text{using (36)} \\ &= p_t L^\alpha S^\beta (1-\delta)^{\beta A} \\ &= p_t L^\alpha S^\beta \phi^A \end{aligned}$$

where p_t can be interpreted as a *period t property price index* and the constant ϕ is defined as follows:

$$(38) \quad \phi \equiv (1-\delta)^\beta.$$

Thus if we deflate V_t by the period t property price index p_t, we obtain the real value or utility u_t of the property with characteristics L and S^*; i.e., we have:

$$(39) \quad V_t/p_t = L^\alpha S^{*\beta} \equiv u_t.$$

Thus $u_t \equiv q_t$ is the *aggregate real value of the property* with characteristics L and S^*.

Define ρ_t as the logarithm of p_t and γ as the logarithm of ϕ; i.e.,

$$(40) \quad \rho_t \equiv \ln p_t \; ; \; \gamma \equiv \ln \phi.$$

---

[40] The early analysis in this section follows that of McMillen (2003; 289-290) and Shimizu, Nishimura and Watanabe (2010; 795). McMillen assumed that $\alpha+\beta = 1$. We follow Shimizu, Nishimura and Watanabe in allowing α and β to be unrestricted. Our depreciation model is somewhat different from the models of these authors. We interpret McMillen's model as a demand side model rather than a supply side model.

After taking logarithms of both sides of the last equation in (37) and using definitions (40), we obtain the following equation:[41]

(41) $\ln V_t = \rho_t + \alpha \ln L + \beta \ln S + \gamma A$.

Suppose that we have N(t) sales of properties in a neighbourhood in period t and the selling price of property n is P_{tn}. Suppose that we have data for T periods. The characteristics associated with property n in period t are L_{tn}, S_{tn} and A_{tn}. Assuming that the property purchaser valuation model defined by (41) holds approximately, the logarithms of the period t property prices will satisfy the following equations:

(42) $\ln P_{tn} = \rho_t + \alpha \ln L_{tn} + \beta \ln S_{tn} + \gamma A_{tn} + \varepsilon_{tn}$;   t = 1,...,T; n = 1,...,N(t);

where the ε_{tn} are independently distributed error terms with 0 means and constant variances. It can be seen that (42) is a traditional log price time dummy hedonic regression model with a minimal number of characteristics. The unknown parameters in (42) are the constant quality log property prices, $\rho_1,...,\rho_T$, and the taste parameters α, β and γ. Once these parameters have been determined, the geometric depreciation rate δ which appears in equation (36) can be recovered from the regression parameter estimates as follows:

(43) $\delta = 1 - e^{\gamma/\beta}$.

We now explain how the hedonic pricing model defined by (39) can be manipulated to provide a decomposition of property value in period t into land and quality adjusted structure components.

From (39), we assume that there is a period t constant quality price of property in period t equal to p_t (as before) and the period t value of a property with characteristics L and S^* is given by the following *period t property valuation function*, $V(p_t, L, S^*)$, defined as follows:

(44) $V(p_t, L, S^*) \equiv p_t L^\alpha S^{*\beta} = p_t u(L, S^*) \equiv V_t$

where the purchaser cardinal utility function u is defined as $u(L, S^*) \equiv L^\alpha S^{*\beta}$ and α and β are positive parameters which parameterize the purchaser's cardinal utility function. P_t is the period t selling price of a property with land and quality adjusted structures equal to L and S^* and V_t is the predicted selling price that is generated using the estimated coefficients from the hedonic regression (42). In our empirical work, our estimates for α and β are always such that $\alpha + \beta$ is always substantially less than 1. This means that a property in a given period that has double the land and quality adjusted structure than another property will sell for less than double the price of the smaller property. This follows from the fact that our Cobb-Douglas utility function, $u(L, S^*) \equiv L^\alpha S^{*\beta}$, exhibits diminishing returns to scale; i.e., we have:

---

[41] Log price hedonic regressions for property prices date back to Bailey, Muth and Nourse (1963).

(45) $u(\lambda L, \lambda S^*) = \lambda^{\alpha+\beta} u(L, S^*)$

for all $\lambda > 0$ where $\alpha+\beta < 1$. This behavior is roughly consistent with our builder's Models 5-7 where there was a tendency for property prices to increase less than proportionally as L and S^* increased.

The *marginal prices of land and constant quality structure* in period t for a property with characteristics L and S^*, $\pi_{tL}(p_t, L, S^*)$ and $\pi_{tS*}(p_t, L, S^*)$, are defined by partially differentiating the property valuation function with respect to L and S^* respectively:

(46) $\pi_{tL}(p_t, L, S^*) \equiv \partial V(p_t, L, S^*)/\partial L \equiv p_t \alpha L^\alpha S^{*\beta}/L = \alpha V(p_t, L, S^*)/L$;
(47) $\pi_{tS}^*(p_t, L, S^*) \equiv \partial V(p_t, L, S^*)/\partial S^* \equiv p_t \beta L^\alpha S^{*\beta}/S^* = \beta V(p_t, L, S^*)/S^*$.

