Alternative Land Price Indexes for Commercial Properties in Tokyo

Erwin Diewert
(University of British Columbia, Canada)

Chihiro Shimizu
(Nihon University, Japan)

Paper prepared for the 35th IARIW General Conference

Copenhagen, Denmark, August 20-25, 2018

Session 4C-2: Prices

Time: Wednesday, August 22, 2018 [14:00-17:00]

Alternative Land Price Indexes for Commercial Properties in Tokyo

Erwin Diewert and Chihiro Shimizu[1] Revised October 2, 2017
Discussion Paper 17-07,
Vancouver School of Economics,
The University of British Columbia,
Vancouver, Canada, V6T 1L4.

Abstract

The SNA (System of National Accounts) requires separate estimates for the land and structure components of a commercial property. Using transactions data for the sales of office buildings in Tokyo, a hedonic regression model (the Builder's Model) was estimated and this model generated an overall property price index as well as subindexes for the land and structure components of the office buildings. The Builder's Model was also estimated using appraisal data on office building REITs for Tokyo. These hedonic regression models also generated estimates for net depreciation rates which can be compared. Finally, the Japanese Government constructs annual official land prices for commercial properties based on appraised values. The paper compares these official land prices with the land prices generated by the hedonic regression models based on transactions data and on REIT data. The results revealed that commercial property indexes based on appraisal and assessment prices lag behind the indexes based on transaction prices.

Key Words

Commercial property price indexes, System of National Accounts, the builder's model, transaction-based indexes, appraisal prices, assessment prices, land and structure price indexes, hedonic regressions, depreciation rates.

Journal of Economic Literature Classification Numbers

C2, C23, C43, D12, E31, R21.

---

[1] W. Erwin Diewert: School of Economics, University of British Columbia, Vancouver B.C., Canada, V6T 1L4 and the School of Economics, University of New South Wales, Sydney, Australia (email: erwin.diewert@ubc.ca) and Chihiro Shimizu, Nihon University, Setagaya, Tokyo, 154-8513, Japan, email: shimizu.chihiro@nihon-u.ac.jp .The authors thank David Geltner for helpful discussions. We gratefully acknowledge the financial support of the SSHRC of Canada and JSPS KAKEN(S) #25220502, Nomura Foundation.

1. Introduction

When estimating commercial property price indexes, we are confronted with the following two problems: how to incorporate quality adjustments in the estimation method, and which data source to use in the estimation procedure.

Research studies on commercial property price indexes have emphasized the problem of data selection when formulating indexes. Traditionally, transaction prices (also called market prices in the literature) have usually been used to estimate price indexes. However, the number of commercial property market transactions is extremely small. Furthermore, even if a sizable number of transaction prices can be obtained, the heterogeneity of the properties is so pronounced that it is difficult to compare like with like and thus the construction of reliable constant quality price indexes becomes very difficult.

Under such circumstances, many commercial property price indexes have been constructed using either appraisal prices from the real estate investment market, or using assessment prices for property tax purposes. The rationale for these price indexes is that, since appraisal prices and assessment prices for property tax purposes are regularly surveyed for the same commercial property, indexes based on these surveys hold most characteristics of the property constant[2], thus greatly reducing the heterogeneity problem as well as generating a wealth of data.

However, while appraisal prices look attractive for the construction of price indexes, they are somewhat subjective; i.e., exactly how are these appraisal prices constructed? Thus these prices lack the objectivity of market selling prices. Such considerations have led to the development of various arguments concerning the precision and accuracy of appraisal and assessment prices when used in measuring price indexes; see Shimizu and Nishimura (2006) on these issues. In particular, the literature on this issue has pointed out that an appraisal based index will typically lag actual turning points in the real estate market.[3] Geltner, Graff and Young (1994) clarified the structure of bias in the NCREIF Property Index, a representative U.S. index based on appraisal prices. In a later study, Geltner and Goetzmann (2000) estimated an index using commercial property transaction prices and demonstrated the magnitude of errors and the degree of smoothing in the NCREIF Property Index. These problems plague not only the NCREIF Property Index, but all indexes based on appraisal prices, including the MSCI-IPD Index.

With specific reference to Japan's real estate bubble period, Nishimura and Shimizu (2003), Shimizu and Nishimura (2006), and Shimizu, Nishimura and Watanabe (2012) estimated hedonic price indexes based on commercial property and residential housing transaction price based indexes and contrasted them with appraisal price based indexes

[2] Two important characteristics which are not held constant are the age of the structure and the amount of capital expenditures on the property between the survey dates. Changes in these characteristics are an important determinant of the property price.

[3] Another problem with appraisal based indexes is that they tend to be smoother than indexes that are based on market transactions. This can be a problem for real estate investors since the smoothing effect will mask the short term riskiness of real estate investments. However, for statistical agencies, smoothing short term fluctuations will probably not be problematic.

and statistically laid out their differences. An examination of the estimated results revealed that during the bubble period, when prices climbed dramatically, indexes based on appraisal prices did not catch up with transaction price increases. Similarly, during the period of falling prices, appraisal based indexes did not keep pace with the decline in prices.

Furthermore, in the case of appraisal prices for investment properties, a systemic factor of appraiser incentives emerges as an additional problem. This problem differs intrinsically from the lagging and smoothing problems that arise in appraisal based methods. Specifically, the incentive problem involves inducing higher valuations from appraisers in order to bolster investment performance; see Crosby, Lizieri and McAllister (2010) on this point.

In this connection, Bokhari and Geltner (2012) and Geltner and Bokhari (2017) estimated quality adjusted price indexes by running a time dummy hedonic regression using transaction price data. Geltner (1997) also used real estate prices determined by the stock market in order to examine the smoothing effects of the use of appraisal prices. Finally, Geltner, Pollakowski, Horrigan and Case (2010), Shimizu, Diewert, Nishimura and Watanabe (2015), Shimizu (2016) and Diewert and Shimizu (2017) proposed various estimation methods for commercial property price indexes using REIT data.

In this paper, we will examine the three alternative data sources suggested in the literature that enable one to construct land price indexes for commercial properties: (i) sales transactions data; (ii) appraisal data for Real Estate Investment Trusts (REITs) and (iii) assessed values of land for property taxation purposes. We will utilize these three sources of data for commercial properties in Tokyo over 44 quarters covering the period Q1:2005 to Q4:2015 and compare the resulting land prices.

Section 2 below explains our data sources. Sections 3 and 4 use sales transactions data and a hedonic regression model that allows us to decompose sale prices into land and structure components. The model of structure depreciation used in Section 3 is a single geometric rate and Section 4 generalizes this model to allow for multiple geometric rates. Section 5 implements the same hedonic regression model using the same transactions data set but we switch to a piece-wise linear depreciation model. Section 6 compares the alternative depreciation schedules.

It will turn out that the land price series that are generated using quarterly transactions data are very volatile and thus they may not be suitable for statistical agency use. Thus in Section 7, we look at some alternative methods for smoothing the raw land price indexes.

Section 8 estimates a hedonic regression model using quarterly appraisal values for 41 Tokyo office buildings over the sample period. Since we have panel data for this application, our hedonic regression model is somewhat different from our earlier models.

Section 9 estimates quality adjusted land prices for commercial properties using tax assessment data. Section 10 compares our land price indexes from the three sources of data. Section 11 constructs overall property price indexes for Tokyo commercial properties using the models estimated in the previous sections; i.e., we combine the land price indexes with a structure price index to obtain overall property price indexes. We also estimate a traditional log price time dummy hedonic regression model and compare the resulting index with our overall indexes. Section 12 concludes.

2. Data Description

This study compiled the following three types of micro-data relating to commercial properties in the Tokyo office market: (i) the transaction price data compiled by the Japanese Ministry of Land, Infrastructure, Transport and Tourism; (ii) the appraisal prices periodically determined in the Tokyo office REIT market; and (iii) the "official land prices" surveyed by the Japanese Ministry of Land, Infrastructure, Transport and Tourism since 1970. Official land prices are based on appraisals that are released on January 1st of each year. In Japan, asset taxes relating to land, such as inheritance taxes and fixed assets taxes, are assessed on the basis of these official land prices. Thus official land prices are considered as assessment data for tax purposes. As official land prices are exclusively based on surveys of land prices, they do not include structure prices.

Table 1. Variables from the Three Data Sources

Symbols	Variables	Contents	Unit
V	Price	Transaction Price and Appraisal price in Total	million yen
L	Total Land area	Land area of building	m^2
S	Total Floor space	Floor space of building.	m^2
A	Age of building at the time of transaction.	Age of building at the time of transaction/appraisal	years
H	Number of stories	Number of stories in the building	stories
DS	Distance to the nearest station	Distance to the nearest station.	meters
TT	Travel time to central business district	Minimum railway riding time in daytime to Tokyo Station.	minutes
WD_k (k=0,…,K)	Location(Ward) dummy	k th ward =1, other district =0.	(0,1)
D_t (t=0,…,T)	Time dummy (quarterly)	*t th* quarter =1, other quarter =0.	(0,1)

Using the first two data sources, land price indexes were estimated using the Builder's Model. These land price indexes will be compared with those estimated using official land prices in Section 5 of the paper.

Our analysis covers the period from 2005 to 2015. The data variables compiled are listed in Table 1.

Table 2 shows a summary of the statistical parameters for the 3 data sources, i.e. transaction prices, REIT appraisal prices, and official land prices. The compiled data consisted of 1,907 MLIT transaction prices, 1,804 REIT prices, and 6,242 MLIT official land prices which we label as Official Land Prices (OLP).

Table 2. Summary Statistics

	MLIT	REIT	OLP
V : Selling Price of Office Building (million yen)	394.18	6686.60	1264.3
	(337.76)	(4055.60)	(1304.1)
S : Structure Floor Area (m^2)	834.00	8509.70	-
	(535.19)	(5463.90)	
L : Land Area (m^2)	239.27	1802.10	229.94
	(135.08)	(1580.20)	(217.18)
H : Total Number of Stories	5.75	10.12	-
	(2.14)	(3.30)	
A : Age (years)	24.23	19.14	-
	(10.61)	(6.80)	
DS : Distance to Nearest Station (meters)	387.65	308.29	347.24
	(238.45)	(170.04)	(254.79)
TT : Time to Tokyo Station (minutes)	19.63	15.88	21.74
	(8.23)	(5.10)	(8.54)
PS : Structure Construction Price per m^2 (million yen)	0.2347	0.2359	-
	(0.0103)	(0.0102)	
Number of Observations	1,907	1,804	6,242

( ): Standard deviation

3. The Builder's Model: Preliminary Results Using Transactions Data

We will use the MLIT commercial office building transactions data in this section and in sections 4-7 below.

The *builder's model* for valuing a commercial property postulates that the value of a commercial property is the sum of two components: the value of the land which the structure sits on plus the value of the commercial 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 meters, times the building cost per square meter, β_t during quarter or year t, plus the cost of the land, which will be equal to the cost per square meter, α_t during quarter or year t, times the area of the land site, L. Now think of a sample of properties of the same general type, which have prices or values V_{tn} in period t[4] 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:[5]

$$(1)\ V_{tn} = \alpha_t L_{tn} + \beta_t S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

Note that the two characteristics in our simple model are the quantities of land L_{tn} and the quantities of structure floor space S_{tn} associated with property n in period t and the two *constant quality prices* in period t are the price of a square meter of land α_t and the price of a square meter of structure floor space β_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 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*:[6]

$$(2)\ V_{tn} = \alpha_t L_{tn} + \beta_t(1 - \delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,...,44;\ n = 1,...,N(t)$$

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 between 2 to 3%.[7] Note that (2) is now a nonlinear regression model whereas (1) was a simple linear regression model.[8] The period t constant quality price of land will be the estimated coefficient for the parameter

[4] The period index t runs from 1 to 44 where period 1 corresponds to Q1 of 2005 and period 44 corresponds to Q4 of 2015.

[5] 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), de Haan and Diewert (2011), Diewert (2008) (2010), Francke (2008; 167), Koev and Santos Silva (2008), Rambaldi, McAllister, Collins and Fletcher (2010), Diewert, Haan and Hendriks (2011) (2015) and Diewert and Shimizu (2015b) (2016) (2017).

[6] This formulation follows that of Diewert (2008) (2010), de Haan and Diewert (2011), Diewert, de Haan and Hendriks (2015) and Diewert and Shimizu (2015b) (2016) (2017) 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. This formulation is designed to be useful for national income accountants who require a decomposition of property value into structure and land components. They also need the structure index which in the hedonic regression model to be consistent with the structure price index they use to construct structure capital stocks. Thus the builder's model is particularly suited to national accounts purposes; see Diewert and Shimizu (2015b) and Diewert, Fox and Shimizu (2016).

[7] 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.

[8] We used Shazam to perform the nonlinear estimations; see White (2004).

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 commercial property n is L_{tn} and the period t quantity of structure for commercial property n, expressed in equivalent units of a new structure, is $(1 - \delta_t)^{A(t,n)}S_{tn}$ where S_{tn} is the space area of commercial property n in period t.

Note that the above model is a *supply side model* as opposed to the *demand side model* of Muth (1971) and McMillen (2003). Basically, we are assuming competitive suppliers of commercial 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 offices but it is less well justified for sales of existing commercial properties.

There is a major 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.[9] Thus in order to deal with the multicollinearity problem, we draw on *exogenous information* on the cost of building new commercial properties from the Japanese Ministry of Land, Infrastructure, Transport and Tourism (MLIT) and we assume that the price of new structures is equal to an official measure of commercial building costs (per square meter of building structure), p_{St}. Thus we replace β_t in (2) by p_{St} for t = 1,...,44. This reduces the number of free parameters in the model by 44.

Experience has also shown that it is difficult to estimate the depreciation rate before obtaining quality adjusted land prices. Thus in order to get preliminary land price estimates, we temporarily assumed that the annual geometric depreciation rate δ in equation 2 was equal to 0.025. The resulting regression model becomes the model defined by (3) below:

$$(3)\ V_{tn} = \alpha_t L_{tn} + p_{St}(1 - 0.025)^{A(t,n)}S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

The final log likelihood for this **Model 1** was −13328.15 and the R^2 was 0.4003.[10]
In order to take into account possible neighbourhood effects on the price of land, we introduce *ward dummy variables*, $D_{W,tnj}$, into the hedonic regression (3). There are 23 wards in Tokyo special district. We made 23 ward or locational dummy variables.[11] These 23 dummy variables are defined as follows: for t = 1,...,44; n = 1,...,N(t); j = 1,...,23:

[9] See Schwann (1998) and Diewert, de Haan and Hendriks (2011) and (2015) on the multicollinearity problem.
[10] Our R^2 concept is the square of the correlation coefficient between the dependent variable and the predicted dependent variable.
[11] The 23 wards (with the number of observations in brackets) are as follows: 1: Chiyoda (191), 2: Chuo (231), 3: Minato (205), 4: Shinjuku (203), 5: Bunkyo (97), 6: Taito (122), 7: Sumida (74), 8: Koto (49), 9: Shinagawa (69), 10: Meguro (28), 11: Ota (64), 12: Setagaya (67), 13: Shibuya (140), 14: Nakano (39), 15: Suginami (39), 16: Toshima (80), 17: Kita (30), 18: Arakawa (42), 19: Itabashi (35), 20: Nerima (40), 21: Adachi (19), 22: Katsushika (18), 23:Edogawa (25).

(4) $D_{W,tnj} \equiv 1$ if observation n in period t is in ward j of Tokyo;
$\equiv 0$ if observation n in period t is not in ward j of Tokyo.

