A Model Based Approach to Produce Residential or Commercial Property Price Indices

16th Ottawa Group Meeting
Rio de Janeiro
Brazil May, 2019

IBGE
Instituto Brasileiro de Geografia e Estatística
83 anos do IBGE

Paulo Fernando Mahaz Simões

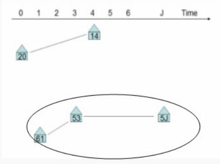
Brazilian Institute of Geography and Statistics – IBGE

paulo.mahaz@ibge.gov.br

HIGHLIGHTS

- **Hedonic Double Imputation Laspeyres House Price Indices**
- **We Link Sold Properties**
- **Create Pseuso Housing Units**
- **Calculate Accurate Indices with Reduced Sample Sizes**

Linking sold properties



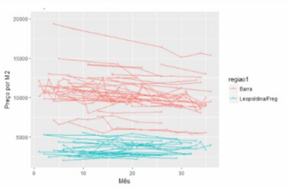
- We create **PSEUDO HOUSING UNITS**

Obtain a longitudinal data Set - Comprised by prices of different properties in distinct months.

Table 1 – Prices of properties (US$)

PSU	jan/14	fev/14	mar/14	abr/14	mai/14	jun/14	jul/14	ago/14	set/14	out/14	nov/14	dez/14
1	300		305	304		308	310			320	316	317
2	420		422	430			436	439		445	447	450
3	500	480	485	500	510			445	442		440	
4	600	610	620		605	608	613		615	618		630
5	730			725		724	728		740		750	743
...	...	...	...	...	...	...	...	...	...	...	...	...
60	800		805			808	805			820		825

Plot of our Longitudinal Data



Questions

- Why some prices increase and other decrease?
- Which factors are related to changes in prices over time?

To answer this questions we apply

Mixed Effects Models

- Interesting technique to analyze longitudinal data because they offer us some perrogatives:

a) Analyze individual trajectories

b) Identify variance components

c) Predictors that explain intraindividual variance and variance among groups

Methodology

- We define a fixed sample (S) - 60 Specific Properties.
- We specify the Model

Ex: To Calculate Results for Jan/2016

- We estimate Model coefficients taking into account data from the last 24 months
- Generate Predict Values for each property in the sample for Dec/15 and Jan/2016
- Calculate the Index for Jan/2016 $\hat{I}^{Jan16} = \frac{\sum \hat{P}_{i \epsilon S}^{Jan16}}{\sum \hat{P}_{i \epsilon S}^{Dez15}}$

Model

Variables

- Size, Month, Condo Characteristics, Neighbor (Zip Code), Distance to the sea;

$$Y_{ij} = \beta_0 + b_{0i} + \beta_1 T_{ij} + b_{1i} T_{ij} + \beta_2 B_i + \beta_3 PC_{ij} + \beta_4 D_{ij} + \epsilon_{ij}$$
$$b_{0i} \sim N(0, \sigma_0^2)$$
$$b_{1i} \sim N(0, \sigma_1^2)$$
$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

Model Estimation

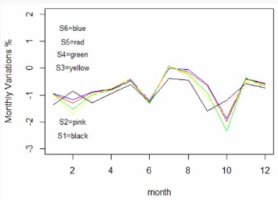
	Variable	Coef	Estimates	sd	p-value
	Intercept	β_0	6378.54	685.37	0.000
	Month	β_1	-40.89	6.68	0.000
neighborhood	Area1 Barra	β_2	–	–	–
	Area2 Bonsuc	β_2	-2818.75	790.40	0.000
	Area3 Freg	β_2	-1459.82	988.11	0.145
	Area4 Olaria	β_2	-2945.64	1098.70	0.009
	Area5 Penha	β_2	-2788.51	903.31	0.003
	Area6 Ramos	β_2	-3579.34	840.71	0.000
	Area7 V. Pen	β_2	-2650.61	902.64	0.005
condo characteristics	Pad Cond 1	β_3	–	–	–
	Pad Cond 2	β_3	461.29	251.66	0.000
	Pad Cond 3	β_3	1160.72	270.71	0.000
	Pad Cond 4	β_3	1917.37	375.06	0.000
	Pad Cond 5	β_3	4394.90	667.27	0.000
Sea distance	Dist Mar longe	β_4	–	–	–
	Dist Mar próxi	β_4	3160.05	635.35	0.000
	Dist Mar orla	β_4	5141.09	755.09	0.000
	Quality of fit	AIC = 3458.81	BIC = 3518.97	LL = -1711.40	