If we multiply the marginal price of land by the amount of land in the property and add to this value of land the product of the marginal price of constant quality structure by the amount of constant quality structure on the property, we obtain the following identity:

(48) $(\alpha+\beta)V(p_t, L, S^*) = \pi_{tL}(p_t, L, S^*)L + \pi_{tS}^*(p_t, L, S^*)^* S$.

If $\alpha+\beta$ is less than one, then using marginal prices to value the land and constant quality structure in a property will lead to a property valuation that is less than its selling price. Thus to make the land and structure components of property value add up to property value, we divide the marginal prices defined by (46) and (47) by $\alpha+\beta$ in order to obtain *adjusted prices of land and structures in period t*, $p_{tL}(p_t, L, S^*)$ and $p_{tS*}(p_t, L, S^*)$:

(49) $p_{tL}(p_t, L, S^*) \equiv \pi_{tL}(p_t, L, S^*)/(\alpha+\beta) = \alpha(\alpha+\beta)^{-1}V(p_t, L, S^*)/L$;
(50) $p_{tS*}(p_t, L, S^*) \equiv \pi_{tS*}(p_t, L, S^*)/(\alpha+\beta) = \beta(\alpha+\beta)^{-1}V(p_t, L, S^*)/S^*$.

The above material outlines a theoretical framework that can generate a decomposition of property value into land and structure components using the results of a traditional log price time dummy hedonic regression model. To complete the analysis, we need to fill in the details of how the individual property land and structure prices that are generated by the model can be aggregated into period t overall land and structure price indexes.

Let P_{tn} denote the selling price of property n in period t and denote the land area, floor space area and age of the structure by L_{tn}, S_{tn} and A_{tn} for $t = 1,...,T$ and $n = 1,...,N(t)$. Run the hedonic regression defined by (42) and denote the resulting parameter estimates by $\rho_1,...,\rho_T$, α, β and γ. Define the constant quality property price index p_t for period t as follows:

(51) $p_t \equiv \exp(\rho_t)$; $t = 1,...,T$.

Define the geometric depreciation rate δ by (43). Once δ has been defined, the amount of quality adjusted structure for property n in period t, S^*_{tn} can be defined as follows:

(52) $\ln(S_{tn}^{*}) \equiv \ln(S_{tn}) + A_{tn}\ln(1-\delta)$; $t = 1,...,T; n = 1,...,N(t).$

Now that p_t, L_{tn}, S_{tn}^{*}, α and β have all been defined, we use these data in equations (44) in order to define the predicted prices for property n sold in period t, V_{tn}:

(53) $V_{tn} \equiv p_t (L_{tn})^{\alpha} (S_{tn}^{*})^{\beta}$; $t = 1,...,T; n = 1,...,N(t).$

Use the following modifications of equations (49) and (50) in order to define *constant quality land and structure prices* for sold property n in period t, p_{tLn} and p_{tS*n}:

(54) $p_{tLn} \equiv \alpha(\alpha+\beta)^{-1}V_{tn}/L_{tn}$; $t = 1,...,T; n = 1,...,N(t);$
(55) $p_{tS*n} \equiv \beta(\alpha+\beta)^{-1}V_{tn}/S_{tn}^{*}$; $t = 1,...,T; n = 1,...,N(t).$

Finally, *unit value constant quality land and structure prices* for all properties sold in period t, p_{tL} and p_{tS*}, are defined as follows:

(56) $p_{tL} \equiv (\Sigma_{n=1}^{N(t)} p_{tLn} L_{tn})/(\Sigma_{n=1}^{N(t)} L_{tn})$; $t = 1,...,T;$
(57) $p_{tS*} \equiv (\Sigma_{n=1}^{N(t)} p_{tS*n}S_{tn}^{*})/(\Sigma_{n=1}^{N(t)} S_{tn}^{*})$; $t = 1,...,T.$

The period t land and structure prices that are defined by (56) and (57) are reasonable summary statistic prices for land and structures sold in period t that are generated by the log price time dummy hedonic regression model defined by (42). However, when we implemented a generalization of this model, the resulting land and structure price indexes were not reasonable as will be seen below.