We now modify the model defined by (3) to allow the *level* of land prices to differ across the Wards. The new nonlinear regression model is the following one:[12]

(5) $V_{tn} = \alpha_t(\sum_{j=1}^{23} \omega_j D_{W,tnj})L_{tn} + p_{St}(1 - 0.025)^{A(t,n)}S_{tn} + \varepsilon_{tn}$; $t = 1,...,44; n = 1,...,N(t).$

Not all of the land time dummy variable coefficients (the α_t) and the ward dummy variable coefficients (the ω_j) can be identified. Thus we impose the following normalization on our coefficients:

(6) $\alpha_1 = 1.$

The final log likelihood for the model defined by (5) and (6) was −12956.60 and the R^2 was 0.5925. Thus there was a large increase in the R^2 and a huge increase in the log likelihood of 371.55 over the previous model. However, many of the wards had only a small number of observations and thus it is unlikely that our estimated ω_j for these wards would be very accurate.

In order to deal with the problem of too few observations in many wards, we used the results of the above model to group the 23 wards into 4 Combined Wards based on their estimated ω_j coefficients. The *Group 1 high priced wards* were 1,2,3 and 13 (their estimated ω_j coefficients were greater than 1), the *Group 2 medium high priced wards* were 4,5,6,9 and 14 ($0.6 < \omega_j \le 1$), the *Group 3 medium low priced wards* were 7,8,10,12,15 and 16 ($0.4 < \omega_j \le 0.6$), and the *Group 4 low priced wards* were 11,17,18,19 20,21,22 and 23 ($\omega_j \le 0.4$).[13] We reran the nonlinear regression model defined by (5) and (6) using just the 4 Combined Wards (call this **Model 2**) and the resulting log likelihood was −12974.31 and the R^2 was 0.5850. Thus combining the original wards into grouped wards resulted in a small loss of fit and a decrease in log likelihood of 17.71 when we decreased the number of ward parameters by 19. We regarded this loss of fit as an acceptable tradeoff.

In our next model, we introduce some nonlinearities into the pricing of the land area for each property. The land plot areas in our sample of properties run from 100 to 790 meters squared. Up to this point, we have assumed that land plots in the same grouped ward sell at a constant price per m^2 of lot area. However, it is likely that there is some nonlinearity in this pricing schedule; for example, it is likely that very large lots sell at an average price that is below the average price of medium sized lots. In order to capture this

---

[12] From this point on, our nonlinear regression models are nested; i.e., we use the coefficient estimates from the previous model as starting values for the subsequent model. Using this nesting procedure is essential to obtaining sensible results from our nonlinear regressions. The nonlinear regressions were estimated using Shazam; see White (2004).

[13] The estimated combined ward relative land price parameters turned out to be: $\omega_1 = 1.3003$; $\omega_2 = 0.75089$; $\omega_3 = 0.49573$ and $\omega_4 = 0.25551$. The sample probabilities of an observation falling in the combined wards were 0.402, 0.278, 0.177 and 0.143 respectively.

nonlinearity, we initially divided up our 1907 observations into 7 groups of observations based on their lot size. The Group 1 properties had lots less than 150 m^2, the Group 2 properties had lots greater than or equal to 150 m^2 and less than 200 m^2, the Group 3 properties had lots greater than or equal to 200 m^2 and less than 300 m^2, ... and the Group 7 properties had lots greater than or equal to 600 m^2. However, there were very few observations in Groups 4 to 7 so we added these groups to Group 4.[14] For each observation n in period t, we define the 4 *land dummy variables*, $D_{L,tnk}$, for k = 1,...,4 as follows:

(7) $D_{L,tnk} \equiv 1$ if observation tn has land area that belongs to group k;
$\equiv 0$ 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:

$$(8)\ f_L(L_{tn}) \equiv D_{L,tn1}\lambda_1 L_{tn} + D_{L,tn2}[\lambda_1 L_1+\lambda_2(L_{tn}-L_1)] + D_{L,tn3}[\lambda_1 L_1+\lambda_2(L_2-L_1)+\lambda_3(L_{tn}-L_2)] + D_{L,tn4}[\lambda_1 L_1+\lambda_2(L_2-L_1)+\lambda_3(L_3-L_2)+\lambda_4(L_{tn}-L_3)]$$

where the λ_k are unknown parameters and $L_1 \equiv 150$, $L_2 \equiv 200$ and $L_3 \equiv 300$. The function $f_L(L_{tn})$ defines a *relative valuation function for the land area of a commercial property* as a function of the plot area.

The new nonlinear regression model is the following one:

$$(9)\ V_{tn} = \alpha_t(\textstyle\sum_{j=1}^4 \omega_j D_{W,tnj})f_L(L_{tn}) + p_{St}(1-\delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

Comparing the models defined by equations (5)[15] and (9), it can be seen that we have added an additional 4 *land plot size parameters*, $\lambda_1,...,\lambda_4$, to the model defined by (5) (with only 4 ward dummy variables). However, looking at (9), it can be seen that the 44 land time parameters (the α_t), the 4 ward parameters (the ω_j) and the 4 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 (9) and (10):

$$(10)\ \alpha_1 \equiv 1;\ \lambda_1 \equiv 1.$$

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 59.65 over the final LL for Model 2 (for adding 3 new marginal price of land parameters) which is a highly significant increase. The R^2 increased to 0.6116 from the previous model R^2 of 0.5850. The parameter estimates turned out to be $\lambda_2 = 1.4297$, $\lambda_3 = 1.2772$ and $\lambda_4 = 0.2973$. For small land plot areas less than 150 m^2, we set the (relative) marginal price of land equal

[14] The sample probabilities of an observation falling in the 7 initial land size groups were: 0.291, 0.234, 0.229, 0.130, 0.050, 0.034 and 0.033.
[15] We compare (9) to the modified equation (5) where we have only 4 combined ward dummy variables in the modified (5) rather than the original 23 ward dummy variables.

to 1 per m^2. As lot sizes increased from 150 m^2 to 200 m^2, the (relative) marginal price of land increased to $\lambda_2 = 1.4297$ per m^2. For the next 100 m^2 of lot size, the (relative) marginal price of land decreased to $\lambda_3 = 1.2772$ per m^2. For lot sizes greater than 200 m^2, the (relative) marginal price of land decreased to 0.2973 per m^2. Thus the average cost of land per m^2 initially increases and then tends to decrease as lot size becomes large.

The *footprint* of a building is the area of the land that directly supports the structure. An approximation to the footprint land for property n in period t is the total structure area S_{tn} divided by the total number of stories in the structure H_{tn}. If we subtract footprint land from the total land area, TL_{tn}, we get *excess land*,[16] EL_{tn} defined as follows:

$$(11)\ EL_{tn} \equiv L_{tn} - (S_{tn}/H_{tn})\ ; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

In our sample, excess land ranged from 1.083 m^2 to 562.58 m^2. We grouped our observations into 5 categories, depending on the amount of excess land that pertained to each observation. Group 1 consists of observations tn where1: $EL_{tn} < 50$; 2: observations such that $50 \leq EL_{tn} < 100$; 3: $100 \leq EL_{tn} < 150$; 4: $150 \leq EL_{tn} < 300$; 5: $EL_{tn} \geq 300$.[17] Now define the excess land dummy variables, $D_{EL,tnm}$, as follows: for t = 1,...,44; n = 1,...,N(t); m = 1,...,5:

$$(12)\ D_{EL,tnm} \equiv 1 \text{ if observation n in period t is in excess land group m;}$$
$$\equiv 0 \text{ if observation n in period t is not in excess land group m.}$$

We will use the above dummy variables as adjustment factors to the price of land. As will be seen, in general, the more excess land a property possessed, the lower was the average per meter squared value of land for that property.[18]

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

$$(13)\ V_{tn} = \alpha_t(\textstyle\sum_{j=1}^{4} \omega_j D_{W,tnj})(\sum_{m=1}^{5} \chi_m D_{EL,tnm}) f_L(L_{tn}) + p_{St}(1-\delta)^{A(t,n)} S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

However, looking at (13) and (8), it can be seen that the 44 land price parameters (the α_t), the 4 combined ward parameters (the ω_j), the 4 land plot size parameters (the λ_k) and the 5 excess land parameters (the χ_m) cannot all be identified. Thus we imposed the following identifying normalizations on these parameters:

---

[16] This is land that is usable for purposes other than the direct support of the structure on the land plot. Excess land was first introduced as an explanatory variable in a property hedonic regression model for Tokyo condominium sales by Diewert and Shimizu (2016; 305).

[17] The sample probabilities of an observation falling in the 4 excess land size groups were: 0.352, 0.343, 0.149, 0.114 and 0.041.

[18] The excess land characteristic was also used by Diewert and Shimizu (2016b) and Burnett-Isaacs, Huang and Diewert (2017) in their studies of condominium prices. The same phenomenon was observed in these studies: the more excess land that a high rise property had, the lower was the per meter land price.

(14) $\alpha_1 \equiv 1$; $\lambda_1 \equiv 1$; $\chi_1 \equiv 1$.

Note that if we set all of the χ_m equal to unity, Model 4 collapses down to Model 3. The final log likelihood for Model 4 was an improvement of 23.99 over the final LL for Model 3 (for adding 4 new excess land parameters) which is a significant increase. The R^2 increased to 0.6207 from the previous model R^2 of 0.6116. The χ_m parameter estimates turned out to be $\chi_2 = 0.9173$, $\chi_3 = 0.7540$, $\chi_4 = 0.7234$ and $\chi_5 = 0.8611$. Thus excess land does reduce the average per meter price of land.

It is likely that the height of the building increases the value of the land plot supporting the building, all else equal. In our sample of commercial property prices, the height of the building (the H variable) ranged from 3 stories to 14 stories. Thus there are 12 building height categories. Thus we define the building height dummy variables, $D_{H,tnh}$, as follows: for t = 1,...,44; n = 1,...,N(t); h = 3,...,14:

(15) $D_{H,tnh} \equiv 1$ if observation n in period t is in a building which has height h;
$\equiv 0$ if observation n in period t is not in a building which has height h.

The new nonlinear regression model is the following one:

$$(16)\; V_{tn} = \alpha_t(\textstyle\sum_{j=1}^{4}\omega_j D_{W,tn,j})(\sum_{m=1}^{5}\chi_m D_{EL,tnm})(\sum_{h=3}^{14}\mu_h D_{H,tnh}) f_L(L_{tn}) + p_{St}(1-\delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}\,; \quad t = 1,...,44;\ n = 1,...,N(t).$$

In addition to the normalizations (14), we also imposed the normalization $\mu_3 \equiv 1$. Note that if we set all of the μ_h equal to unity, the new model collapses down to Model 4.

The final log likelihood for the new model was −12,667.58, a big improvement of 247.08 over the final log likelihood for Model 4 (for adding 11 new height parameters). The R^2 increased to 0.6980 from the previous model R^2 of 0.6207. The μ_4 to μ_{14} parameter estimates turned out to be 1.106, 1.342, 1.448, 1.559, 2.012, 2.303, 2.672, 2.554, 2,773, 3.690 and 2.237 respectively. It can be seen that the land price of a property tended to increase linearly (approximately) as the height of the building increases. Thus in order to conserve degrees of freedom, we decided to replace the 12 categorical height parameters (the μ_h) by a single parameter μ associated with a continuous variable, the height H. Thus **Model 5** is the following nonlinear regression model (where H_{tn} is the number of stories of the structure for property n sold during period t):

$$(17)\; V_{tn} = \alpha_t(\textstyle\sum_{j=1}^{4}\omega_j D_{W,tnj})(\sum_{m=1}^{5}\chi_m D_{EL,tnm})(1+\mu(H_{tn}-3)) f_L(L_{tn}) + p_{St}(1-\delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}\,; \quad t = 1,...,44;\ n = 1,...,N(t).$$

Not all of the parameters in (17) can be identified so we again impose the normalizations (14). The final log likelihood for Model 5 was −12685.19, a big improvement of 205.47 over the final log likelihood for Model 4 (for adding 1 new height parameters). The R^2 increased to 0.6923 from the Model 4 R^2 of 0.6207. The height parameter μ turned out to be 0.2358. Thus the land value of the property increased 23.58% for each extra story of structure. This is a very substantial height premium.

Model 6 is the same as Model 5 except that we estimated the annual geometric depreciation rate δ instead of assuming that it was equal to 2.5%. The final log likelihood for Model 6 was −12680.66, an improvement of 4.53 over the final log likelihood for Model 4 (for adding 1 new parameter). The R^2 increased marginally to 0.6938 from the previous model R^2 of 0.6923. The estimated depreciation rate was 4.76% with a standard error of 0.009.

Recall that we used building height as a quality adjustment factor for the land area of the property. In our next model, we will use building height as a quality adjustment factor for the structure component of the property. Recall that the 12 building height dummy variables $D_{H,tnh}$ were defined by (15) above for h = 3,4, ..., 14. Due to the small number of observations in the last 5 height categories, we combined these dummy variables into a single height category that included all buildings of height 10 to 14 stories; i.e., the new $D_{H,tn10}$ was defined as $\Sigma_{h=10}^{14} D_{H,tnh}$. **Model 7** is defined as the following nonlinear regression model:

$$(18)\quad V_{tn} = \alpha_t(\Sigma_{j=1}^{4}\omega_j D_{W,tnj})(\Sigma_{m=1}^{5}\chi_m D_{EL,tnm})(1+\mu(H_{tn}-3))f_L(L_{tn}) + p_{St}(1-\delta)^{A(t,n)}(\Sigma_{h=3}^{10}\phi_h D_{H,tnh})S_{tn} + \varepsilon_{tn}\,; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

In addition to the normalizations in (14), we also imposed the normalization $\phi_3 = 1$ in order to insure a reasonable split between structure and land values. Thus the number of unknown parameters in Model 7 increased by 7 over the number of parameters in Model 6.

The final log likelihood for Model 7 was − 12640.40, an improvement of 40.26 over the final log likelihood for Model 6 (for adding 7 new parameters). The R^2 increased to 0.7063 from the previous model R^2 of 0.6938. The estimated depreciation rate δ was 3.41% with a standard error of 0.0077. The estimated $\phi_4,...,\phi_{10}$ were equal to 1.11, 1.31, 1.32, 1.11, 1.83, 2.01 and 2.12 (recall that ϕ_3 was set equal to 1). Thus as the height of the structure increased, the quality adjusted quantity of the structure increased (except for buildings with 7 stories; i.e., ϕ_7 was less than ϕ_6).

This completes our description of our preliminary hedonic regression models for Tokyo office buildings. In the following section, we will extend these preliminary models by estimating more complex depreciation schedules.

