Estimating and smoothing appraisal-based commercial real-estate performance indexes

Paulo Picchetti

paulo.picchetti@fgv.br

Instituto Brasileiro de Economia/Escola de Economia de São Paulo
Fundação Getulio Vargas - Brazil

March 5, 2013

Abstract

Differences in frequencies of property evaluations and index construction impose biases in the formulas usually employed to calculate commercial rea-estate investment performance. Here we propose a statistical methodology to estimate values for properties at the time where such an index is computed, in the cases of properties whose last re-evaluations occurred in the past. The statistical model is a hierarchical regression, which captures the main hypotheses used by standard practices of external valuation for commercial real-estate. The statistical results are employed in the construction of an index for Brazilian commercial real-estate performance.
JEL: C23, C43, C81, R33
Keywords: Longitudinal Data, Index Numbers, Real Estate Markets, Methodology for Collecting and Estimating Microeconomic Data.

1 Introduction

Indexes for measuring the performance of commercial real-estate portfolios have been regularly produced in several countries following a standard formula, such as the one adopted by NECREIF in the US and by IPD in the UK. The essential information contained in this formulation is the evolution of property values across time. Ideally, these should reflect value resulting from market transactions. In practice, as these transactions tend to be very infrequent, property values are typically evaluated by regularly conducted appraisals, based on a set of approved methodologies.

A further complication arises from the fact that these appraisals are usually not conducted at the same frequency in which the indexes are calculated. Whereas for constructing past values of a series, some sort of interpolation of the data between appraisal dates can be used, the problem remains when the latest available information refers to an appraisal conducted at some period in the past.

The alternatives in this scenario involve different compromises. One could convince market participants who provide information to increase the frequency of their appraisals, which involves costs and time. Alternatively, the performance index could only be calculated at the frequency of the appraisals, which in turn would eliminate the possibility of continuous timely information.

A third possibility, chosen by the index for commercial real-estate in Brazil (IGMI-C), is to employ econometric methods to estimate property values prior to eventual re-appraisals, using information from the rest of the market and also from regularly updated variables from individual properties, such as the evolution of rents. In this paper, we present and discuss the different biases introduced in the index by infrequent appraisals, and compare them to the results of the index calculated from the econometrically estimated property values.

The econometric method which allows us to estimate property values is based on a hierarchical model, and eventually adjusts for total variation when appraisals for individual properties are eventually carried out, therefore eliminating biases on the long-run. Specifically, we employ the conditioning modeling framework to represent longitudinal (time-ordered) data, considering individual properties as the level-2 units of analysis and observations at different time points as the level-1 units. Our goal is the assessment of change in property prices, in the context of latent growth-curve models. Characteristics of the individual properties constitute yet another level to the model, in terms of location and type of usage. At each point in time where a particular property was neither transacted nor re-evaluated, its value is then estimated by associating its characteristics parametrically to the other similar (in conditional terms) properties that were re-evaluated at the period, and to the evolution of its own observed performance in terms of rents.

2 Formulas

The standard formulas for commercial real-estate investment performance, such as the ones adopted by NECREIF in the U.S., and by IPD in the U.K. and more than 30 other countries is:

$$\text{Income Return} = \frac{NOI_t}{V_{t-1} + I_t}$$

$$\text{Capital Return} = \frac{(V_t - V_{t-1}) - I_t + A_t}{V_{t-1} + I_t}$$

$$\text{Total Return} = \frac{(V_t - V_{t-1}) - I_t + A_t + NOI_t}{V_{t-1} + I_t}$$

where:

NOI Net Operational Income - Income received by the investor, mainly in the form of rents, deducing operational costs such as taxes and administration fees.

V Value - Price of the property, typically provided by an external evaluator following standard practices.

A Sales - Partial or total sales of the property, evaluated by market transaction values.