The problem with this hedonic regression model shows up in equations (54) and (55). Suppose a representative property (with the same characteristics) sold in periods t−1 and t. Let the period t−1 and t property values be $V_{t-1,r}$ and V_{tr} and let the constant quality amounts of land and structures in both periods be $L_{t-1,r} = L_{tr} \equiv L_r$ and $S_{t-1,r}^{*} = S_{tr}^{*} \equiv S_r^{*}$. Substitute these values into equations (54) and (55) and calculate (one plus) the rate of increase in constant quality land value and constant quality structure value. We obtain the following equations:

(58) $p_{tLr} L_{tr}/p_{t-1,Lr} L_{t-1,r} = p_{tS*r} S_{tr}^{*}/p_{t-1,S*r} S_{t-1,r}^{*} = V_{tr}/V_{t-1,r}.$

Thus the log price hedonic model predicts that the rate of increase going from one period to the next in the land value of a representative constant quality property is equal to the corresponding rate of increase in the structure value of the property and these two rates of increase are in turn equal to the overall rate of increase in property value of the representative property. This is clearly not true for Richmond properties during our sample period: the value of a unit of land increased far more than the value of a unit of constant quality structure.[42]

---

[42] Our interpretation of the log price hedonic regression model is a utility driven purchaser side model as opposed to the builder's model which is a supply side model. Strictly speaking, we can be reasonably sure that the supply side model applies to new houses but it is not so certain that it applies to older structures

We indicate how the above very simple log price hedonic regression model can be generalized to include additional (discrete) characteristics of the properties. Suppose that properties have been classified to 6 postal zones. If property n in period t belongs to postal zone j, then define the dummy variable $d_{j,tn}$ for observation n in period t to equal 1 and if property n in period t does not belongs to postal zone k, then define the dummy variable $d_{j,tn}$ for observation n in period t to equal 0. Next, suppose that properties have been classified according to the number of bathrooms m in the structure where 6 is the maximum number of bathrooms. If property n in period t has i bathrooms, then define the dummy variable $d_{1i,tn}$ for observation n in period t to equal 1 and if property n in period t does not have i bathrooms, then define the dummy variable $d_{1i,tn}$ for observation n in period t to equal 0. Finally, suppose that properties have been classified according to the number of bedrooms k in the structure where the number of bedrooms ranges from 3 to 7. If property n in period t has k bedrooms, then define the dummy variable $d_{2k,tn}$ for observation n in period t to equal 1 and if property n in period t does not have k bedrooms, then define the dummy variable $d_{2k,tn}$ for observation n in period t to equal 0. Now consider the following generalization of the hedonic regression model defined by (42):

$$(59) \ \ln P_{tn} = \rho_t + \alpha \ln L_{tn} + \beta \ln S_{tn} + \gamma A_{tn} + \Sigma_{j=1}^{6} \omega_j d_{j,tn} + \Sigma_{i=1}^{6} \eta_i d_{1i,tn}$$
$$+ \Sigma_{k=3}^{7} \theta_k d_{2k,tn} + \varepsilon_{tn} ; \qquad\qquad t = 1,...,T; n = 1,...,N(t).$$

The ω_j parameters affect the quality of the land component of property value while the last two sets of dummy variables affect the quality of the structure component of property value. Not all of the parameters ρ_t, ω_j, η_i and θ_k can be identified; i.e., there is exact multicollinearity associated with the dummy variables associated with these parameters.[43] Thus to identify all of the remaining parameters, we make the following normalizations that are counterparts to the normalizations (25) for Model 7:

$$(60) \ \omega_4 \equiv 0; \ \eta_3 \equiv 0; \ \theta_5 = 0.$$

Once the linear regression defined by equations (59) and the normalizations (60) has been run, define the amount of *quality adjusted land* associated with property n in period t, L_{tn}^{*}, as follows:

$$(61) \ L_{tn}^{*} \equiv L_{tn} \exp[\Sigma_{j=1}^{6} \omega_j d_{j,tn}/\alpha] ; \qquad\qquad t = 1,...,T; n = 1,...,N(t).$$

Define the amount of *quality adjusted structure* associated with property n in period t, S_{tn}^{*}, as follows:

$$(62) \ S_{tn}^{*} \equiv S_{tn} \exp[(\gamma A_{tn} + \Sigma_{i=1}^{6} \eta_i d_{1i,tn} + \Sigma_{k=3}^{7} \theta_k d_{2k,tn})/\beta] ; \qquad t = 1,...,T; n = 1,...,N(t).$$

---

that are sold since both the land and structure are fixed factors once the structure is built. Thus it is possible that the purchaser side model is more appropriate for older structures. However, it is clear that the purchaser side model fails for new houses.

[43] The dummy variables associated with the ρ_t in equations (59) have been suppressed.

Now replace L_{tn} in equations (52)-(58) by $L_{tn}{}^*$ defined by (61) and replace the $S_{tn}{}^*$ in equations (52)-(58) by the new $S_{tn}{}^*$ defined by (62). Equations (58) still hold for this more general hedonic regression model except that L_{tn} is replaced by $L_{tn}{}^*$ defined by (61) and we use the new definition (62) for $S_{tn}{}^*$. Thus the new more general hedonic regression model still does not capture economic reality; i.e., the rates of growth in the values of land and structures are far from being equal for Richmond houses sold during our sample period.