4. The Builder’s Model with Multiple Geometric Depreciation Rates

In the following model, we allowed the geometric depreciation rates to differ after each 10 year interval (except for the last interval).[19] We divided up our 1907 observations into

[19] The analysis in this section and the subsequent section follows the approach taken by Diewert, Huang and Burnett-Isaacs (2017). Geltner and Bokhari (2017) estimate a much more flexible model of

5 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 3 properties had structure ages greater than or equal to 20 years but less than 30 years, the Group 4 properties had structure ages greater than or equal to 30 years but less than 40 years and the Group 5 properties had structure ages greater than or equal to 40 years.[20] For each observation n in period t, we define the 5 *age dummy variables*, $D_{A,tni}$, for i = 1,...,5 as follows:

(19) $D_{A,tni} \equiv 1$ 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 *aging function*, $g_A(A_{tn})$, defined as follows:[21]

$$(20)\ g_A(A_{tn}) \equiv D_{A,tn1}(1-\delta_1)^{A(t,n)} + D_{A,tn2}(1-\delta_1)^{10}(1-\delta_2)^{(A(t,n)-10)} + D_{A,tn3}(1-\delta_1)^{10}(1-\delta_2)^{10}(1-\delta_3)^{(A(t,n)-20)} + D_{A,tn4}(1-\delta_1)^{10}(1-\delta_2)^{10}(1-\delta_3)^{10}(1-\delta_4)^{(A(t,n)-30)} + D_{A,tn5}(1-\delta_1)^{10}(1-\delta_2)^{10}(1-\delta_3)^{10}(1-\delta_4)^{10}(1-\delta_5)^{(A(t,n)-40)}\ .$$

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

The new **Model 8** nonlinear regression model is the following one:

$$(21)\ V_{tn} = \alpha_t(\sum_{j=1}^{4}\omega_j D_{W,tnj})(\sum_{m=1}^{5}\chi_m D_{EL,tnm})(1+\mu(H_{tn}-3))f_L(L_{tn}) + p_{St}g_A(A_{tn})(\Sigma_{h=3}^{10}\ \phi_h\ D_{H,tnh})S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,...,44;\ n = 1,...,N(t).$$

We imposed the normalizations $\alpha_1 \equiv 1$, $\lambda_1 \equiv 1$, $\chi_1 \equiv 1$ and $\phi_3 \equiv 1$. Note that Model 8 collapses down to Model 7 if $\delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta$. Thus the number of unknown parameters in Model 8 increased by 4 over the number of parameters in Model 7. The final log likelihood for Model 8 was −12631.21, an improvement of 9.19 over the final log likelihood for Model 7 (for adding 4 additional parameters). The R^2 increased to 0.7091 from the previous model R^2 of 0.7063. The estimated depreciation rates (with standard errors in brackets) were as follows: $\delta_1 = 0.0487$ (0.0111), $\delta_2 = 0.0270$ (0.0097), $\delta_3 = 0.0096$ (0.0106),[22] $\delta_4 = 0.0403$ (0.0154), $\delta_5 = -0.0319$ (0.0185). Thus properties with

commercial property depreciation using US transaction data by allowing an age dummy variable for each age of building. This methodological approach generates a combined land and structure depreciation rate whereas our approach will generate depreciation rates that apply only to the structure portion of property value.

[20] There were only 28 properties which had age greater than 50 years so these properties were combined with the age 40 to 50 properties.

[21] A_{tn} is the same as A(t,n). The aging function $g_A(A_{tn})$ quality adjusts a building of age A_{tn} into a comparable number of units of a new building.

[22] Recall that these depreciation rates are net depreciation rates. As surviving structures approach their middle age, renovations become important and thus a decline in the net depreciation rate is plausible. The pattern of depreciation rates is similar to the comparable geometric depreciation rates that were observed

structures which are over 40 years old tended to have a negative depreciation rate; i.e., the value of the structure tends to *increase* by 3.19% per year.[23]

There are two additional explanatory variables in our data set that may affect the price of land. Recall that DS was defined as the distance to the nearest subway station and TT as the subway running time in minutes to the Tokyo station from the nearest station. DS ranges from 0 to 1,500 meters while TT ranges from 1 to 48 minutes. Typically, as DS and TT increase, land value decreases.[24] **Model 9** introduces these new variables into the previous nonlinear regression model (21) in the following manner:

$$(22)\ V_{tn} = \alpha_t(\Sigma_{j=1}^{4}\omega_j D_{W,tnj})(\Sigma_{m=1}^{5}\chi_m D_{EL,tnm})(1+\mu(H_{tn}-3))(1+\eta(DS_{tn}-0))(1+\theta(TT_{tn}-1))\times f_L(L_{tn}) + p_{St}g_A(A_{tn})(\Sigma_{h=3}^{10}\phi_h D_{H,tnh})S_{tn} + \varepsilon_{tn}\ ; \quad t = 1,...,44;\ n = 1,...,N(t).$$

Thus two new parameters, η and θ, are introduced. If these new parameters are both equal to 0, then Model 9 collapses down to Model 8.

The final log likelihood for Model 9 was −12614.70, an improvement of 16.51 over the final log likelihood for Model 8 (for adding 2 additional parameters). The R^2 increased to 0.7142 from the previous model R^2 of 0.7091. The estimated walking distance parameter was $\eta = -0.00023$ (0.000066), which indicates that commercial property land value does tend to decrease as the walking distance to the nearest subway station increases. However, the estimated travel time to Tokyo Central Station parameter was $\theta = 0.0209$ (0.0053) which indicates that land value increases on average as the travel time to the central station increases, a relationship which was not anticipated. All of the estimated parameter coefficients and their t statistics are listed in Table 3 below.[25]

Table 3: Estimated Coefficients for Model 9

Coef	**Estimate**	**t Stat**	**Coef**	**Estimate**	**t Stat**	**Coef**	**Estimate**	**t Stat**
α_2	**1.5461**	**7.183**	α_{25}	**0.7937**	**3.739**	ω_4	**0.1229**	**3.510**
α_3	**1.6315**	**8.815**	α_{26}	**0.8785**	**3.185**	λ_2	**1.4212**	**4.160**
α_4	**1.5339**	**8.763**	α_{27}	**1.2128**	**4.792**	λ_3	**1.6805**	**7.867**
α_5	**1.4198**	**6.250**	α_{28}	**1.2489**	**7.593**	λ_4	**0.5771**	**4.199**
α_6	**1.7653**	**7.833**	α_{29}	**1.3320**	**6.819**	χ_2	**0.9070**	**24.025**
α_7	**2.1552**	**7.893**	α_{30}	**1.0175**	**4.038**	χ_3	**0.7634**	**17.948**
α_8	**1.7566**	**7.338**	α_{31}	**1.1255**	**5.402**	χ_4	**0.8278**	**16.228**

for Richmond (a suburb of Vancouver, Canada) detached houses by Diewert, Huang and Burnett-Isaacs (2016).

[23] This phenomenon has been observed in the housing literature before; i.e., older heritage houses that have been extensively renovated may increase in value over time rather than depreciate as they age. Diewert, Huang and Burnett-Isaacs (2016) found that Richmond house structures appreciated by 2.4% per year after age 50.

[24] See Diewert and Shimizu (2015b) where these relationships also held for Tokyo detached houses.

[25] Standard errors can be obtained by dividing the estimated coefficient by the corresponding t statistic.

α_{9}	**2.2697**	**7.743**	α_{32}	**1.5853**	**6.808**	χ_5	**0.9551**	**10.654**
α_{10}	**2.6226**	**7.593**	α_{33}	**1.2028**	**5.776**	μ	**0.0602**	**2.675**
α_{11}	**2.4724**	**7.271**	α_{34}	**1.4807**	**5.755**	η	**–0.0002**	**–3.533**
α_{12}	**2.4234**	**7.194**	α_{35}	**1.4105**	**7.359**	θ	**0.0209**	**3.924**
α_{13}	**2.4672**	**7.425**	α_{36}	**1.5614**	**6.432**	ϕ_4	**1.3063**	**6.503**
α_{14}	**1.7139**	**5.754**	α_{37}	**1.6905**	**7.738**	ϕ_5	**1.6760**	**10.260**
α_{15}	**1.7080**	**7.011**	α_{38}	**1.3886**	**6.174**	ϕ_6	**1.7117**	**10.687**
α_{16}	**2.0576**	**7.009**	α_{39}	**1.7169**	**7.528**	ϕ_7	**1.6865**	**10.216**
α_{17}	**1.4671**	**6.664**	α_{40}	**1.9941**	**7.675**	ϕ_8	**2.3615**	**13.392**
α_{18}	**0.8818**	**3.656**	α_{41}	**1.3744**	**6.291**	ϕ_9	**2.5553**	**14.103**
α_{19}	**0.6900**	**3.341**	α_{42}	**1.6397**	**6.161**	ϕ_{10}	**2.8092**	**13.714**
α_{20}	**1.1983**	**5.128**	α_{43}	**1.5263**	**6.219**	δ_1	**0.0484**	**5.221**
α_{21}	**0.9835**	**5.383**	α_{44}	**2.0459**	**7.031**	δ_2	**0.0252**	**2.955**
α_{22}	**1.1742**	**4.981**	ω_1	**0.6518**	**6.535**	δ_3	**0.0060**	**0.656**
α_{23}	**1.2670**	**5.864**	ω_2	**0.3483**	**5.455**	δ_4	**0.0389**	**2.925**
α_{24}	**0.9392**	**4.903**	ω_3	**0.2493**	**4.841**	δ_5	**–0.0312**	**–1.876**

Recall that α_1 was set equal to 1. The sequence of coefficients α_1, α_2, ..., α_{44} comprise our estimated quarterly commercial office building price index for the land component of property value. It can be seen that this land price index is quite volatile due to the sparseness of commercial property sales and the heterogeneity of the properties. In a subsequent section, we will show how this volatile land price index can be smoothed in a fairly simple fashion.

Turning to the other estimated coefficients, it can be seen that the ward relative land price parameters, ω_1-ω_4, decline (substantially) in magnitude as we move from the first more expensive composite ward to the less expensive composite wards. The marginal value of land parameters, λ_1 (set equal to 1), λ_2, λ_3 and λ_4, exhibit the same inverted U pattern that emerged in Model 3 (and persisted through all of the subsequent models). The excess land parameters, χ_1 (set equal to 1), χ_2, χ_3, χ_4 and χ_5, show that excess land is generally valued less than footprint land but the decline in land value as excess land increases is not monotonic. The building height land parameter $\mu = 0.0602$ is no longer as large as it was in Model 5 but an extra story of building height still adds 6% to the land value of the structure which is a significant premium for extra building stories. The walking distance to the nearest subway station parameter $\eta = -0.00023$ seems small but it tells us if the property is 1000 meters away from the nearest station, then the land value of the property is expected to fall by 23% compared to a nearby property. The travel time to Tokyo station parameter $\theta = 0.0209$ has a counterintuitive sign; it is possible that this variable is correlated with other land price determining characteristics and hence is not reliably determined. The height parameters, $\phi_3 = 1$ and ϕ_4-ϕ_{10}, are very significant determinants of structure value; the value of the structure increases almost monotonically as the number of stories increases. Finally, the decade by decade estimated geometric depreciation rates, δ_1-δ_5, show much the same pattern as was shown by the results for the previous model. Overall, the results of Model 9 seem to be reasonable.

In the following section, we will see if changing the model of depreciation from a multiple geometric depreciation rates model to a piece-wise linear model of depreciation leads to a significant change in our estimated land price index.

5. The Builder's Model with Piece-Wise Linear Depreciation Rates

Thus far, we have assumed that geometric depreciation models can best describe our data. In this section, we check the robustness of our results by assuming alternative depreciation models.

Recall that the structure aging (or survival) function for Model 9, $g_A(A_{tn})$, was defined by (20) above. In this section, we switch from a geometric model of depreciation to a straight line or linear depreciation model. Thus for **Model 10**, we defined the aging function as follows:

$$(23)\ g_A(A_{tn}) \equiv (1-\delta A_{tn})$$

where δ is the straight line depreciation rate. Our new nonlinear regression model is the same as the previous model defined by equations (22) except that the function g_A is defined by (23). The starting parameter values were taken to be the final parameter values from Model 7 except that the initial δ was set equal to 0.01 and the initial values for the parameters η and θ were set equal to 0.

The final log likelihood for Model 10 was −12635.83 and the R^2 was 0.7078. The estimated straight line depreciation rate was $\delta = 0.01357$ (0.00127). This model generated reasonable parameter estimates and the imputed value of the structure component of property value was positive for all observation.[26]

The straight line model of depreciation is not very flexible. Thus following the approach used by Diewert and Shimizu (2015b), we implement a piece-wise linear depreciation model. Recall definitions (19) above which defined the 5 age dummy variables, $D_{A,tni}$, for $i = 1,...,5$. We use these age dummy variables to define the piece-wise linear aging function, $g_A(A_{tn})$, as follows:

$$\begin{aligned}(24)\ g_A(A_{tn}) \equiv\ & D_{A,tn1}(1-\delta_1 A_{tn}) + D_{A,tn2}(1-10\delta_1-\delta_2(A_{tn}-10)) \\ & + D_{A,tn3}(1-10\delta_1-10\delta_2-\delta_3(A_{tn}-20)) + D_{A,tn4}(1-10\delta_1-10\delta_2-10\delta_3-\delta_4(A_{tn}-30)) \\ & + D_{A,tn5}(1-10\delta_1-10\delta_2-10\delta_3-10\delta_4-\delta_5(A_{tn}-40)).\end{aligned}$$

The **Model 11** nonlinear regression model is the same as the model defined by equations (22) except that the function g_A is defined by (24). The starting parameter values were taken to be the final parameter values from Model 10 except that the new depreciation

[26] This does not always happen for straight line depreciation models; i.e., for properties with very old structures, the imputed value of the structure can become negative if the estimated depreciation rate is large enough. This phenomenon cannot occur with geometric depreciation models, which is an advantage of assuming this form of depreciation.

parameters $\delta_1,...,\delta_5$ were all set equal to the final straight line depreciation rate δ estimated in Model 10. If all 5 δ_i are set equal to a common δ, then Model 11 collapses down to Model 10.

The final log likelihood for Model 11 was −12614.35, which was an increase in log likelihood of 21.48 over the Model 10 log likelihood. The R^2 for Model 11 was 0.7143.[27] The estimated piecewise linear depreciation rates (with standard errors in brackets) were as follows: δ_1 = 0.0393 (0.0057), δ_2 = 0.0125 (0.0049), δ_3 = 0.0302 (0.0041),[28] δ_4 = 0.0159 (0.0054), δ_5 = −0.0135 (0.0074). Thus as was the case with the multiple geometric depreciation rate model, properties with structures which are over 40 years old tended to *increase* in value by 1.35% per year. All of the estimated parameter coefficients for Model 11 and their t statistics are listed in Table 4 below.

Table 4: Estimated Coefficients for Model 11

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
α_2	**1.5529**	**7.227**	α_{25}	**0.7911**	**3.552**	ω_4	**0.1211**	**3.669**
α_3	**1.6342**	**8.878**	α_{26}	**0.8702**	**3.090**	λ_2	**1.4159**	**4.134**
α_4	**1.5352**	**8.846**	α_{27}	**1.2100**	**4.945**	λ_3	**1.6789**	**8.081**
α_5	**1.4210**	**6.373**	α_{28}	**1.2474**	**7.587**	λ_4	**0.5667**	**4.002**
α_6	**1.7646**	**7.895**	α_{29}	**1.3276**	**6.854**	χ_2	**0.9090**	**23.966**
α_7	**2.1559**	**7.705**	α_{30}	**1.0091**	**3.855**	χ_3	**0.7654**	**17.928**
α_8	**1.7560**	**7.255**	α_{31}	**1.1229**	**5.239**	χ_4	**0.8313**	**15.173**
α_9	**2.2680**	**7.632**	α_{32}	**1.5835**	**6.670**	χ_5	**0.9644**	**9.593**
α_{10}	**2.6239**	**7.483**	α_{33}	**1.1892**	**5.837**	μ	**0.0595**	**2.671**
α_{11}	**2.4706**	**7.299**	α_{34}	**1.4767**	**5.744**	η	**−0.0002**	**−3.973**
α_{12}	**2.4236**	**7.196**	α_{35}	**1.4063**	**7.278**	θ	**0.0212**	**3.972**
α_{13}	**2.4715**	**7.362**	α_{36}	**1.5521**	**6.370**	ϕ_4	**1.2947**	**6.828**
α_{14}	**1.7192**	**5.818**	α_{37}	**1.6954**	**7.796**	ϕ_5	**1.6571**	**10.556**
α_{15}	**1.7004**	**7.001**	α_{38}	**1.3819**	**6.117**	ϕ_6	**1.6872**	**10.906**
α_{16}	**2.0585**	**7.234**	α_{39}	**1.7117**	**7.498**	ϕ_7	**1.6708**	**10.098**
α_{17}	**1.4660**	**6.624**	α_{40}	**1.9944**	**7.622**	ϕ_8	**2.3273**	**13.861**
α_{18}	**0.8829**	**3.554**	α_{41}	**1.3680**	**6.354**	ϕ_9	**2.5212**	**14.271**
α_{19}	**0.6842**	**3.258**	α_{42}	**1.6344**	**6.043**	ϕ_{10}	**2.7717**	**13.774**
α_{20}	**1.1944**	**5.085**	α_{43}	**1.5117**	**6.324**	δ_1	**0.0393**	**6.920**
α_{21}	**0.9789**	**5.400**	α_{44}	**2.0454**	**6.897**	δ_2	**0.0125**	**2.524**
α_{22}	**1.1714**	**5.079**	ω_1	**0.6493**	**6.794**	δ_3	**0.0030**	**0.737**

[27] Recall that the log likelihood for the comparable geometric model of depreciation, Model 9, was −12614.70 and the R^2 for Model 9 was 0.7142. Thus the descriptive power of both models is virtually identical.