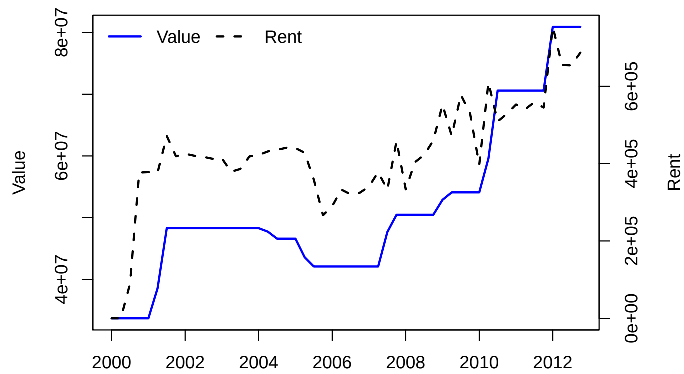
I Investment - Expenses paid by the investor which potentially affect the value of the property.

2.1 Biases related to the frequency of re-evaluations

As can be readily seen in the above formulas, the effects of the lagged information on values caused by the infrequent re-evaluations is to under-estimate the capital return, and on the other hand to over-estimate the income return, supposing an increase in property prices not yet accounted for. The net result for the total return depends on the relative intensity of each of the biaeses, and is therefore uncertain. These biases are naturally undesirable, and therefore introduced the compromises mencioned in the introduction when such an index is calculated with this information shortcoming.

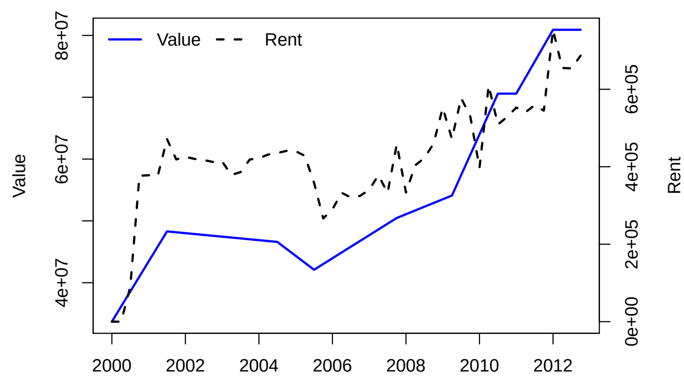
When dealing with this fact in the historical series, one possibility is to smooth the property values by linear interpolating between periods of appraisal. The graph below displays the time evolution of rents and values for a typical property in the IGMI-C database.

Before in−sample Value Interpolation



The step-function nature of the value accross time becomes evident, which generates the above mentioned biases in the calculation of the index. Linear interpolation of the values between appraisal dates is a possibility for atenuating these biases:

After in−sample Value Interpolation



However, linear interpolation is not an option at the end of the sample. Therefore, using actual reported values will introduce biases in the latest figures provided by the straightforward computation of the index using the provided information.

The next section presents a model which essentially tries to anticipate the information contained in new appraisals for the properties that have not yet been re-evaluated in the most recent past. Standard appraisal methodologies seek to input a price to a commercial real-estate property. On one hand, this is done by analyzing the evolution of its income generation. On the other hand, additional information is produced by a comparison with other properties whose values are considered to be known (either by more recent appraisals or even better, by actual market transactions), and which are considered to be similar (in terms of location and type of usage). The next section presents a statistical model designed to capture this type of procedure, using the longitudinal information contained in the index database.

3 Statistical Model

In statistics, a hierarchical model is a type of random-coefficient regression model, where both intercepts and slopes are allowed to vary between levels, and conditionally on values of covariates. These levels refer to types of aggregation defining meaningful clusters of individuls under analysis. In our case, individuals will be considered as commercial real-estate units. A first logical type of aggregation for real-estate is geographical location, which in our case will be the particular State of Brazil where the property is located. A second form of aggregation is type of usage of the commercial property, which in our case will stratify information in terms of commercial offices, shopping centers, logistics, hotels, and others.