We illustrate the above theoretical results by running the log price linear regression model defined by (59) and (60) using our Richmond data. The R^2 for this regression turned out to be 0.8550 which is comparable to the R^2 we obtained for our Model 7 nonlinear regressions. The estimated coefficients this traditional log price time dummy linear regression along with their t statistics are listed in Table 8.[44]

Table 8: Estimated Coefficients for the Log Price Time Dummy Regression

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
ρ_1	5.937	426.9	ρ_1	6.226	382.3	γ	-0.0079	-44.63
ρ_2	5.957	424.5	ρ_2	6.189	372.7	β	0.2604	25.76
ρ_3	5.926	346.8	ρ_3	6.189	410.5	α	0.3806	52.20
ρ_4	5.834	290.6	ρ_4	6.215	447.6	ω_1	-0.2589	-44.47
ρ_5	5.832	383.4	ρ_5	6.215	451.2	ω_2	-0.0318	-5.59
ρ_6	5.879	448.3	ρ_6	6.233	427.9	ω_3	-0.1145	-29.05
ρ_7	5.929	456.6	ρ_7	6.257	449.8	ω_5	-0.0531	-14.42
ρ_8	5.989	426.5	ρ_8	6.284	463.0	ω_6	-0.3273	-41.20
ρ_9	6.085	453.1	ρ_9	6.293	457.6	η_1	0.1151	11.62
ρ_{10}	6.096	445.7	ρ_{10}	6.316	442.1	η_2	0.0255	4.95
ρ_{11}	6.081	423.3	ρ_{11}	6.363	477.7	η_4	0.0489	9.93
ρ_{12}	6.113	441.2	ρ_{12}	6.427	505.2	η_5	0.0631	9.39
ρ_{13}	6.283	479.7	ρ_{13}	6.496	486.1	η_6	0.1162	14.40
ρ_{14}	6.265	455.9	ρ_{14}	6.556	488.4	θ_3	0.0207	4.02
ρ_{15}	6.240	431.1	ρ_{15}	6.750	513.6	θ_4	0.0177	4.63
ρ_{16}	6.239	400.4	ρ_{16}	6.797	519.3	θ_6	-0.0568	-10.99
ρ_{17}	6.283	435.8	ρ_{17}	6.771	436.2	θ_7	-0.1004	-9.33
ρ_{18}	6.249	418.8	ρ_{18}	6.717	311.7			

Using the estimates for γ and β that are listed in Table 8, we can use equation (43) in order to compute the geometric depreciation rate that is implicit in this hedonic regression model. The resulting estimate turns out to be $\delta = 1 - e^{\gamma/\beta} = 0.03003$; i.e., using our demand side Cobb-Douglas property valuation model, the implied annual geometric depreciation rate turns out to be 3.00% per year, which is very close to the single geometric depreciation rate that we obtained for Model 4 in Section 3, which was 2.98% per year. This is a very useful result: it says that it may be the case that we can obtain

---

[44] Note that the ρ_t listed in Table 8 correspond to the logarithms of the α_t which appeared in Tables 3 and 7 and the ω_j, η_i and θ_k which appear in Table 8 correspond to the logarithms of the same parameters which appeared in Tables 3 and 7.

reasonably accurate structure depreciation rates using the results of traditional log price time dummy hedonic regressions, which are widespread in the literature.[45]

Using the parameter estimates that are listed in Table 8, we used definitions (51) in order to calculate the sequence of quarter t quality adjusted property prices, p_t. The normalized sequence of property prices was defined as $P_t \equiv p_t/p_1$. This series is listed in Table 9 below. We can use definitions (61) and (62) in order to form estimates for the quantity of quality adjusted land and structure, L_{tn}^* and S_{tn}^*, for each property in our Richmond sample of properties. Equations (53)-(57) were then used in order to form the constant quality unit value quarter t prices, p_{tL*} and p_{tS*}.[46] These prices were normalized to equal 1 in quarter 1; i.e., define $P_{Lt} \equiv p_{tL*}/p_{1L*}$ and $P_{St} \equiv p_{tS*}/p_{1S*}$ for t = 1,...,36. These constant quality price series for land and structures are also listed in Table 9.