[28] Recall that these depreciation rates are net depreciation rates. As surviving structures approach their middle age, renovations become important and thus a decline in the net depreciation rate is plausible. The pattern of depreciation rates is similar to the comparable geometric depreciation rates that were observed for Richmond (a suburb of Vancouver, Canada) detached houses by Diewert, Huang and Burnett-Isaacs (2017).

α_{23}	**1.2619**	**5.899**	ω_2	**0.3460**	**5.720**	δ_4	**0.0159**	**2.962**
α_{24}	**0.9390**	**5.026**	ω_3	**0.2473**	**5.061**	δ_5	**−0.0135**	**−1.826**

Comparing the estimated coefficients in Tables 3 and 4, it can be seen that the parameter estimates for Models 9 and 11 were very similar except that there were some differences in the estimated depreciation rates δ_1 to δ_5. However, in the following section, we will show that these two multiple depreciation rate models generate aging functions g_A that approximate each other reasonably well. Thus both models describe the underlying data to the same degree of approximation.

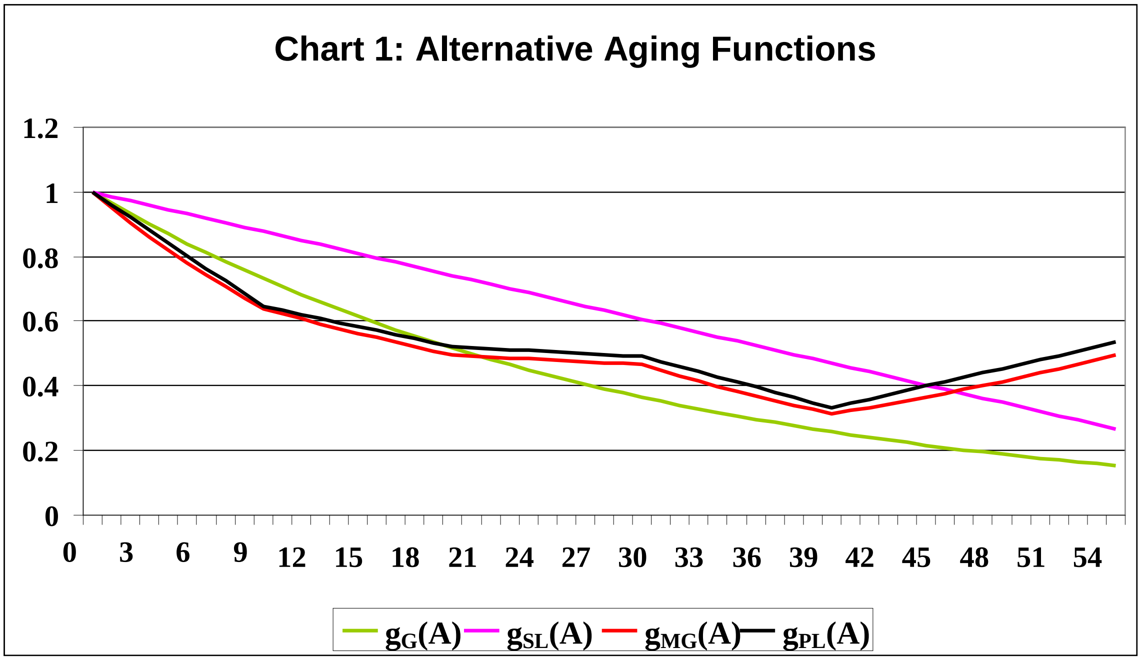
6. Comparing Alternative Models of Depreciation

The determination of depreciation schedules for commercial office buildings is important for tax purposes, for investors and for the estimation of commercial office structure stocks, which in turn feed into the computation of the Multifactor Productivity of the Commercial Office Sector. Thus in this section, we compare the single geometric rate (Model 7), the straight line (Model 10), the multiple geometric rate (Model 8) and piece-wise linear (Model 11) depreciation schedules. These depreciation schedules are equal to the ageing functions $g_A(A)$ defined by $g_G(A) \equiv (1-\delta)^A$, $g_{SL}(A) \equiv (1-\delta A)$, $g_{MG}(A)$ where g_{MG} is equal to g_A defined by (20) and $g_{PL}(A)$ where g_{PL} is is the g_A defined by (24) and the age variable A = 0,1,2,...,54. The resulting depreciation schedules are listed in Table 5 and plotted on Chart 1.

Table 5: Geometric, Straight Line, Multiple Geometric and Piece-Wise Linear Ageing Functions

A	$g_G(A)$	$g_{SL}(A)$	$g_{MG}(A)$	$g_{PL}(A)$
0	1.0000	1.0000	1.0000	1.0000
1	0.9659	0.9864	0.9516	0.9607
2	0.9329	0.9729	0.9055	0.9214
3	0.9011	0.9593	0.8617	0.8821
4	0.8703	0.9457	0.8200	0.8427
5	0.8406	0.9321	0.7803	0.8034
6	0.8119	0.9186	0.7425	0.7641
7	0.7842	0.9050	0.7066	0.7248
8	0.7574	0.8914	0.6724	0.6855
9	0.7316	0.8779	0.6398	0.6461
10	0.7066	0.8643	0.6237	0.6337
11	0.6825	0.8507	0.6080	0.6212
12	0.6592	0.8371	0.5927	0.6087
13	0.6367	0.8236	0.5777	0.5962
14	0.6150	0.8100	0.5632	0.5838
15	0.5940	0.7964	0.5490	0.5713
16	0.5737	0.7829	0.5352	0.5588
17	0.5541	0.7693	0.5217	0.5463
18	0.5352	0.7557	0.5085	0.5339
19	0.5170	0.7421	0.4957	0.5214

20	**0.4993**	**0.7286**	**0.4927**	**0.5184**
21	**0.4823**	**0.7150**	**0.4898**	**0.5153**
22	**0.4658**	**0.7014**	**0.4868**	**0.5123**
23	**0.4499**	**0.6878**	**0.4839**	**0.5093**
24	**0.4346**	**0.6743**	**0.4810**	**0.5063**
25	**0.4197**	**0.6607**	**0.4781**	**0.5032**
26	**0.4054**	**0.6471**	**0.4752**	**0.5002**
27	**0.3916**	**0.6336**	**0.4723**	**0.4972**
28	**0.3782**	**0.6200**	**0.4695**	**0.4942**
29	**0.3653**	**0.6064**	**0.4667**	**0.4912**
30	**0.3528**	**0.5928**	**0.4485**	**0.4753**
31	**0.3408**	**0.5793**	**0.4311**	**0.4594**
32	**0.3291**	**0.5657**	**0.4143**	**0.4435**
33	**0.3179**	**0.5521**	**0.3982**	**0.4276**
34	**0.3071**	**0.5386**	**0.3827**	**0.4117**
35	**0.2966**	**0.5250**	**0.3678**	**0.3958**
36	**0.2865**	**0.5114**	**0.3535**	**0.3799**
37	**0.2767**	**0.4978**	**0.3397**	**0.3640**
38	**0.2672**	**0.4843**	**0.3265**	**0.3481**
39	**0.2581**	**0.4707**	**0.3138**	**0.3323**
40	**0.2493**	**0.4571**	**0.3236**	**0.3457**
41	**0.2408**	**0.4436**	**0.3336**	**0.3592**
42	**0.2326**	**0.4300**	**0.3441**	**0.3727**
43	**0.2246**	**0.4164**	**0.3548**	**0.3862**
44	**0.2170**	**0.4028**	**0.3659**	**0.3997**
45	**0.2096**	**0.3893**	**0.3773**	**0.4132**
46	**0.2024**	**0.3757**	**0.3890**	**0.4266**
47	**0.1955**	**0.3621**	**0.4012**	**0.4401**
48	**0.1888**	**0.3485**	**0.4137**	**0.4536**
49	**0.1824**	**0.3350**	**0.4266**	**0.4671**
50	**0.1762**	**0.3214**	**0.4399**	**0.4806**
51	**0.1702**	**0.3078**	**0.4536**	**0.4941**
52	**0.1643**	**0.2943**	**0.4678**	**0.5076**
53	**0.1587**	**0.2807**	**0.4824**	**0.5210**
54	**0.1533**	**0.2671**	**0.4974**	**0.5345**



The straight line depreciation schedule is represented by the aging function $g_{SL}(A)$; it is the straight line in Chart 1. The depreciation schedule for the geometric model of depreciation is represented by the convex curved line in Chart 1. It can be seen that these single rate depreciation schedules are rather different. The multiple geometric rate depreciation schedule is the lower of the two broken lines in Chart 1 while the piece-wise linear depreciation schedule is the slightly higher broken line. It can be seen that these two multiple depreciation rate schedules approximate each other fairly well.[29] It can also be seen that the single geometric rate depreciation schedule provides a rough approximation to the two multiple rate schedules up to age 40 but then the schedules diverge.

The sequence of parameters α_t for t = 2,3,...,44 (along with $\alpha_1 \equiv 1$) listed in Tables 3 and 4 above provide alternative land price indexes generated by the MLIT transaction data. It can be seen that these alternative indexes are virtually identical (they cannot be distinguished on a chart) and hence only one of these alternative models of depreciation needs to be considered in what follows. Since the log likelihood of the piece-wise linear depreciation model (Model 11) was slightly higher than the multiple geometric depreciation rate model (Model 10), we will use the α_t sequence generated by Model 11 as our MLIT land price series in subsequent sections. We will label this series for quarter t as PL_{MLIT}^t.

7. Smoothing the MLIT Land Price Series

[29] This is to be expected. As the number of separate depreciation rates in each of these models tends to 43, the two schedules will converge to a common schedule.

In Chart 2 below, it can be seen that our Model 11 estimated land price series, $PL_{MLIT}{}^{t} \equiv \alpha_t$, is extremely volatile. This is due to the fact that commercial properties are very heterogeneous and we have relatively few transactions per quarter. Thus the raw series PL_{MLIT} does not accurately represent the *trend* in commercial land prices in Tokyo; the raw series requires some smoothing in order to model the trends in land prices. Patrick (2017) found the same problem for Irish house price sales and we will follow his example and smooth the raw series.[30]

We used the LOWESS nonparametric smoother in Shazam in order to construct a preliminary smoothed land price series, PL_S, using PL_{MLIT} as the input series.[31] We used the cross-validation criterion to choose the smoothing parameter which turned out to be f = 0.12. The series PL_{MLIT} and PL_S are listed in Table 6 and plotted in Chart 2 below.

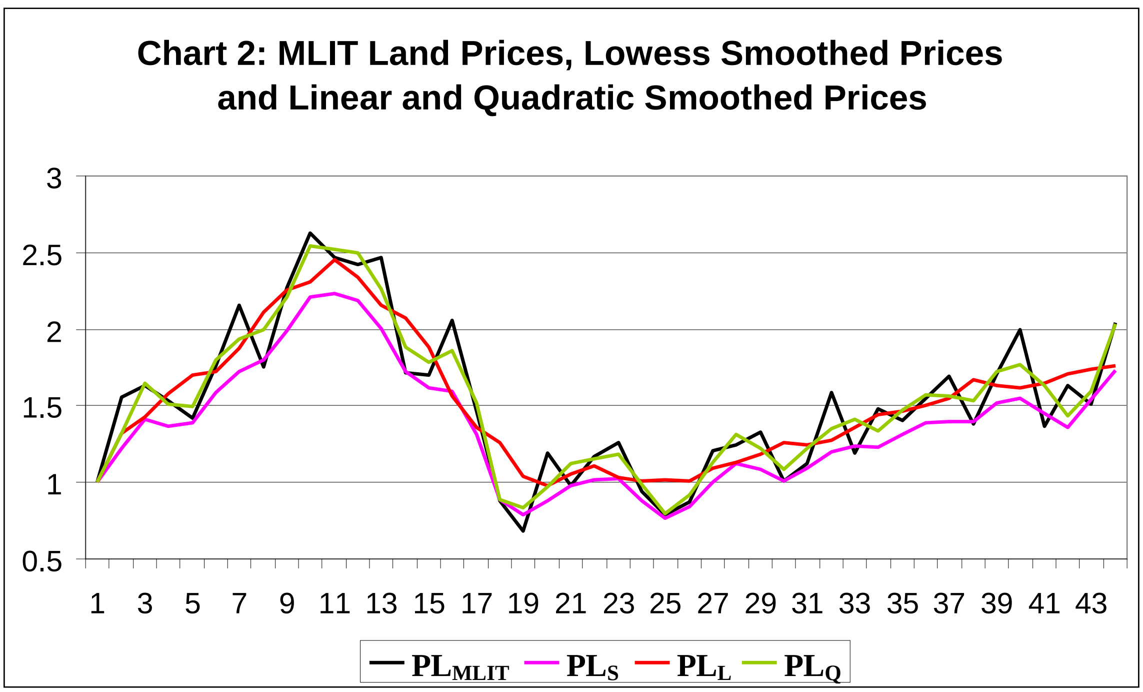
Table 6: MLIT Land Prices PL_{MLIT}, Lowess Smoothed Land Prices PL_S and Linear and Quadratic Smoothed Land Prices PL_L and PL_Q

Quarter	PL_{MLIT}	PL_S	PL_L	PL_Q
1	1.00000	1.00000	1.00000	1.00000
2	1.55293	1.22256	1.31711	1.31711
3	1.63422	1.41343	1.42867	1.64505
4	1.53523	1.36721	1.58159	1.50881
5	1.42096	1.38616	1.70218	1.49669
6	1.76462	1.58816	1.72654	1.80134
7	2.15588	1.72313	1.87309	1.93802
8	1.75601	1.80146	2.11368	1.99351
9	2.26798	1.98565	2.25488	2.20673
10	2.62393	2.21322	2.30843	2.54089
11	2.47061	2.23370	2.45153	2.52436
12	2.42362	2.18881	2.34178	2.49936
13	2.47153	2.00272	2.15709	2.26334
14	1.71923	1.72301	2.07466	1.88127
15	1.70045	1.61503	1.88313	1.78365
16	2.05848	1.59538	1.56540	1.86242
17	1.46597	1.31167	1.35840	1.51611
18	0.88287	0.88680	1.25719	0.88721
19	0.68422	0.79234	1.04127	0.83499
20	1.19442	0.88117	0.98237	0.97427
21	0.97889	0.97860	1.05818	1.11980
22	1.17144	1.02013	1.10914	1.15441

[30] Patrick initially smoothed his series by taking a three month rolling average of the raw index prices for Ireland. He found that the resulting index was still too volatile to publish and he ended up using a double exponential smoothing procedure.

[31] The initial smoothed series was divided by the Quarter 1 value so that the resulting normalized series equalled 1 in Quarter 1. Recall that Quarter 1 is the first quarter in 2005 and Quarter 44 is the last quarter in 2015.