Latent growth-curve models are a special case of random-coefficient models where it is the coefficient of time that varies randomly between subjects. It models the shape of trajectories of individuals over time, and how these trajectories vary, both systematically, due to time-level and subject-level covariates, and randomly. We will interpret level-1 as time-periodos (quarters) over which individual properties are evaluated in terms of value and rents. Level-2 will be considered as clusters defined over usage types and level-3 will be associated with location.

Therefore, the value of each property at each point in time will be considered as an observation drawn from a distribution where values of other properties are related in terms of usage and location. When one property is re-evaluated, this will be considered to be its value in the end of the sample, whereas for properties that were evaluated at past periods, their values will be predicted using the information provided by the model's estimated parameters, each individual property receiving a value reflecting the evolution of its own rent values, as well as the evolution of values observed in similar properties, defined by its position on the distributions conditional on usage and location.

The proposed model is:

$$y_{ijt} = \beta_0 + \sum_{k=1}^{3} \beta_k t + \sum_{l=1}^{18} \beta_l state_i + \theta_i + \eta_{1ij} + \eta_{2ij} x_{ijt} + \epsilon_{ijt}$$

$$\begin{aligned}
E[\theta] &= 0 & Var[\theta] &= \sigma^2_\theta \\
E[\eta_{1ij}|\theta_i] &= 0 & Var[\eta_{1ij}] &= \sigma^2_{\eta_1} \\
E[\eta_{2ij}|\theta_i] &= 0 & Var[\eta_{2ij}] &= \sigma^2_{\eta_2} \\
E[\epsilon_{ijt}|\theta_i, \eta_{1ij}, \eta_{2ij}] &= 0 & Var[\epsilon_{ijt}] &= \sigma^2_\epsilon
\end{aligned}$$

where the subscripts represent the different levels for each observation, i for property, j for type of usage, and t for time (measured in quarters). *state* are dummy variables relating each property to its location in one of the 18 Brazilian states in the database. y_{ijt} is the rate of quarterly change in value for a particular property i, with usage type j at time t. x_{ijt} represents the rate of change in rents received again by a particular property i, with usage type j at time t. These rates of change are smoothed by a two-quarters moving average to filter common oscillation in rent receipts that arise not from economic fundamentals, but rather by random payment delays with subsequent compensations. θ are individual-level residuals, η_1 and η_2 are type of usage level residuals, and finally ϵ are overall residuals.

The estimated coefficients for the sample composed by the IGMI-C database are shown in the table below (state dummies coefficients omitted).

Table 1: Model Results

	Number of	Observations per Group		
Group Variable	Groups	Minimum	Average	Maximum
Type of Usage	6	60	959.2	2648
Property	152	2	33.9	51

Number of Observations = 5155

Value	Coef.	Std. Err.	z	P-value
t	-.254	.028	-9.10	0.000
t^2	.001	.0002	8.98	0.000
t^3	-2.38e-06	2.69e-07	-8.85	0.000
$const.$	15.806	1.72	9.19	0.000
LR test vs. linear regression: $\chi^2(3) = 996.38$ Prob $> \chi^2 = 0.0000$				

Random-effects Parameters	Estimate	Std. Err.	95% Conf.	Interval
Property				
σ_{η_1}	.064	.005	.055	.074
σ_{η_2}	.023	.002	.019	.027
Type of Usage				
σ_θ	.003	.012	.014	.064
σ_ϵ	.043	.0004	.042	.044

3.1 Assigning values to the random intercepts

Once we estimate the model's parameters and consider them as if they are the true values for these parameters, we can estimate the values for the individual random components recursively accross levels. For clarity, if we consider only two levels, represented by the individual properties and time, we can substitute the estimated βs to obtain

$$y_{it} = \hat{\beta} + \underbrace{\eta_t + \epsilon_{it}}_{\xi_{it}}$$

and then estimate total residuals $\hat{\xi}_i t$ as

$$\hat{\xi}_{it} = y_{it} - \hat{\beta} = \eta_t + \epsilon_{it}$$

The maximum-likelihood estimate of η_t is the sample mean of this estimated total residual over the n_t periods of each individual property