A final set of price indexes was computed using the Cobb-Douglas property valuation model. Instead of running a single log price regression with 36 time dummy variables, it is possible to run a sequence of 35 *adjacent period log price regressions*. Each one of these regressions uses only the data for two adjacent periods but otherwise, the methodology is exactly the same as the one regression methodology that uses the entire sample. Each adjacent period regression can be used to update the overall property price index level of the previous period as well as the implied land and structure price levels. Thus this alternative methodology uses a chaining methodology. The advantage of this alternative methodology is that the taste parameters are only held constant for two successive periods for each of the regressions and so this approach allows for taste changes over time. The disadvantage of the adjacent period approach is that it can introduce a certain amount of volatility into the price indexes. In any case, we implemented this approach and the resulting overall, land and structure price indexes, P_{CHt}, P_{CHLt} and P_{CHSt}, are listed in Table 9. For comparison purposes, we also listed the Model 7 builder's model counterpart price indexes, P_{7t}, P_{7Lt} and P_{7St}, from Table 3.

Table 9: Single Regression Cobb-Douglas Property, Land and Structure Price Indexes, P_t, P_{Lt} and P_{St}, their Adjacent Period Counterpart Indexes, P_{CHt}, P_{CHLt} and P_{CHSt}, and the Corresponding Model 7 Geometric Depreciation Indexes, P_{7t}, P_{7Lt} and P_{7St}

t	P_t	P_{CHt}	P_{7t}	P_{Lt}	P_{CHLt}	P_{7Lt}	P_{St}	P_{CHSt}	P_{7St}
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0204	1.0188	1.0103	1.0334	1.0350	1.0203	0.9929	0.9769	1.0020
3	0.9896	0.9875	0.9805	1.0113	0.9952	0.9574	0.9520	0.9487	0.9990
4	0.9026	0.9046	0.8945	0.9253	0.9400	0.8103	0.8048	0.7982	0.9610
5	0.9011	0.8907	0.8729	0.9493	0.9652	0.8407	0.7718	0.7143	0.9006
6	0.9442	0.9340	0.9236	0.9888	1.0029	1.0238	0.9060	0.8889	0.8509
7	0.9925	0.9874	0.9703	1.0184	1.0384	1.1002	0.9197	0.9023	0.8743

---

[45] Note that the logarithms of the land and structure areas must be used in the log price hedonic regression in order for our suggested methodology for estimating structure depreciation rates to work. Usually, the log likelihood that is obtained using the logarithms of L and S as independent variables is higher than the LL obtained using the levels of L and S as independent variables; this was the case for our Richmond data.
[46] As mentioned earlier, L_{tn} in equations (52)-(58) is replaced by the quality adjusted land, L_{tn}^*, defined by (61).

8	1.0542	1.0508	1.0357	1.0614	1.0833	1.2080	0.9692	0.9507	0.9064
9	1.1600	1.1590	1.1468	1.1219	1.1448	1.4243	1.1365	1.1211	0.9318
10	1.1734	1.1740	1.1794	1.1566	1.1831	1.4772	1.2250	1.2149	0.9474
11	1.1557	1.1521	1.1649	1.1304	1.1535	1.4514	1.1663	1.1566	0.9425
12	1.1932	1.1896	1.1892	1.1626	1.1871	1.5120	1.1649	1.1348	0.9357
13	1.4139	1.4069	1.4266	1.2780	1.3057	2.0130	1.5629	1.5598	0.9367
14	1.3891	1.3905	1.4059	1.2993	1.3255	1.9642	1.5518	1.5806	0.9425
15	1.3547	1.3576	1.3625	1.3315	1.3561	1.8743	1.3365	1.3801	0.9406
16	1.3530	1.3570	1.3721	1.2642	1.2862	1.9018	1.5324	1.5792	0.9337
17	1.4143	1.4176	1.4350	1.3618	1.3765	2.0372	1.4936	1.4938	0.9298
18	1.3667	1.3723	1.3880	1.3768	1.3937	1.9384	1.3389	1.2863	0.9279
19	1.3353	1.3381	1.3431	1.3680	1.3819	1.8390	1.1946	1.1858	0.9279
20	1.2877	1.2989	1.2930	1.2994	1.3198	1.7315	1.1864	1.1639	0.9240
21	1.2871	1.2989	1.3057	1.3150	1.3353	1.7668	1.2070	1.1575	0.9181
22	1.3207	1.3378	1.3307	1.2938	1.3115	1.8264	1.2583	1.2186	0.9142
23	1.3208	1.3322	1.3321	1.3586	1.3821	1.8330	1.1612	1.1531	0.9113
24	1.3448	1.3579	1.3501	1.3552	1.3809	1.8824	1.1764	1.1731	0.9045
25	1.3774	1.3896	1.3830	1.3649	1.3889	1.9599	1.2456	1.2493	0.9016
26	1.4160	1.4371	1.4228	1.3772	1.4028	2.0559	1.4360	1.4614	0.8938
27	1.4290	1.4481	1.4359	1.4392	1.4617	2.0867	1.2727	1.3241	0.8918
28	1.4612	1.4748	1.4458	1.3901	1.3967	2.1027	1.3864	1.4910	0.8967
29	1.5312	1.5475	1.5290	1.4824	1.4898	2.2868	1.4101	1.5177	0.8947
30	1.6331	1.6524	1.6169	1.5851	1.5945	2.4650	1.5761	1.7077	0.9064
31	1.7506	1.7687	1.7439	1.6640	1.6752	2.7270	1.7406	1.8781	0.9172
32	1.8576	1.8398	1.8420	1.7766	1.7285	2.9253	1.7896	1.9868	0.9289
33	2.2560	2.2133	2.2110	2.0547	2.0053	3.6870	2.4491	2.6565	0.9493
34	2.3634	2.3393	2.3401	2.2650	2.2110	3.9361	2.4260	2.6634	0.9717
35	2.3039	2.2950	2.2802	2.2769	2.2261	3.7888	2.1920	2.3749	0.9883
36	2.1827	2.1617	2.1449	2.1611	2.1049	3.4972	2.1311	2.2771	0.9873