23	**1.26194**	**1.02133**	**1.02848**	**1.18481**
24	**0.93901**	**0.88339**	**1.00673**	**0.98500**
25	**0.79111**	**0.76458**	**1.01445**	**0.79266**
26	**0.87016**	**0.84384**	**1.01155**	**0.92135**
27	**1.21003**	**1.00334**	**1.08928**	**1.13215**
28	**1.24743**	**1.12503**	**1.13287**	**1.31487**
29	**1.32764**	**1.08376**	**1.18341**	**1.21856**
30	**1.00910**	**1.01178**	**1.25811**	**1.08766**
31	**1.12286**	**1.09153**	**1.24647**	**1.21854**
32	**1.58349**	**1.19563**	**1.27629**	**1.34877**
33	**1.18925**	**1.23645**	**1.35573**	**1.41007**
34	**1.47675**	**1.22737**	**1.44159**	**1.33842**
35	**1.40632**	**1.31129**	**1.46396**	**1.47429**
36	**1.55214**	**1.38556**	**1.50250**	**1.57230**
37	**1.69536**	**1.39737**	**1.54949**	**1.56217**
38	**1.38194**	**1.39867**	**1.66709**	**1.53537**
39	**1.71167**	**1.51667**	**1.63026**	**1.72640**
40	**1.99436**	**1.54872**	**1.61806**	**1.76603**
41	**1.36798**	**1.45003**	**1.64401**	**1.63230**
42	**1.63437**	**1.36017**	**1.71076**	**1.43488**
43	**1.51167**	**1.54133**	**1.73534**	**1.59740**
44	**2.04541**	**1.73013**	**1.75991**	**2.03579**



The jagged black line in Chart 2 represents the unsmoothed land price index PL_{MLIT} that we estimated from Model 11 while the lowest line represents the Lowess nonparametric smoothed series PL_S that was generated using Shazam. It can be seen that while PL_S is reasonably smooth, it is not quite centered; i.e., it is consistently below the raw series. Thus we considered some alternative methods for smoothing the raw series.

Henderson (1916) was the first to realize that various moving average smoothers could be related to rolling window least squares regressions that would exactly reproduce a polynomial curve. Thus we apply his idea to derive the moving average weights that would be equivalent to fitting a linear function to 5 consecutive quarters of a time series, which we represent by the vector $Y^T \equiv [y_1,...,y_5]$ where Y^T denotes the transpose of a vector Y. Define the 5 dimensional column vectors X_1 and X_2 as $X_1 \equiv [1,1,1,1,1]^T$ and $X_2 \equiv [-2,-1.0,1,2]^T$. Define the 5 by 2 dimensional X matrix as $X \equiv [X_1,X_2]$. Denote the linear smooth of the vector Y by Y^*. Then least squares theory tells us that $Y^* = X(X^TX)^{-1}X^TY$. Thus the 5 rows of the 5 by 5 projection matrix $X(X^TX)^{-1}X^T$ give us the weights that can be used to convert the raw Y series into the smoothed Y^* series. For our particular example, the 5 rows of the projection matrix are as follows: Row 1 = (1/10)[6,4,2,0,−2]; Row 2 = (1/10)[4,3,2,1,0]; Row 3 = (1/5)[1,1,1,1,1]; Row 4 = (1/10)[0,1,2,3,4]; Row 5 = (1/10)[−2,0,2,4,6]. Note that Row 3 tells us that the third component of the smoothed vector Y^* is equal to $y_3^* = (1/5)(y_1+y_2+y_3+y_4+y_5)$. a simple equally weighted moving average of the raw data for 5 periods. Thus the way this smoothing method could be applied in practice to 44 consecutive quarters of PL_{MLIT} data is as follows. The first 3 components of the smoothed series would use the inner products of the first 3 rows of the projection matrix $X(X^TX)^{-1}X^T$ times the first 5 components of the PL_{MLIT} series. This would generate the first 3 components of the smoothed series, PL_L^t for t = 1,2,3. For t = 3,4,...,42, define $PL_L^t \equiv (1/5)[PL_{MLIT}^{t-2} + PL_{MLIT}^{t-1} + PL_{MLIT}^t + PL_{MLIT}^{t+1} + PL_{MLIT}^{t+2}]$. Thus for all observations t except for the first two and last two observations, the smoothed series PL_t would be defined as the simple centered moving average of 5 consecutive PL_{MLIT} observations with equal weights. The final two observations would be defined as the inner products of Rows 4 and 5 of $X(X^TX)^{-1}X^T$ with the last 5 observations in the PL_{MLIT} series. In practice, as the data of a subsequent period became available, the last two observations in the existing series would be revised but after receiving the data of two subsequent periods, there would be no further revisions; i.e., the final smoothed value of an observation would be set equal to the centered 5 period moving average of the raw data.

We implemented the above procedure but the above algorithm does not ensure that the value of the smoothed series in the first quarter of the sample is equal to 1 and so the generated series had to be divided by a constant to ensure that the first observation in the smoothed series is equal to unity. We found that this division caused the smoothed series to lie below the raw series for the most part.[32] Patrick (2017; 25-26) found that a similar problem occurred with his initial simple moving average smoothing method. He solved the problem by setting the smoothed values equal to the actual values for the first two observations when he applied his second smoothing method. We solved the centering problem in a similar manner: we set the initial value of the smooth equal to the

[32] A similar problem of a lack of centering occurred when we implemented the Lowess smoothing procedure; i.e., we had to divide by a constant to make the first component of the smoothed series equal to one. As a result, the Lowess smooth tended to lie below the raw series as can be seen in Chart 2.

corresponding raw number (so that $PL_L^1 \equiv PL_{MLIT}^1$) and we set the second value of the smooth equal to the average of the first and third observations in the raw series (so that $PL_L^2 \equiv (1/2)[PL_{MLIT}^1 + PL_{MLIT}^3]$. For the Quarter 3 value of the smooth, we used the simple 5 term centered moving average so that $PL_L^3 \equiv (1/5)[PL_{MLIT}^1 + PL_{MLIT}^2 + PL_{MLIT}^3 + PL_{MLIT}^4 + PL_{MLIT}^5]$ and we carried on using this moving average until Quarters 43 and 44 where we used Rows 4 and 5 of the matrix $X(X^TX)^{-1}X^T$ defined above for our moving average weights. The resulting smoothed series PL_L^t is listed in Table 6 and plotted in Chart 2 above. It can be seen that it does a good job of smoothing the initial PL_{MLIT}^t series.

We also applied the same least squares methodology to a rolling window 5 term quadratic regression model. Define the 5 dimensional column vectors X_1 and X_2 as before and define $X_3 \equiv [4,1,0,1,4]^T$. Define the 5 by 3 dimensional X matrix as $X \equiv [X_1,X_2.X_3]$. Denote the quadratic smooth of the vector Y by Y^{**}. Again least squares theory tells us that $Y^{**} = X(X^TX)^{-1}X^TY$. The 5 rows of the new 5 by 5 projection matrix $X(X^TX)^{-1}X^T$ give us the weights that can be used to convert the raw Y series into the smoothed Y^{**} series. The 5 rows of the new projection matrix are as follows: Row 1 = (1/35)[31,9,–3,–5,3]; Row 2 = (1/35)[9,13,12,6,–5]; Row 3 = (1/35)[–3,12,17,12,–3]; Row 4 = (1/35)[,–5,6,12,13,9]; Row 5 = (1/35)[3,–5,–3,9,31]. Now repeat the steps that were used to construct the linear smooth PL_L^t to construct a preliminary quadratic smooth PL_Q^t, except that the new 5 by 5 projection matrix $X(X^TX)^{-1}X^T$ replaces the previous one. A final PL_Q^t series was constructed by replacing the first 2 values in the smoothed series by the same initial 2 values that we used to construct the final versions of PL_L^1 and PL_L^2. The resulting smoothed series PL_Q^t is listed in Table 6 and plotted in Chart 2 above. It can be seen that PL_Q^t is not nearly as smooth as the linear smoothed series PL_L^t but of course, it is a lot closer to the unadjusted series PL_{MLIT}^t. For our particular data set, we would recommend the linear smoother over the quadratic smoother.[33]

We turn now to the construction of land prices using commercial property appraisal data.

8. The Builder's Model Using Property Appraisal Data

As was indicated in Section 2 above, we have quarterly appraisal data for 41 commercial office REIT office buildings located in Tokyo for the 44 quarters starting at Q1:2005 and ending at Q4:2015, which of course, is the same period that was covered by the MLIT selling price data. We will implement the builder's model for this data set in this section.

The builder's model using appraisal data is somewhat different from the builder's model using selling price data. The panel nature of the REIT data means that we can use a single property specific dummy variable as a variable that concentrates all of the location attributes of the property into a single variable; i.e., we do not have to worry about how

[33] A quadratic Henderson type smoother would be much smoother if we lengthened the window. But a longer window would imply a longer revision period before the series would be finalized. Since the linear smoother with window length 5 seems to do a nice job of smoothing, we would not recommend moving to a longer window length for this particular application.

close to a subway line the property is or how many stories the building has or how much excess land is associated with the property. The single property specific dummy variable will take all of these characteristics into account.

There are 41 separate properties in our REIT data set. For each of our 44 quarters, we assume that the 41 properties appear in the appraised property value for property n in period t, V_{tn}, in the same order. Our initial regression model is the following one where the variables have the same definitions as in equations (2) above except that ω_n is now the *property n sample average land price* (per m^2) rather than a Ward n relative price of land:

(25) $V_{tn} = \sum_{n=1}^{41} \omega_n L_{tn} + p_{St}(1 - 0.025)^{A(t,n)} S_{tn} + \varepsilon_{tn}$; $\quad t = 1,...,44; n = 1,...,41.$

Thus in **Model 1** above, there are no quarter t land price parameters in this very simple model with 41 unknown property average land price ω_n parameters to estimate. Note that the geometric (net) depreciation rate in the model defined by (25) was assumed to be 2.5% per year.

The final log likelihood for this model was −14968.77 and the R^2 was 0.9426. Thus the 41 property average price parameters ω_n explain a large part of the variation in the data.

In **Model 2**, we introduce quarterly land prices α_t into the above model. The new nonlinear regression model is the following one:

(26) $V_{tn} = \sum_{n=1}^{41} \alpha_t \omega_n L_{tn} + p_{St}(1 - 0.025)^{A(t,n)} S_{tn} + \varepsilon_{tn}$; $\quad t = 1,...,44; n = 1,...,41.$

Not all of the quarterly land price parameters (the α_t) and the average property price parameters (the ω_n) can be identified. Thus we impose the following normalization on our coefficients:

(27) $\alpha_1 = 1$.

We used the final parameter values for the ω_n from Model 1 as starting coefficient values for Model 2 (with all α_t initially set equal to 1).[34] The final log likelihood for Model 2 was −13999.00, a huge improvement of 969.77 for adding 43 new parameters. The R^2 was 0.9804. Thus the 41 property average price parameters ω_n and the 43 quarterly average land price parameters α_t explain most of the variation in the data.

---

[34] The reader may well wonder why we estimated the ω_n in Model 1 rather than first estimating the α_t in Model 1. When this alternative strategy was implemented, we found that the resulting Model 2 did not converge to the "right" parameter values; i.e., the resulting R^2 was very low. This is the reason for following our nested estimation methodology where each successive model uses the final coefficient values from the previous model. It is not possible to simply estimate our final models in one step and obtain sensible results.

Model 3 is the following nonlinear regression model:

(28) $V_{tn} = \alpha_t\omega_n L_{tn} + p_{St}(1-\delta)^{A(t,n)}S_{tn} + \varepsilon_{tn}$; $t = 1,...,44; n = 1,...,41$

where δ is the annual geometric (net) depreciation rate. The normalization (27) is also imposed. Thus Model 3 is the same as Model 2 except that we now estimate the single geometric depreciation rate δ.

We used the final parameter values for the α_t and ω_n from Model 2 as starting coefficient values for Model 3 (with δ initially set equal to 0.025). The final log likelihood for this model was −13993.47, and increase of 5.53 for one additional parameter, and the R^2 was 0.9806. The estimated geometric (net) depreciation rate was $\delta = 0.01353$.[35] The estimated coefficients and their t statistics are listed in Table 7. Recall that α_1 was set equal to 1. The sequence of land price (per m^2) α_t, for t =1,2,...,44 is our estimated sequence of quarterly Tokyo land prices, $PL_{REIT}{}^t$, which appears in Chart 3 below.

Table 7: Estimated Coefficients for Model 3 Using REIT Data

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
α_2	1.0268	2.121	α_{31}	1.0645	1.403	ω_{15}	2.1471	10.600
α_3	1.0637	1.479	α_{32}	1.0594	1.429	ω_{16}	5.8157	44.703
α_4	1.1045	1.541	α_{33}	1.0498	1.466	ω_{17}	5.8961	40.816
α_5	1.1499	1.532	α_{34}	1.0393	1.334	ω_{18}	4.0615	34.128
α_6	1.1987	1.607	α_{35}	1.0341	1.380	ω_{19}	5.4266	39.534
α_7	1.2473	1.569	α_{36}	1.0294	1.242	ω_{20}	5.7298	36.905
α_8	1.2994	1.636	α_{37}	1.0277	1.226	ω_{21}	1.0098	40.629
α_9	1.3450	1.735	α_{38}	1.0299	1.200	ω_{22}	4.0731	47.336
α_{10}	1.3882	1.904	α_{39}	1.0350	1.253	ω_{23}	2.0521	25.603
α_{11}	1.4422	2.027	α_{40}	1.0425	1.315	ω_{24}	2.5844	38.974
α_{12}	1.4904	2.052	α_{41}	1.0564	1.270	ω_{25}	1.0869	41.419
α_{13}	1.5082	1.995	α_{42}	1.0700	1.244	ω_{26}	1.2409	22.634
α_{14}	1.4990	2.075	α_{43}	1.0874	1.286	ω_{27}	2.0714	23.201
α_{15}	1.4751	2.103	α_{44}	1.1078	1.827	ω_{28}	0.7289	32.147
α_{16}	1.4419	2.238	ω_1	3.8704	31.383	ω_{29}	0.6271	7.422
α_{17}	1.3976	1.721	ω_2	4.8678	48.918	ω_{30}	3.1068	39.453
α_{18}	1.3423	1.838	ω_3	1.7514	27.642	ω_{31}	1.7773	32.149
α_{19}	1.2892	1.705	ω_4	2.3099	26.957	ω_{32}	5.8748	41.597
α_{20}	1.2428	1.522	ω_5	1.8451	27.751	ω_{33}	1.5201	20.558
α_{21}	1.2108	1.634	ω_6	3.7399	30.589	ω_{34}	3.4731	30.059
α_{22}	1.1766	1.583	ω_7	2.6487	32.409	ω_{35}	2.1225	27.539

[35] We also estimated the straight line depreciation model counterpart to Model 3. The resulting estimated straight line depreciation rate δ was equal to 0.01317 (t statistic = 45.73). The R^2 for this model was 0.9806 and the final log likelihood was −13989.83. The resulting land price series was very similar to the land price series generated by Model 3 above.