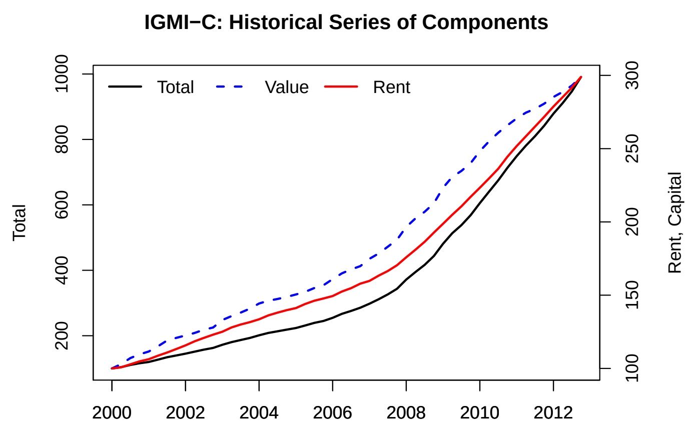
$$\hat{\eta}_j = \frac{1}{n_t} \sum_{i=1}^{n_t} \hat{\xi}_{it}$$

Proceding recursively for all levels provides us with estimates for each of the random components for every individual property. This allow us finally to estimate the value of each property that has not been re-evaluated in the end of the sample.

3.2 Estimation of the smoothed index

At each point in time, the index can be computed using the standard formulas, with values for each individual property given by their actual appraisal values (if the appraisal is conducted during this period), or by the value estimated by the hierarchical regression model. When properties with estimated values are eventually re-evaluated at a later period, an adjustment is made to compensate for potential errors in the estimation process, so that values accumulated accross time are note distorted.

The calculated index along with its components can be seen on the graph below, and a table with the actual figures can be found on the Statistical Appendix.



4 Conclusion and further developments

The methodology presented here was developed to circumvent the limitations imposed by the available information for a construction of a commercial real-estate performance index in Brazil. These limitations have two dimensions. First, properties are re-evaluated at a longer frequency compared to the production of a quarterly index. This is not a particularity of Brazilian data, but the solution employed in other countries faces the second kind of limitation of information in the Brazilian case. If a sample large enough is available, so that at each point in time one can consider sub-samples which are both statistically representative and also comprised solely of properties re-evaluated at the period, than one can construct an index based on a pseudo-panel of the properties. However, Brazilian data includes up to this point around 350 observations, which are not evenly distributed accross the relevant stratification variables of location and type of usage. Therefore, the solution of a pseudo-panel is not yet available for the Brazilian database. However, a continuous effort is being made to increase the number of participants in the index database, and also measures for increasing the frequency of new appraisals, including specific legislation, are being seriously considered. In the meantime, the proposed methodology will benefit from the increase in sample size and frequency of appraisals, providing an important missing information for this market.

References

Edelstein, R. and Quan, D. (2006), *How Does Appraisal Smoothing Bias Real Estate Returns Measuerment ?*, The Journal of Real Estate Finance and Economics, Springer.

Fisher, J., D. Geltner, and R. Webb. (1994). *Value Indices of Commercial Real Estate: A Comparison of Index Construction Methods*, Journal of Real Estate Finance and Economics 9, 137-164.

Gelman, Andrew and Hill, Jennifer (2007), *Data Analysis using Regression and Multilevel/Hierarchical Models* Cambridge University Press.

Geltner, D., and W. Goetzmann. (2000). *Two Decades of Commercial Property Returns: A Repeated- Measure Regression-Based Version of the NCREIF Index*, Journal of Real Estate Finance and Economics 21(1), 5-21.

IPD Australia (2009), *Methodology changes to the PCA/IPD Australia Property Index*, IPD Australia.