The single regression and the adjacent period overall property price indexes P_t and P_{CHt} are plotted on Chart 3 along with our final geometric depreciation property price index P_{7t} and the mean and median price indexes P_{Mean} and P_{Median} for comparison purposes.

Chart 3: Alternative Overall Property Price Indexes for Richmond
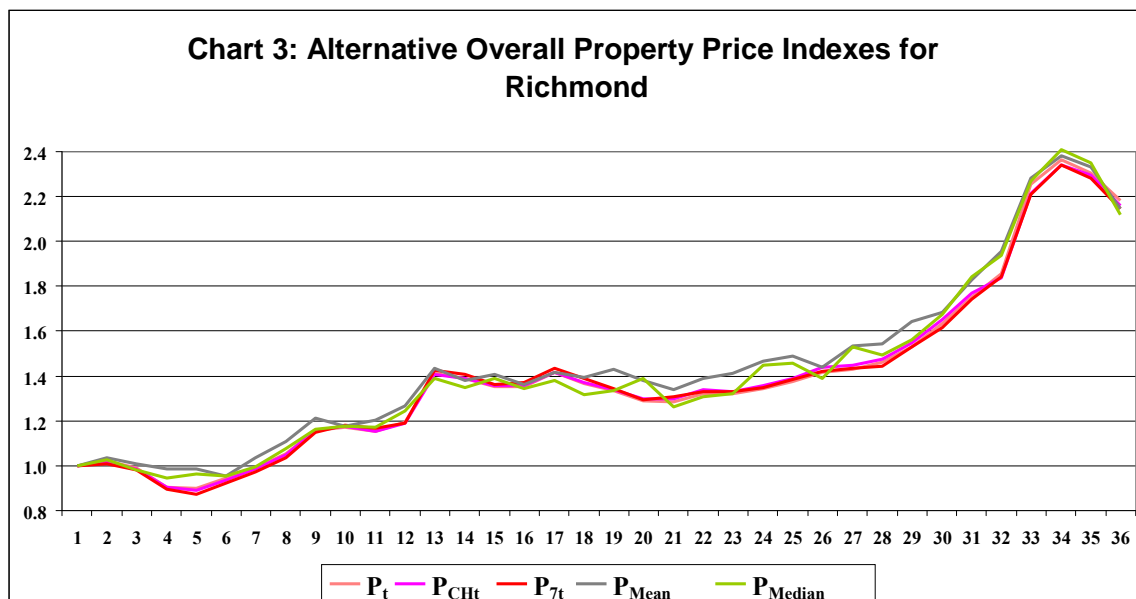
From viewing Chart 3, it can be seen that the two traditional log price time dummy property price indexes, P_t and P_{CHt}, can barely be distinguished from our final geometric depreciation property price index P_{7t}. This is a very encouraging result: it means that it is possible for the traditional time dummy log price hedonic regressions to generate overall property price indexes that are consistent with the overall indexes generated by the builder's model. From Chart 3, it is also evident that simple mean and median property price indexes do capture the trends in property prices to a reasonable degree of approximation but the volatility of these indexes as compared to the hedonic indexes is apparent.

Finally, Chart 4 plots the three land price indexes listed in Table 9, P_{Lt}, P_{CHLt} and P_{7Lt}, as well as the three constant quality structure price indexes listed in Table 9, P_{St}, P_{CHSt} and P_{7St}.