α_{23}	**1.1543**	**1.419**	ω_8	**3.2710**	**28.703**	ω_{36}	**6.2429**	**48.012**
α_{24}	**1.1375**	**1.485**	ω_9	**4.8665**	**47.654**	ω_{37}	**4.2053**	**22.829**
α_{25}	**1.1166**	**1.414**	ω_{10}	**4.9867**	**41.462**	ω_{38}	**2.6778**	**21.825**
α_{26}	**1.1007**	**1.479**	ω_{11}	**1.1427**	**16.089**	ω_{39}	**3.0139**	**23.805**
α_{27}	**1.0967**	**1.314**	ω_{12}	**2.3817**	**20.345**	ω_{40}	**2.9460**	**12.591**
α_{28}	**1.0908**	**1.401**	ω_{13}	**1.1255**	**15.765**	ω_{41}	**1.8028**	**15.349**
α_{29}	**1.0799**	**1.437**	ω_{14}	**0.8444**	**14.470**	δ	**0.0135**	**4.437**
α_{30}	**1.0683**	**1.349**						

Note that the implied standard errors on the quarterly land price coefficients, the α_t, are fairly large whereas they are fairly small for the property coefficients, the ω_n. This means that our estimated land price indexes, $PL_{REIT}{}^t = \alpha_t$, are not reliably determined. Note also that our estimated geometric depreciation rate δ is only 1.35% per year which is much lower than our estimated depreciation rate from Model 7 in Section 3 above which was 3.41% per year. One factor which may help to explain this divergence in estimates of wear and tear depreciation is that appraisers take into account capital expenditures on the properties. However, our current data base did not have information on capital expenditures and it is likely that not having capital expenditures as an explanatory factor affected our estimates for the depreciation rate. In our previous study of land prices using REIT data for Tokyo, Diewert and Shimizu (2017), we adjusted our nonlinear regressions for capital expenditures and found that the resulting estimated quarterly wear and tear geometric depreciation rate was 0.005 which implied an annual (single) geometric depreciation rate of about 2%.[36]

In the following section, we will estimate our final land price series for Tokyo commercial office buildings using official estimates for the land values of commercial properties for taxation purposes.

9. Estimating Land Prices for Commercial Properties using Tax Assessment Data

In this section, we will use the Official Land Price (OLP) data described in section 2 above. We have 6242 annual assessed values for the land components of commercial properties in Tokyo covering the 11 years 2005-2015. We will label these years as t = 1,2,...,11. The assessed land value for property n in year t is denoted as V_{tn}.[37] We have

[36] In the multiple geometric depreciation rate model estimated by Diewert and Shimizu (2017), the various rates averaged out to an annual rate of 2.6% per year. Our earlier study covered the 22 quarters starting at Q1 of 2007 and ending at Q2 of 2012. The correlation coefficient between the price of land series in this model in Diewert and Shimizu (2017) and the above Model 3 price of land series for the overlapping 22 quarters is 0.9901 so these two studies using REIT appraisal data show much the same trends in Tokyo commercial property land prices even though the estimated wear and tear depreciation rates are different. Note that in addition to wear and tear depreciation, depreciation due to the early demolition of a structure before it reaches “normal” retirement age should be taken into account. Our current study does not estimate this extra component of depreciation. However, Diewert and Shimizu (2017) estimated demolition depreciation for Tokyo commercial office buildings at 1.2% per year.

[37] The units of measurement used in this section are in 100,000 yen.

information on which Ward each property is located and the ward dummy variables $D_{W,tnj}$ are defined by definitions (4) above. The land plot area of property n in year t is denoted by L_{tn} and the subway variables DS_{tn} and TT_{tn} are defined as in section 2 above. The number of observations in year t is N(t).

Our initial regression model is the following one where we regress property land value on the ward dummy variables times the land plot area:

(29) $V_{tn} = (\sum_{j=1}^{23} \omega_j D_{W,tnj})L_{tn} + \varepsilon_{tn}$; $t = 1,...,11; n = 1,...,N(t).$

Thus in **Model 1** above, there are no year t land price parameters in this very simple model and ω_j is an estimate of the average land price (per m^2) in Ward j for j = 1,...,23. The final log likelihood for this model was –67073.91 and the R^2 was 0.3647.

In **Model 2**, we introduce annual land prices α_t into the above model. The new nonlinear regression model is the following one:

(30) $V_{tn} = \alpha_t(\sum_{j=1}^{23} \omega_j D_{W,tnj})L_{tn} + \varepsilon_{tn}$; $t = 1,...,11; n = 1,...,N(t).$

Not all of the 11 annual land price parameters (the α_t) and the 23 Ward average property relative price parameters (the ω_n) can be identified. Thus we impose the normalization $\alpha_1 = 1$.

We used the final parameter values for the ω_n from Model 1 as starting coefficient values for Model 2 (with all α_t initially set equal to 1). The final log likelihood for Model 2 was –67022.90, an increase of 51.01 for adding 43 new parameters. The R^2 was 0.3748.

In our next model, we allowed the price of land to vary as the lot size increased. We divided up our 6242 observations into 5 groups of observations based on their lot size. The Group 1 properties had lots less than 100 m^2, the Group 2 properties had lots greater than or equal to 100 m^2 and less than 150 m^2, the Group 3 properties had lots greater than or equal to 150 m^2 and less than 200 m^2, the Group 4 properties had lots greater than or equal to 200 m^2 and less than 300 m^2 and the Group 5 properties had lots greater than or equal to 300 m^2.[38] For each observation n in period t, we define the 5 *land dummy variables*, $D_{L,tnk}$, for k = 1,...,5 as follows:

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

Define the constants L_1-L_4 as 100, 150, 200 and 300 respectively. These constants and the dummy variables defined by (31) are used in the definition of the following piecewise linear function of L_{tn}, $f(L_{tn})$:

[38] The sample probabilities of an observation falling in the 5 land size groups were: 0.171, 0.285, 0.175, 0.178 and 0.191.

$$(32)\ f(L_{tn}) \equiv D_{L,tn1}\lambda_1 L_{tn} + D_{L,tn2}[\lambda_1 L_1+\lambda_2(L_{tn}-L_1)] + D_{L,tn3}[\lambda_1 L_1+\lambda_2(L_2-L_1)+\lambda_3(L_{tn}-L_2)]$$
$$+ D_{L,tn4}[\lambda_1 L_1+\lambda_2(L_2-L_1)+\lambda_3(L_3-L_2)+\lambda_4(L_{tn}-L_3)]$$
$$+ D_{L,tn5}[\lambda_1 L_1+\lambda_2(L_2-L_1)+\lambda_3(L_3-L_2)\ )+\lambda_4(L_4-L_3)+\lambda_5(L_{tn}-L_4)].$$

Model 3 was defined as the following nonlinear regression model:

$$(33)\ V_{tn} = \alpha_t(\textstyle\sum_{j=1}^{23} \omega_j D_{W,tn,j})f(L_{tn}) + \varepsilon_{tn}\ ; \qquad t = 1,...,11;\ n = 1,...,N(t).$$

We imposed the normalizations $\alpha_1 = 1$ and $\lambda_1 = 1$ so that all of the remaining parameters in (33) could be identified. These normalizations were also imposed in Model 4 below.

We used the final parameter values for the α_t and ω_j from Model 2 as starting coefficient values for Model 3 (with all λ_k initially set equal to 1). Thus Model 3 adds the 4 new marginal prices of land, λ_2, λ_3, λ_4 and λ_5 to Model 2. The final log likelihood for Model 3 was −66044.02, an increase of 978.88 for adding 4 new parameters. The R^2 was 0.4668.

Our final land price model added the subway variables to Model 3. Thus **Model 4** was defined as the following nonlinear regression model:[39]

$$(34)\ V_{tn} = \alpha_t(\textstyle\sum_{j=1}^{23} \omega_j D_{W,tnj})\ (1+\eta(DS_{tn}-50))(1+\theta(TT_{tn}-4))f(L_{tn}) + \varepsilon_{tn} \qquad t = 1,...,11;\ n = 1,...,N(t).$$

Thus Model 4 has added two new subway parameters, η and θ, to Model 3. We used the final parameter values for the α_t, ω_j and λ_k from Model 3 as starting coefficient values for Model 4 (with η and θ initially set equal to 0). The final log likelihood for Model 4 was −65584.56, an increase of 459.46 for adding 2 new parameters. The R^2 was 0.5401. The estimated coefficients for this model are listed in Table 8. The α_t sequence of estimated parameters (along with $\alpha_1 \equiv 1$) forms an annual (quality adjusted) Official Land Price series. For comparison purposes, we repeat each α_t four times and convert the annual Official Land Price series into the quarterly Official Land Price series, PL_{OLP}^t. It will be listed and compared with our final transactions based MLIT land price series PL_{MLIT}^t and its linear smooth PL_L^t along with our final REIT based land price series PL_{REIT}^t in the following section.

Table 8: Estimated Coefficients for Model 4 Using Annual Tax Assessment Data

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
ω_1	181.560	26.429	ω_{14}	172.950	14.037	α_5	1.2595	27.892
ω_2	164.660	25.788	ω_{15}	146.350	9.085	α_6	1.1865	27.452
ω_3	236.800	26.434	ω_{16}	213.590	21.443	α_7	1.1486	28.356
ω_4	263.190	24.031	ω_{17}	91.988	10.182	α_8	1.1382	27.710

[39] The minimum value for the distance to the nearest subway station DS_{tn} is 50 meters and the minimum value for the subway running time from the nearest station to the central Tokyo subway station was 4 minutes.

ω_5	**124.920**	**17.418**	ω_{18}	**83.365**	**17.515**	α_9	**1.1120**	**26.484**
ω_6	**126.740**	**22.236**	ω_{19}	**145.590**	**7.685**	α_{10}	**1.0919**	**25.122**
ω_7	**77.712**	**7.866**	ω_{20}	**193.350**	**6.975**	α_{11}	**1.1154**	**25.699**
ω_8	**84.417**	**8.496**	ω_{21}	**86.169**	**9.314**	λ_2	**0.7011**	**6.032**
ω_9	**137.330**	**18.302**	ω_{22}	**87.688**	**11.699**	λ_3	**–0.3331**	**–2.907**
ω_{10}	**230.320**	**14.959**	ω_{23}	**64.602**	**6.739**	λ_4	**0.3568**	**7.247**
ω_{11}	**101.550**	**12.494**	α_2	**1.0751**	**27.758**	λ_5	**0.1440**	**18.745**
ω_{12}	**195.970**	**13.212**	α_3	**1.1643**	**28.318**	η	**–0.000740**	**–27.937**
ω_{13}	**385.910**	**23.858**	α_4	**1.3399**	**28.916**	θ	**–0.022807**	**–45.349**

It can be seen that the standard errors on the estimated annual land prices α_t are fairly small; recall that they were fairly large for the RIET based quarterly land price series, PL_{REIT}^t. Except for λ_3, it can be seen that the λ_k monotonically decrease as k increases; this indicates that the marginal price of land decreases as the land plot size increases. The two estimated subway parameters, η and θ, both have the expected negative sign and are reasonable in magnitude. Since we do not have additional information on the height or size of the buildings, we cannot add more explanatory variables to the Model 4 regression.

10. Comparing Land Price Indexes from Different Sources

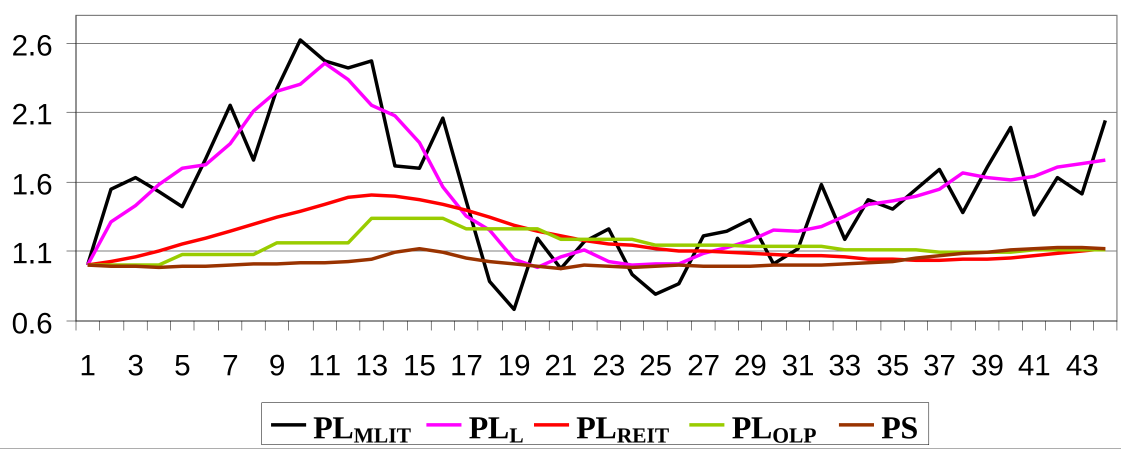
Table 6 lists the MLIT transactions based land price series PL_{MLIT}^t and its linear smooth, PL_L^t. Table 7 lists the REIT based land price series PL_{REIT}^t and the Official Land Price series PL_{OLP}^t can be constructed using the estimated α_t listed in Table 8. These 4 series along with the official construction price series P_{St} are listed in Table 9 and plotted in Chart 3.

Table 9: Alternative Land Price Series and the Price of Structures

Quarter t	PL_{MLIT}^t	PL_L^t	PL_{REIT}^t	PL_{OLP}^t	P_{St}
1	1.00000	1.00000	1.00000	1.00000	1.00000
2	1.55293	1.31711	1.02676	1.00000	0.99636
3	1.63422	1.42867	1.06369	1.00000	0.99211
4	1.53523	1.58159	1.10454	1.00000	0.98941
5	1.42096	1.70218	1.15004	1.07513	0.99184
6	1.76462	1.72654	1.19883	1.07513	0.99790
7	2.15588	1.87309	1.24744	1.07513	1.00351
8	1.75601	2.11368	1.29953	1.07513	1.00918
9	2.26798	2.25488	1.34499	1.16432	1.01241
10	2.62393	2.30843	1.38812	1.16432	1.01770
11	2.47061	2.45153	1.44194	1.16432	1.02208
12	2.42362	2.34178	1.49019	1.16432	1.02971
13	2.47153	2.15709	1.50818	1.33985	1.04572
14	1.71923	2.07466	1.49977	1.33985	1.09410
15	1.70045	1.88313	1.47635	1.33985	1.11708
16	2.05848	1.56540	1.44330	1.33985	1.09343
17	1.46597	1.35840	1.39905	1.25946	1.05114
18	0.88287	1.25719	1.34400	1.25946	1.02499

19	**0.68422**	**1.04127**	**1.29130**	**1.25946**	**1.01141**
20	**1.19442**	**0.98237**	**1.24531**	**1.25946**	**0.99425**
21	**0.97889**	**1.05818**	**1.21368**	**1.18646**	**0.98011**
22	**1.17144**	**1.10914**	**1.18021**	**1.18646**	**1.00021**
23	**1.26194**	**1.02848**	**1.15830**	**1.18646**	**0.99197**
24	**0.93901**	**1.00673**	**1.14173**	**1.18646**	**0.98385**
25	**0.79111**	**1.01445**	**1.12141**	**1.14862**	**0.99586**
26	**0.87016**	**1.01155**	**1.10598**	**1.14862**	**1.00424**
27	**1.21003**	**1.08928**	**1.10225**	**1.14862**	**0.99826**
28	**1.24743**	**1.13287**	**1.09666**	**1.14862**	**0.99692**
29	**1.32764**	**1.18341**	**1.08618**	**1.13820**	**0.99776**
30	**1.00910**	**1.25811**	**1.07504**	**1.13820**	**1.00624**
31	**1.12286**	**1.24647**	**1.07151**	**1.13820**	**1.00058**
32	**1.58349**	**1.27629**	**1.06681**	**1.13820**	**1.00290**
33	**1.18925**	**1.35573**	**1.05778**	**1.11199**	**1.01027**
34	**1.47675**	**1.44159**	**1.04788**	**1.11199**	**1.02160**
35	**1.40632**	**1.46396**	**1.04320**	**1.11199**	**1.02960**
36	**1.55214**	**1.50250**	**1.03916**	**1.11199**	**1.05012**
37	**1.69536**	**1.54949**	**1.03814**	**1.09194**	**1.07326**
38	**1.38194**	**1.66709**	**1.04095**	**1.09194**	**1.08818**
39	**1.71167**	**1.63026**	**1.04657**	**1.09194**	**1.09886**
40	**1.99436**	**1.61806**	**1.05460**	**1.09194**	**1.11577**
41	**1.36798**	**1.64401**	**1.06887**	**1.11544**	**1.12204**
42	**1.63437**	**1.71076**	**1.08289**	**1.11544**	**1.12769**
43	**1.51167**	**1.73534**	**1.10053**	**1.11544**	**1.12651**
44	**2.04541**	**1.75991**	**1.12109**	**1.11544**	**1.11855**

Chart 3: Alternative Land Price Series and the Price of Structures



It can be seen that the land price series based on transactions data, PL_{MLIT}^t and its linear smooth, PL_L^t, paint a very different picture of land price movements as compared to the series based on appraisal values for commercial land in Tokyo, PL_{REIT}^t, and the series based on property tax assessed values, PL_{OLP}^t. As was noted in section 1 above, appraisal prices tend to lag behind the movements in transaction prices and they also smooth the sales data. The same phenomenon evidently applies to assessed value prices. Chart 3 shows that the appraisal and assessed value based price indexes for commercial land fluctuate far less than the index based actual transactions prices. However, it can be seen that the appraisal and assessed value series do tend to move in the same direction as the transactions prices but with a lag. The Chart also shows the problem with the transactions based series: its quarter to quarter fluctuations are massive. But it also can be seen that the linear smoothed series PL_L^t (which is essentially a centered five quarter moving average of the unsmoothed series PL_{MLIT}^t) captures the trend in transactions prices quite well. This series can be finalized after a two quarter delay. Our preferred land price series is the linear smoothed transaction series PL_L^t.