McAllister, P., A. Baum, N. Crosby, P. Gallimore, and A. Gray. (2003). *Appraiser Behaviour and Appraisal Smoothing: Some Qualitative and Quantitative Evidence*, Journal of Property Research 20(3), 261Y-80.

Rabe-Hesketh, Sophia and Skrondal (2012), Anders, *Multilevel and Longitudinal Modeling using Stata*, Third Edition, Stata Press.

5 Statistical Appendix

IGMI-C is based on information from a group of data suppliers, including institutional investors and institutions related to the real-estate sector, such as pension funds, consultants, managers of real-estate funds, and builders, among others. An initial effort allowed the construction of a historical series, spanning in this release the first quarter of 2000 until the fourth quarter of 2012. The index in its three versions is calculate on a quarterly basis, beginning with a value of 100 in the first quarter of 2000, being updated at each semester according to the expression $\text{IGMI}_t = \text{IGMI}_{t-1} \times (1+\text{Return}_t)$. The database in the last quarter of 2012 comprises 352 individual properties, divided among offices, shopping-centers, commercial activities, hotels, industrial and logistic facilities, among others. The biggest shares are represented by offices (around 50% of the total properties) and shopping-centers (around 25% of the total). On a regional level, the biggest shares are in the State of São Paulo (around 37% of the total properties), and Rio de Janeiro (around 26% of the total).

Table 2 shows the available historical record for the index.

	quarter	total	capital	income
1	2000.00	100.00	100.00	100.00
2	2000.25	103.89	102.94	100.95
3	2000.50	110.58	107.06	103.41
4	2000.75	115.82	109.62	105.84
5	2001.00	119.77	111.53	107.60
6	2001.25	126.84	115.10	110.51
7	2001.50	134.24	119.09	113.12
8	2001.75	139.68	120.87	116.03
9	2002.00	145.18	122.50	119.02
10	2002.25	151.37	124.23	122.41
11	2002.50	157.53	126.44	125.22
12	2002.75	162.71	127.93	127.87
13	2003.00	172.26	133.03	130.27
14	2003.25	180.39	135.71	133.79
15	2003.50	186.86	138.11	136.23
16	2003.75	193.23	140.82	138.20
17	2004.00	201.32	144.32	140.55
18	2004.25	208.50	146.19	143.74
19	2004.50	213.52	147.39	146.02

20	2004.75	218.62	148.83	148.09
21	2005.00	223.55	150.47	149.79
22	2005.25	231.19	152.25	153.15
23	2005.50	239.23	154.86	155.84
24	2005.75	245.11	156.79	157.72
25	2006.00	254.73	160.96	159.73
26	2006.25	266.87	164.89	163.44
27	2006.50	275.79	167.59	166.22
28	2006.75	285.60	169.81	169.93
29	2007.00	297.88	174.82	172.22
30	2007.25	311.55	178.45	176.55
31	2007.50	326.45	183.11	180.39
32	2007.75	343.62	187.88	185.18
33	2008.00	371.69	196.61	191.70
34	2008.25	394.70	202.37	197.94
35	2008.50	416.59	206.91	204.49
36	2008.75	443.58	212.69	212.02
37	2009.00	481.10	223.33	219.35
38	2009.25	513.30	230.70	226.79
39	2009.50	538.21	234.83	233.74
40	2009.75	568.49	240.13	241.61
41	2010.00	605.38	248.30	249.07
42	2010.25	640.38	254.97	256.78
43	2010.50	675.01	260.87	264.72
44	2010.75	713.45	265.93	274.66
45	2011.00	748.68	270.61	283.39
46	2011.25	780.81	274.53	291.45
47	2011.50	809.99	277.18	299.53
48	2011.75	842.12	280.71	307.60
49	2012.00	878.79	285.25	316.02
50	2012.25	910.98	288.71	323.76
51	2012.50	946.61	293.02	331.59
52	2012.75	991.02	298.92	340.46

Table 2: IGMI-C historical record