Chart 4: Alternative Land and Structure Price Indexes
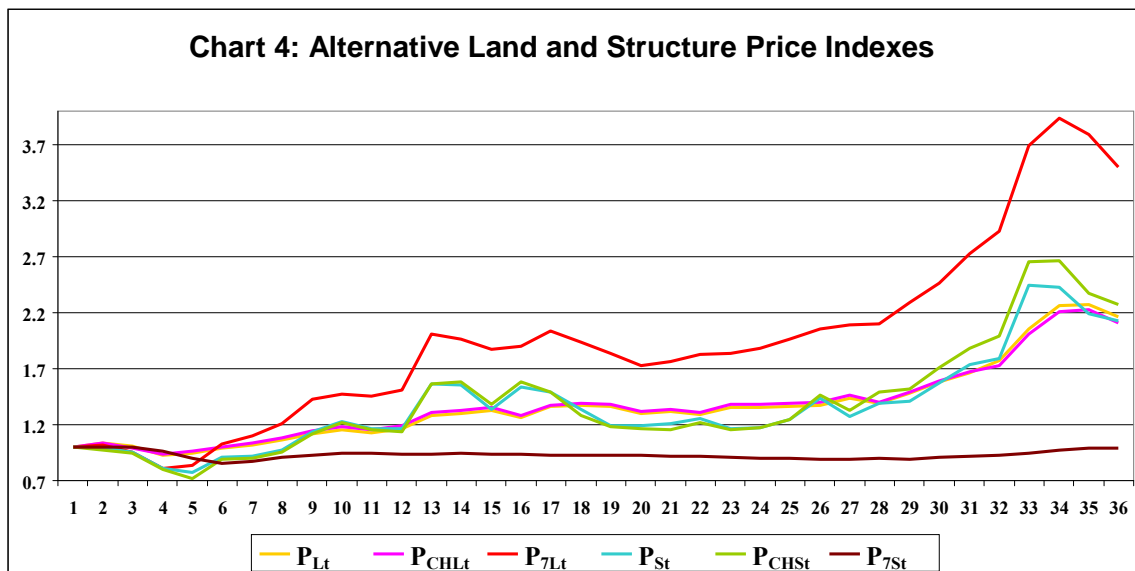
Viewing Chart 4, it can be seen that the builder's Model 7 land price series, P_{7Lt}, lies far above the counterpart log price time dummy series P_{Lt} and P_{CHLt}. These last two series approximate each other very closely and are fairly smooth. On the other hand, the builder's Model 7 structure price series, P_{7St}, lies far below the counterpart log price time dummy series P_{St} and P_{CHSt}. These last two series approximate each other fairly closely up to the final 10 quarters of our sample period when the chained index moves well above its single regression counterpart. Interestingly, these two structure price indexes based on traditional time dummy hedonic regressions are much more volatile than their counterpart time dummy land price indexes while the Statistics Canada structure price index, P_{7St}, is extremely smooth. Note that, as predicted, the structure and land price indexes based on traditional time dummy hedonic regressions are all basically similar.

Given our knowledge of the Greater Vancouver property market, we conclude that our imputation procedure applied to traditional log price time dummy regressions does not

generate reasonable housing price subindexes for the land and structure components of overall property price.

8. Conclusion

Some of the conclusions we can draw from the paper are as follows:

- The builder's model can generate reasonable overall house price indexes along with reasonable land and structure subindexes using just four characteristics: the land plot area L, the structure floor space area S, the age of the structure A and the location of the property (typically the postal code).
- Introducing spline segments for the land and structure area of a property does lead to a massive improvement in the fit of the builder's model.
- It is useful to introduce multiple depreciation rates for different ages of the structure in terms of improving the fit of the model. However, a single geometric depreciation rate does provide an adequate approximation to the more complex models for the Richmond data.
- Once we introduce multiple age dependent depreciation rates, the geometric and piecewise linear depreciation models generate virtually identical indexes. This conclusion does not hold if we have only a single depreciation rate for both types of model.
- In countries with rapidly rising residential land prices (such as Australia and Canada), the inflation rate that is generated by estimating a sales price index will tend to understate the corresponding house price inflation rate for the stock of housing.
- Simple mean and median property price indexes generated property price indexes that captured the trend in our constant quality property price indexes.[47] However, these indexes are much more volatile than our hedonic property price indexes (as is well known).
- The traditional log price time dummy hedonic regression approach generated overall property price indexes which were virtually identical to our builder's model overall property price indexes when both types of model used the maximum number of characteristics.[48]
- The traditional log price time dummy hedonic regression approach generated an implied geometric structure depreciation rate which was virtually identical to the single geometric depreciation rate generated by the builder's model. This is a very encouraging result.
- However, our demand side utility interpretation of the traditional log price time dummy hedonic regression approach did not generate reasonable land and

---

[47] This conclusion does not always hold but it is somewhat encouraging that it can hold given that many published property price indexes are essentially trimmed mean or median sales price indexes.
[48] This result is useful in the context of trying to use existing information on residential property price indexes along with information on the capital stock of residential structures in order to provide a complete decomposition of residential property value into land and structure subcomponents; see Davis and Heathcote (2007) for a possible methodology.

structure subindexes (and we explained why this result holds). The builder's supply side model seems to generate much more reasonable results.