In the following section, we will use the MLIT and REIT data to construct alternative commercial property price indexes; i.e., we will aggregate the land and structure price data into overall property price indexes and compare these indexes with other indexes which are simpler to construct.

11. A Comparison of Alternative Commercial Property Price Indexes

Recall that in Section 3 above, the MLIT value of property n in quarter t was defined as V_{tn} in period t and the corresponding property land and structure areas were defined as S_{tn} and L_{tn} for n = 1,...,N(t) and t = 1,...,44. In the property price literature, a frequently used index of overall property prices is the period average of the individual property values V_{tn} divided by the corresponding structure area S_{tn}. Thus define the (preliminary) quarter t mean property price P_{MEANP}^t as follows:

$$(35)\ P_{MEANP}^t \equiv (1/N(t))\Sigma_{n=1}^{N(t)} V_{tn}/S_{tn}\ ; \qquad t = 1,...,44.$$

The final mean property price index for quarter t, P_{MEAN}^t, is defined as the corresponding preliminary index P_{MEANP}^t divided by P_{MEANP}^1; i.e., we normalize the series defined by (35) to equal 1 in quarter 1.

Table 10: Alternative Overall Commercial Property Price Indexes

Quarter t	P_{MEAN}^t	P_{MEANS}^t	P_{FMLIT}^t	P_{FMLITS}^t	P_{FREIT}^t	P_{LPHED}^t	P_{LPHEDS}^t
1	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
2	1.18211	1.11644	1.25260	1.15058	1.01858	1.12578	1.12971
3	1.23289	1.16886	1.26945	1.18790	1.04447	1.25942	1.16157
4	1.22061	1.21225	1.21852	1.23585	1.07364	1.24404	1.20967
5	1.20868	1.24444	1.18141	1.28993	1.10761	1.17860	1.24030
6	1.21694	1.28112	1.33886	1.32389	1.14501	1.24052	1.27364
7	1.34307	1.30365	1.51703	1.40608	1.18219	1.27892	1.32785
8	1.41632	1.43708	1.33632	1.46794	1.22201	1.42612	1.42052

9	**1.33326**	**1.43627**	**1.56545**	**1.56057**	**1.25638**	**1.51511**	**1.45011**
10	**1.87582**	**1.46519**	**1.76501**	**1.63832**	**1.28961**	**1.64193**	**1.51268**
11	**1.21289**	**1.47075**	**1.64041**	**1.63363**	**1.33051**	**1.38846**	**1.54175**
12	**1.48766**	**1.44451**	**1.64183**	**1.61144**	**1.36814**	**1.59179**	**1.49237**
13	**1.44413**	**1.34393**	**1.66991**	**1.55087**	**1.38555**	**1.57148**	**1.43715**
14	**1.20203**	**1.36900**	**1.35804**	**1.48591**	**1.39181**	**1.26820**	**1.43887**
15	**1.37291**	**1.33842**	**1.37911**	**1.45152**	**1.38040**	**1.36580**	**1.35784**
16	**1.33824**	**1.25163**	**1.52799**	**1.32192**	**1.34982**	**1.39708**	**1.27652**
17	**1.33477**	**1.18789**	**1.23700**	**1.19171**	**1.30617**	**1.18664**	**1.21455**
18	**1.01018**	**1.10527**	**0.95988**	**1.12128**	**1.25869**	**1.16489**	**1.16196**
19	**0.88336**	**1.01512**	**0.86424**	**1.02333**	**1.21607**	**0.95832**	**1.08358**
20	**0.95981**	**0.92967**	**1.08210**	**0.98875**	**1.17747**	**1.10288**	**1.05599**
21	**0.88749**	**0.90702**	**0.97954**	**1.01588**	**1.15025**	**1.00516**	**1.03546**
22	**0.90749**	**0.90880**	**1.07696**	**1.04978**	**1.13033**	**1.04872**	**1.04410**
23	**0.89697**	**0.88131**	**1.11188**	**1.00911**	**1.11178**	**1.06220**	**1.02744**
24	**0.89223**	**0.88767**	**0.96497**	**0.99329**	**1.09728**	**1.00156**	**1.03771**
25	**0.82235**	**0.91494**	**0.90797**	**1.00329**	**1.08506**	**1.01955**	**1.06638**
26	**0.91929**	**0.93288**	**0.94799**	**1.00717**	**1.07552**	**1.05652**	**1.06417**
27	**1.04386**	**1.01361**	**1.09552**	**1.04126**	**1.07103**	**1.19205**	**1.11638**
28	**0.98669**	**1.06931**	**1.10571**	**1.05758**	**1.06634**	**1.05119**	**1.15239**
29	**1.29586**	**1.08824**	**1.14663**	**1.08420**	**1.05854**	**1.26258**	**1.16835**
30	**1.10084**	**1.08174**	**1.00751**	**1.11102**	**1.05216**	**1.19961**	**1.15112**
31	**1.01394**	**1.10295**	**1.05406**	**1.10498**	**1.04790**	**1.13634**	**1.17367**
32	**1.01135**	**1.04434**	**1.27150**	**1.13668**	**1.04475**	**1.10587**	**1.16224**
33	**1.09274**	**1.05850**	**1.08869**	**1.15609**	**1.03958**	**1.16397**	**1.16253**
34	**1.00285**	**1.08167**	**1.21316**	**1.19959**	**1.03469**	**1.20540**	**1.18833**
35	**1.17163**	**1.13272**	**1.19967**	**1.22353**	**1.03293**	**1.20105**	**1.23764**
36	**1.12980**	**1.14552**	**1.26047**	**1.24145**	**1.03474**	**1.26535**	**1.25125**
37	**1.26657**	**1.17903**	**1.36487**	**1.30179**	**1.03940**	**1.35244**	**1.27450**
38	**1.15674**	**1.19541**	**1.22411**	**1.34261**	**1.04489**	**1.23201**	**1.30450**
39	**1.17042**	**1.20649**	**1.36521**	**1.33301**	**1.05141**	**1.32164**	**1.32267**
40	**1.25353**	**1.21222**	**1.51213**	**1.35536**	**1.06127**	**1.35104**	**1.35608**
41	**1.18520**	**1.22743**	**1.23270**	**1.34508**	**1.07319**	**1.35622**	**1.40102**
42	**1.29522**	**1.23749**	**1.35747**	**1.38881**	**1.08475**	**1.51951**	**1.44634**
43	**1.23277**	**1.23569**	**1.31046**	**1.40686**	**1.09736**	**1.45668**	**1.49583**
44	**1.22073**	**1.23388**	**1.55012**	**1.42906**	**1.11051**	**1.54825**	**1.54532**

As could be expected, the mean property price series $P_{MEAN}{}^{t}$ is rather volatile and so in order to capture the trends in Tokyo commercial property prices, it is necessary to smooth this series. We used the same linear smoothing procedure that was explained in Section 7 above to construct the smoothed land price series $PL_L{}^{t}$. Thus we set the initial value of the smoothed mean series, $P_{MEANS}{}^{1}$, equal to the corresponding unsmoothed value $P_{MEAN}{}^{1}$. We set the quarter 2 value of the smooth equal to the average of the first and third observations in the raw series (so that $P_{MEANS}{}^{2} \equiv (1/2)[\, P_{MEAN}{}^{1} + P_{MEAN}{}^{3}]$. For the Quarter 3 value of the smooth, we used the simple 5 term centered moving average so that $P_{MEANS}{}^{3} \equiv (1/5)[P_{MEAN}{}^{1} + P_{MEAN}{}^{2} + P_{MEAN}{}^{3} + P_{MEAN}{}^{4} + P_{MEAN}{}^{5}]$ and we carried on using this 5 term centered moving average until Quarters 43 and 44 where we used Rows 4 and 5 of the matrix $X(X^TX)^{-1}X^T$ defined in Section 7 for our Henderson linear regression smoother. The resulting smoothed mean price series, P_{MEANS}^t, is listed in Table 10 and plotted in Chart 4 below. We note that the average value of the unsmoothed series P_{MEAN}^t is 1.1644 while the average value of the corresponding smoothed series P_{MEANS}^t is 1.1614.

Table 9 in the previous section lists the land price index PL_{MLIT}^t based on the builder's model using the MLIT transactions data. Table 9 also lists the quarter t structure price indexes, P_{St}. We can use the predicted values from the Model 11 regression explained in Section 5 above in order to construct the imputed value of land sold during quarter t. This quarter t value of land is defined as follows:

$$(36)\ V_L^t \equiv \alpha_t\Sigma_{n=1}^{N(t)} (\Sigma_{j=1}^{4}\omega_j D_{W,tn,j})(\Sigma_{j=1}^{5}\chi_m D_{EL,tnm})(1+\mu(H_{tn}-3))(1+\eta(DS_{tn}-0))\times (1+\theta(TT_{tn}-1))f_L(L_{tn})\ ; \qquad t = 1,...,44.$$

In a similar fashion, we can use the predicted values from the Model 11 regression in order to define the impute value of structures sold during quarter t, V_S^t, as follows:

$$(37)\ V_S^t \equiv p_{St}\Sigma_{n=1}^{N(t)} g_A(A_{tn})(\Sigma_{h=3}^{10} \phi_h D_{H,tnh})S_{tn} \qquad t = 1,...,44.$$

The *quality adjusted quarter t quantities of land and of structures*, Q_L^t and Q_S^t, are defined as follows:

$$(38)\ Q_L^t \equiv V_L^t/PL_{MLIT}^t\ ;\ Q_S^t \equiv V_S^t/P_{St}\ ; \qquad t = 1,...,44.$$

With the prices and quantities of land and structures defined for each quarter, we calculated Fisher (1922) property price indexes, which are listed as P_{FMLIT}^t in Table 10 above and plotted on Chart 4 below.[40]

From viewing Chart 4, it can be seen that the Fisher property price indexes using MLIT data, P_{FMLIT}^t, are quite volatile (due of course to the volatility of the MLIT land price component indexes, P_{LMLIT}^t). The Henderson linear regression smooth of the unsmoothed land price series P_{LMLIT}^t was listed as PL_L^t in Table 9. We use this smoothed land price series along with the new land quantities defined as $Q_L^t \equiv V_L^t/PL_L^t$ in order to define the smoothed Fisher property price index, P_{FMLITS}^t, which is listed in Table 10 above and plotted on Chart 4 below. This series is our preferred measure of overall commercial property prices for Tokyo.

Recall Model 3 in Section 8 above that used REIT data to implement a version of the builder's model. We can use the predicted values from the Model 3 regression in order to

[40] The Laspeyres and Paasche indexes for quarter t are defined as $P_L^t \equiv [P_{LMLIT}^t Q_L^1 + P_{St}Q_S^1]/[P_{LMLIT}^1 Q_L^1 + P_{S1}Q_S^1]$ and $P_P^t \equiv [P_{LMLIT}^t Q_L^t + P_{St}Q_S^t]/[P_{LMLIT}^1 Q_L^t + P_{S1}Q_S^t]$ respectively. The quarter t Fisher index is defined as $P_{FMLIT}^t \equiv [P_L^t P_P^t]^{1/2}$ for t = 1,...,44. See Fisher (1922) for additional materials on these indexes. The Fisher index has strong economic and axiomatic justifications; see Diewert (1976) (1992). We also calculated chained Fisher property price indexes using the same data but these indexes were virtually the same as the Fisher fixed base indexes listed in Table 10.

construct the imputed value of land sold during quarter t. This quarter t value of land is defined as follows:

(39) $V_L{}^t \equiv \sum_{n=1}{}^{41} \alpha_t\omega_n L_{tn}$; $\qquad$ t = 1,...,44.

In a similar fashion, we can use the predicted values from the Model 3 REIT regression in order to define the impute value of structures sold during quarter t, $V_S{}^t$, as follows:

(40) $V_S{}^t \equiv \sum_{n=1}{}^{41} p_{St}(1 - \delta)^{A(t,n)}S_{tn}$ $\qquad$ t = 1,...,44.

The (REIT data based) quality adjusted land price for quarter t is the α_t which appears in (39) and is listed as $PL_{REIT}{}^t$ in Table 9 above. The price of structures is $P_S{}^t = p_{St}$ where p_{St} is the official construction price index. The corresponding period t quantities of land and structure are defined as follows:

(41) $Q_L{}^t \equiv V_L{}^t/PL_{REIT}{}^t$; $Q_S{}^t \equiv V_S{}^t/P_S{}^t$; $\qquad$ t = 1,...,44.

The overall REIT based property price index for quarter t is defined as the Fisher index $P_{FREIT}{}^t$ using the above prices and quantities for land and structures as the basic building blocks. The REIT based overall property price series $P_{FREIT}{}^t$ is listed in Table 10 above and plotted in Chart 4. It can be seen that this series is not volatile and does not require any smoothing.

Our final property price index will be generated by a traditional log price time dummy hedonic regression using the MLIT data.

We use the same notation and definitions of variables as was used in Section 4 above. Define the natural logarithms of V_{tn}, L_{tn} and S_{tn} as LV_{tn}, LL_{tn} and LS_{tn} for t = 1,...,44 and n = 1,...,N(t). The log price time dummy hedonic regression model is the following linear regression model:

(42) $LV_{tn} = \beta_t + \sum_{j=2}{}^4\omega_j D_{W,tnj} + \gamma A_{tn} + \lambda LL_{tn} + \mu LS_{tn} + \Sigma_{h=4}{}^{10}\, \phi_h\, D_{H,tnh} + \eta DS_{tn}$
$+ \theta TT_{tn} + \varepsilon_{tn}$; $\qquad$ t = 1,...,44; n = 1,...,N(t).

The 4 combined ward dummy variables $D_{W,tnj}$ were defined by equations (4) and the discussion around Model 2 in Section 3. The building height dummy variables, $D_{H,tnh}$, were defined just above equations (19) in Section 3. However, due to the small number of observations in the heights equal to 10-14 stories, all buildings in this range were aggregated into the height 10 stories category. As usual, A_{tn} is the age of building n sold in quarter t and DS_{tn} and TT_{tn} are the two subway variables pertaining to building n in quarter t. The 44 time dummy variable coefficients are $\beta_1,...,\beta_{44}$. Note that the dummy variable for the first combined ward, $D_{W,tn1}$, is not included in the linear regression defined by (42) in order to prevent multicollinearity. Similarly, the dummy variable for building height equal to 3 was also excluded from the regression to prevent multicollinearity. There are 59 unknown parameters in the regression. The R^2 for this regression was 0.7593. This is higher than our Model 9 and Model 11 R^2 using the same data, which were 0.7091 and 0.7143 respectively. The estimated coefficients and their t statistics are listed in Table 11.