We conclude the paper with a cautionary comment: In order to obtain sensible results from a hedonic property price regression, it is essential to trim the tails of the distributions of both the dependent and independent variables. If this is not done, nonsensical results can emerge from the regressions.

References

Altus Group (2015), *Construction Cost Guide 2015*, available on the web at:
    http://www.altusgroup.com/media/4099/costguide_2015_web.pdf

Bailey, M.J., R.F. Muth and H.O. Nourse (1963), "A Regression Method for Real Estate Price Construction", *Journal of the American Statistical Association* 58, 933-942.

Bostic, R.W., S.D. Longhofer and C.L. Readfearn (2007), "Land Leverage: Decomposing Home Price Dynamics", *Real Estate Economics 35:2*, 183-2008.

Burnett-Issacs, K., N. Huang and W.E. Diewert (2016), "Developing Land and Structure Price Indexes for Ottawa Condominium Apartments", Discussion Paper 16-09, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada.

Clapp, J.M. (1980), "The Elasticity of Substitution for Land: The Effects of Measurement Errors", Journal of Urban Economics 8, 255-263.

Davis, M.A. and J. Heathcote (2007), "The Price and Quantity of Residential Land in the United States", Journal of Monetary Economics 54. 2595-2620.

de Haan, J., W.E. Diewert (eds.) (2011), *Residential Property Price Indices Handbook*, Luxembourg: Eurostat.
http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/methodology/hps/rppi_handbook

Diewert, W.E. (2008), "OECD Workshop on Productivity Analysis and Measurement: Conclusions and Future Directions", pp. 11-36 in *Proceedings From the OECD Workshops on Productivity Measurement and Analysis*, Paris: OECD.

Diewert, W.E. (2010), "Alternative Approaches to Measuring House Price Inflation", Discussion Paper 10-10, Department of Economics, The University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E. (2011), "The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Further Directions", pp. 87-116 in *Price and Productivity Measurement, Volume 1*, Housing, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Publishing.

Diewert, W. E. and C. Shimizu (2015), "Residential Property Price Indexes for Tokyo," *Macroeconomic Dynamics* 19, 1659-1714.

Diewert, W. E. and C. Shimizu (2016), "Hedonic Regression Models for Tokyo Condominium Sales," *Regional Science and Urban Economics* 60, 300-315.

Diewert, W. E. and C. Shimizu (2017), "Alternative Approaches to Commercial Property Price Indexes for Tokyo," *Review of Income and Wealth*, forthcoming.

Diewert, W.E., J. de Haan and R. Hendriks (2011), "The Decomposition of a House Price Index into Land and Structures Components: A Hedonic Regression Approach", *The Valuation Journal* 6, 58-106.

Diewert, W.E., J. de Haan and R. Hendriks (2015), "Hedonic Regressions and the Decomposition of a House Price index into Land and Structure Components", *Econometric Reviews* 34, 106-126.

Fisher, I. (1922), *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Francke, M.K. (2008), "The Hierarchical Trend Model", pp. 164-180 in *Mass Appraisal Methods: An International Perspective for Property Valuers*, T. Kauko and M. Damato (eds.), Oxford: Wiley-Blackwell.

Francke, M. K. and G. A. Vos. (2004), "The Hierarchical Trend Model for Property Valuation and Local Price Indices," *Journal of Real Estate Finance and Economics* 28, 179-208.

Hulten, C. and F. Wykoff (1981), "The Estimation of Economic Depreciation Using Vintage Asset Prices", *Journal of Econometrics* 15, 367-396.

Koev, E. and J.M.C. Santos Silva (2008), "Hedonic Methods for Decomposing House Price Indices into Land and Structure Components", unpublished paper, Department of Economics, University of Essex, England, October.

Lowe, J. (1823), The Present State of England in Regard to Agriculture, Trade and Finance, second edition, London: Longman, Hurst, Rees, Orme and Brown.

McMillen, D.P. (2003), "The Return of Centralization to Chicago: Using Repeat Sales to Identify Changes in House Price Distance Gradients", *Regional Science and Urban Economics* 33, 287–304.

Rambaldi, A.N., R.R.J McAllister, K. Collins and C.S. Fletcher (2010), "Separating Land from Structure in Property Prices: A Case Study from Brisbane Australia", School of Economics, The University of Queensland, St. Lucia, Queensland 4072, Australia.

Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economy* 82, 34-55.

Shimizu, C.,K.G. Nishimura and T. Watanabe (2010), "Housing Prices in Tokyo: A Comparison of Hedonic and Repeat Sales Measures", *Journal of Economics and Statistics* 230, 792-813.

Schwann, G.M. (1998), "A Real Estate Price Index for Thin Markets", *Journal of Real Estate Finance and Economics* 16:3, 269-287.

White, K.J. (2004), *Shazam: User's Reference Manual*, Version 10, Vancouver, Canada: Northwest Econometrics Ltd.