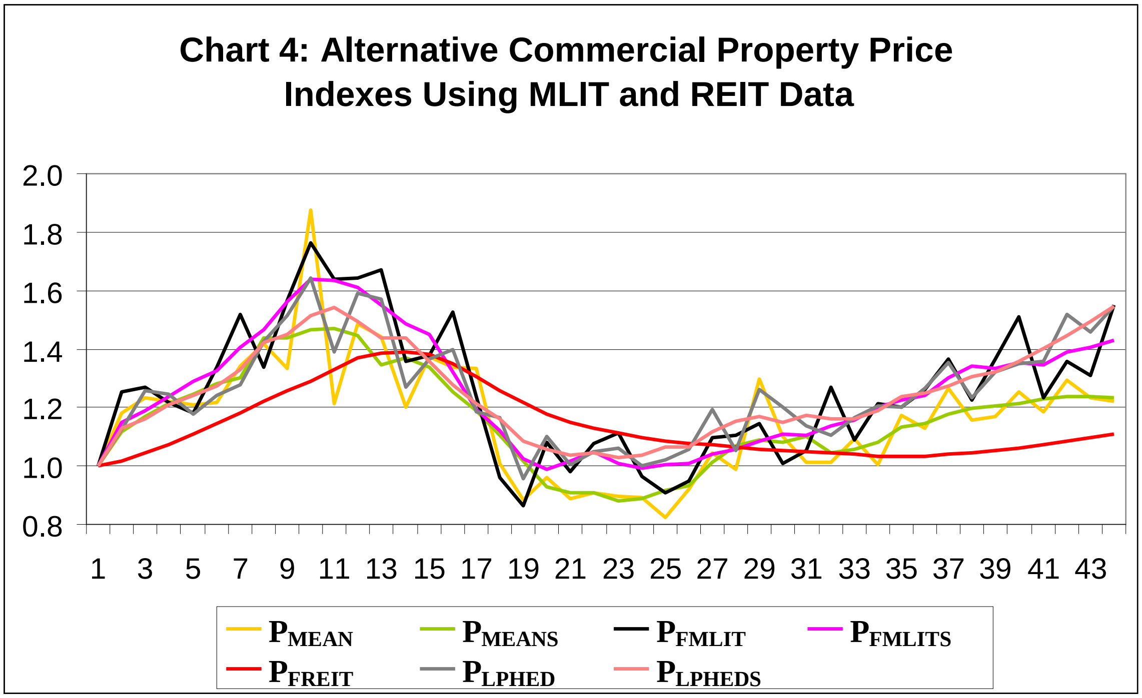
Table 11: Estimated Coefficients for the Log Price Time Dummy Hedonic Regression Model

Coef	Estimate	t Stat	Coef	Estimate	t Stat	Coef	Estimate	t Stat
β_1	-0.20529	-1.542	β_{21}	-0.20015	-1.393	β_{41}	0.09941	0.718
β_2	-0.08681	-0.607	β_{22}	-0.15772	-1.076	β_{42}	0.21310	1.472
β_3	0.02537	0.194	β_{23}	-0.14495	-1.024	β_{43}	0.17087	1.205
β_4	0.01308	0.100	β_{24}	-0.20373	-1.462	β_{44}	0.23184	1.554
β_5	-0.04096	-0.299	β_{25}	-0.18593	-1.248	γ	-0.00970	-10.490
β_6	0.01025	0.075	β_{26}	-0.15031	-0.981	μ	0.49390	12.890
β_7	0.04072	0.301	β_{27}	-0.02962	-0.203	λ	0.52956	14.200
β_8	0.14967	1.077	β_{28}	-0.15537	-1.098	ω_2	-0.32900	-13.990
β_9	0.21020	1.524	β_{29}	0.02786	0.195	ω_3	-0.49677	-16.620
β_{10}	0.29059	2.097	β_{30}	-0.02329	-0.160	ω_4	-0.73092	-21.530
β_{11}	0.12291	0.844	β_{31}	-0.07748	-0.531	ϕ_4	0.02949	0.868
β_{12}	0.25957	1.811	β_{32}	-0.10465	-0.698	ϕ_5	0.12691	3.384
β_{13}	0.24673	1.738	β_{33}	-0.05345	-0.378	ϕ_6	0.12419	2.766
β_{14}	0.03231	0.227	β_{34}	-0.01848	-0.132	ϕ_7	0.18867	3.748
β_{15}	0.10645	0.751	β_{35}	-0.02209	-0.160	ϕ_8	0.31515	5.874
β_{16}	0.12910	0.861	β_{36}	0.03006	0.214	ϕ_9	0.45593	7.579
β_{17}	-0.03416	-0.239	β_{37}	0.09662	0.712	ϕ_{10}	0.55094	8.261
β_{18}	-0.05266	-0.323	β_{38}	0.00336	0.023	η	-0.00018	-4.234
β_{19}	-0.24786	-1.715	β_{39}	0.07358	0.522	θ	0.00067	0.458
β_{20}	-0.10737	-0.737	β_{40}	0.09558	0.680			

The standard errors for the time coefficients β_t were fairly large (in the 0.13 to 0.15 range). Define the unnormalized land price for quarter t, α_t, as the exponential of β_t; i.e., $\alpha_t \equiv \exp(\beta_t)$ for t = 1,...,44. The log price hedonic regression property price for quarter t, $P_{LPHED}{}^t$ is defined as α_t/α_1 for t = 1,...,44. This traditional hedonic regression model property price index $P_{LPHED}{}^t$ is listed in Table 10 above and graphed in Chart 4 below.

It is interesting to note that our estimated λ and μ parameters almost sum to unity. Thus a generic commercial property sold in quarter t at price P with land and structure areas L and S respectively has a price that is approximately proportional to the Cobb-Douglas function $\alpha_t L^{\lambda} S^{\mu}$ which has returns to scale that are approximately equal to one. Note also that the estimated ω_k follow the same pattern that we saw in Sections 3-5 for land prices; i.e., the composite Ward 1 is the most expensive ward, Ward 2 the next most expensive, Ward 3 less expensive again and Ward 4 has the lowest level of property prices. The height dummy variables exhibit the same trends that we saw in our MLIT builder's models: the higher the height of the structure, the higher is the price of the property. Finally, the distance from the nearest subway station parameter η is significantly negative indicating that property value falls as the distance increases. The subway travel time parameter θ has an unexpected positive sign but is not significantly different from 0. Finally, it is possible to convert the estimated age coefficient γ into an estimate for a geometric rate of structure depreciation, δ. The formula for this conversion is $\delta \equiv 1 - e^{\gamma/\beta}$.[41] When this conversion formula is utilized, we find that our estimated δ is 0.01945; i.e., the traditional hedonic regression model generates an implied annual geometric depreciation rate equal to 1.945% per year, which is a reasonable estimate.

Viewing Table 10 or Chart 4, it can be seen that the time dummy hedonic regression model implied property price index P_{LPHED}^t is just as volatile as the corresponding builder's model property price index P_{FMLIT}^t. Thus we apply our modified Henderson linear smoothing operator to P_{LPHED}^t which produces the smoothed series, P_{LPHEDS}^t, which is also listed in Table 10 and plotted in Chart 4 below.



The two top jagged lines are the Fisher property price index using the builder's model, P_{FMLIT}^t, and the log price time dummy hedonic regression property price index, P_{LPHED}^t. Both of these series use the MLIT sales transaction data. Their linear smooths are P_{FMLITS}^t and P_{LPHEDS}^t. It can be seen that these two smoothed series approximate each other reasonably well.[42] What is somewhat surprising is that the smoothed mean index P_{MEANS}^t

---

[41] See McMillen (2003; 289-290), Shimizu, Nishimura and Watanabe (2010; 795) and Diewert, Huang and Burnett-Isaacs (2107; 24) for derivations of this formula.

[42] Diewert (2010) noticed that the Fisher property price index generated by the builder's model frequently approximated the traditional log price time dummy property price index using the same data. However, the

(which uses the same transactions data) approximates the two smoothed hedonic indexes reasonably well but the series gradually diverge due to the fact that an index based on average prices per m^2 cannot take depreciation into account.[43] The hills and valleys in the $P_{MEANS}{}^t$ series are less pronounced than the corresponding fluctuations in the $P_{FMLITS}{}^t$ and $P_{LPHEDS}{}^t$ series but the turning points are the same. Finally, it can be seen that the Fisher property price series that is based on appraised values of properties, $P_{FREIT}{}^t$, does not provide a good approximation to the two smoothed series based on transactions, the $P_{FMLITS}{}^t$ and $P_{LPHEDS}{}^t$ series. The fluctuations in $P_{FREIT}{}^t$ are too small and the turning points in this series lag well behind our preferred series.

12. Conclusion

Here are our main conclusions:

- It is possible to construct a quarterly transactions based commercial property price index that can be decomposed into land and structure components.
- The main characteristics of the properties that are required in order to implement our approach are: (i) the property location (or neighbourhood); (ii) the floor space area of the structure on the property; (iii) the area of the land plot; (iv) the age of the structure and (v) the height of the building. We also require an appropriate exogenous commercial property construction price index.
- The land price index that our hedonic regression model generates may be too volatile and hence may need to be smoothed. We found that a slightly modified five quarter moving average of the raw land price indexes did an adequate job of smoothing. This means that the final land price index could be produced with a two quarter lag.
- We found that a smoothed version of a traditional log price time dummy hedonic regression model produced an acceptable approximation to our preferred smoothed builder's model overall price index.
- We also found that a very simple overall price index which is proportional to the quarterly arithmetic average of each property price divided by the corresponding structure area provided a rough approximation to our preferred price index. This model cannot take depreciation into account and hence will in general have an downward bias but it has the advantage of requiring information on only a single property characteristic (the structure floor space area) in order to be implemented.
- The price indexes that were based on appraisal and assessed value information were not satisfactory approximations to the transactions based indexes. The turning points in these series lagged our preferred series and the appraisal based series smoothed the data based series to an unacceptable degree.[44]

---

key to a successful approximation is that the time dummy model must generate a reasonable implied structure depreciation rate, which is the case for our particular data set.

[43] If the age structure of the quarterly sales of properties remains reasonably constant, then this neglect of depreciation will not be a factor.

[44] These points are well known in the real estate literature; see Chapter 25 in Geltner, Miller, Clayton and Eichholtz (2014).

- The two versions of the builder's model that estimated multiple (net) depreciation rates produced virtually the same indexes and virtually identical depreciation schedules. These rates of depreciation changed materially as the structure aged and the depreciation rates became appreciation rates for structures over age 40.

Our overall conclusion is that it should be possible for national income accountants to construct acceptable commercial land price series using transactions data on the sales of commercial properties. The required information on the characteristics of the properties is being collected by some private sector businesses. It should be possible for government statisticians to collect the same information using building permit, land registry and property assessment data.

References

Bokhari,S and D. Geltner (2012), "Estimating Real Estate Price Movements for High Frequency Tradable Indexes in a Scarce Data Environment," *Journal of Real Estate Finance and Economics* 45(2), 522-543.

Bostic, R.W., S.D. Longhofer and C.L. Redfearn (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.

Crosby, N, C. Lizieri and P. McAllister (2010), "Means, Motive and Opportunity? Disentangling Client Influence on Performance Measurement Appraisals", *Journal of Property Research* 27(2), 181–201.

de Haan, Jan and W.E. Diewert (eds.) (2011), *Residential Property Price Handbook*, Luxembourg: Eurostat.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1992), "Fisher Ideal Output, Input and Productivity Indexes Revisited", *Journal of Productivity Analysis* 3, 211-248.

Diewert, W.E. (2008), "The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Further Directions", pp. 11-36 in *Proceedings From the OECD Workshop 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., 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.

Diewert, W. E., K. Fox and C. Shimizu (2016), "Commercial Property Price Indexes and the System of National Accounts," *Journal of Economic Surveys* 30(5), 913-943.

Diewert, W.E., N. Huang and K. Burnett-Isaacs (2017), "Alternative Approaches for Resale Housing Price Indexes", Discussion Paper 17-05, Vancouver School of Economics, The University of British Columbia, Vancouver, Canada, V6T 1L4.

Diewert, W. E. and C. Shimizu (2015a), "A Conceptual Framework for Commercial Property Price Indexes," *Journal of Statistical Science and Application* 3(9-10),131-152.

Diewert, W. E. and C. Shimizu (2015b), "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.

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.

Geltner, D. (1997), "The Use of Appraisals in Portfolio Valuation and Index", *Journal of Real Estate Finance and Economics* 15, 423-445.

Geltner, D. and S. Bokhari (2017), "Commercial Buildings Capital Consumption and the United States National Accounts", *Review of Income and Wealth* (forthcoming).

Geltner, D. and W. Goetzmann (2000), "Two Decades of Commercial Property Returns: A Repeated-Measures Regression-Based Version of the NCREIF Index", *Journal of Real Estate Finance and Economics* 21, 5-21.

Geltner, D., R.A. Graff and M.S. Young (1994), "Random Disaggregate Appraisal Error in Commercial Property, Evidence from the Russell-NCREIF Database", *Journal of Real Estate Research* 9(4), 403-419.

Geltner, D.M., N.G. Miller, J. Clayton and P. Eichholtz (2014), *Commercial Real Estate Analysis and Investments*, Third Edition, Mason Ohio: OnCourse Learning.

Geltner, D., H. Pollakowski, H. Horrigan, and B. Case (2010), "REIT-Based Pure Property Return Indexes", United States Patent Application Publication, Publication Number: US 2010/0174663 A1, Publication Date: July 8, 2010.

Henderson, R. (1916), "Note on Graduation by Adjusted Average", *Actuarial Society of America Transactions* 17, 43-48.

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.

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.

Muth, R.F. (1971), "The Derived Demand for Urban Residential Land", *Urban Studies* 8, 243-254.

Nishimura, K.G and C. Shimizu (2003), "Distortion in Land Price Information: Mechanism in Sales Comparables and Appraisal Value Relation", Discussion paper 195, Center for International Research on the Japanese Economy, University of Tokyo.

Patrick, G. (2017), "Redeveloping Ireland's Residential Property Price Index (RPPI)", paper presented at the Ottawa Group Meeting at Eltville, Germany, May10.

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.

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

Shimizu, C (2016), "Microstructure of Asset Prices, Property Income, and Discount Rates in Tokyo Residential Market", *International Journal of Housing Markets and Analysis* (forthcoming), IRES-NUS(National University of Singapore) Working Paper 2016-002.

Shimizu, C., W.E. Diewert, K.G. Nishimura and T.Watanabe (2015), "Estimating Quality Adjusted Commercial Property Price Indexes Using Japanese REIT Data", *Journal of Property Research* 32(3), 217-239.

Shimizu, C. and K.G. Nishimura (2006), "Biases in Appraisal Land Price Information: the Case of Japan", *Journal of Property Investment & Finance*, 24(2), 150- 175.

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.

Shimizu, C., K.G. Nishimura and T.Watanabe (2012), "Biases in Commercial Appraisal-Based Property Price Indexes in Tokyo Lessons from Japanese Experience in Bubble Period", RIPESS (Reitaku Institute of Political Economics and Social Studies), No.48, (presented at the International Conference on Commercial Property Price Indicators on 10-11, May 2012, the European Central Bank in Frankfurt).

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