Hedonic Double Imputation Laspeyres Price Indices

6 different Sample Sizes

Mês	S1	S2	S3	S4	S5	S6
Jan	-1.3616	-1.2257	-0.9809	-0.9678	-0.9665	-0.9399
Fev	-0.8593	-0.8289	-1.7305	-1.5259	-1.2960	-1.1849
Mar	-1.2923	-1.2788	-1.0337	-0.9837	-0.9106	-0.8921
Abr	-0.9461	-0.9675	-0.7374	-0.8255	-0.7968	-0.7808
Mai	-0.6097	-0.3694	-0.3694	-0.4981	-0.4772	-0.4240
Jun	-1.2654	-1.2507	-1.2477	-1.3173	-1.2772	-1.2045
Jul	-0.3982	0.1136	0.0307	0.0897	0.0275	-0.0115
Ago	-0.4451	-0.3675	-0.2513	-0.2249	-0.1297	-0.0572
Set	-1.5931	-1.6205	-1.1224	-0.9200	-0.6637	-0.6117
Out	-1.1872	-1.2907	-1.7590	-2.3445	-1.9898	-1.8832
Nov	-0.5819	-0.6782	-0.4999	-0.3636	-0.4039	-0.3805
Dez	-0.7385	-0.6300	-0.6215	-0.6777	-0.5903	-0.5849
Acum:	-10.7215	-9.9261	-9.8645	-10.0846	-9.0894	-8.6112

Indices for 2016 from different samples



Bootstrap Confidence Intervals for Sample S6

Month	Variation (%)	Lower	Upper
Jan/16	-0.93997576	-1.0850158	-0.79493576
Fev/16	-1.18498173	-1.8455017	-0.52446173
Mar/16	-0.89217714	-1.0489771	-0.73537714
Abr/16	-0.78083960	-0.9905596	-0.57111960
Mai/16	-0.42403444	-0.6788344	-0.16923444
Jun/16	-1.20455500	-1.5397150	-0.86939500
Jul/16	-0.01156223	-0.3173222	0.29419777
Ago/16	-0.05720701	-0.3806070	0.26619299
Set/16	-0.61173888	-1.1605389	-0.06293888
Out/16	-1.88321976	-2.6201798	-1.14625976
Nov/16	-0.38051861	-0.6196386	-0.14139861
Dez/16	-0.58491403	-0.7652340	-0.40459403

Concluding Remarks

- **We calculate Quality Adjusted Indices**
- **Longitudinal models allows more accurate results than other methods with the same sample size**
- **Results based on transaction prices**
- **We are analysing alternative data sources and methods (Using Appraisal Prices)**

Appendix

Mixed Effects Models

Extension of a standard linear model:
$Y_{ij} = \beta_0 + \beta_1 X_{ij} + b_{0i} + b_{1i} X_{ij} + \epsilon_{ij}$
Matrix Notation: $Y_i = X_i \beta + Z_i b_i + \varepsilon_i$
We can specify distributions: $b_i \sim NMV(0, \Sigma_i)$ $\varepsilon_i \sim NMV(0, R_i)$
$R_i = \sigma_i^2 I_{n_i}$
Cov Structure given by: $Z_i \Sigma_i Z_i' + R_i$.
Intraindividual Variance: (R_i)
Between Variance: $(Z_i \Sigma_i Z_i')$
PS: Covariance Pattern Models (CPMs): $Y_i = X_i \beta + \varepsilon_i$, where $cov(Y_i) = \Omega_i$

References

Bailey, M. J., Muth, R. F., and Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, 58(304):933–942.

Case, B. and Quigley, J. (1991). The dynamics of real estate prices. *The Review of Economics and Statistics*, 73(1):50–58.

Diewert, W. E. (2009). The paris oecd-imf workshop on real estate price indexes: Conclusions and future directions. In *Price and Productivity Measurement*, volume 01, pages 87–116.

Dorsey, R. E., Hu, H., Mayer, W. J., and chen Wang, H. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, 19(2):75 – 93.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1–26.

Guo, X., Zheng, S., Geltner, D., and Liu, H. (2014). A new approach for constructing home price indices: The pseudo repeat sales model and its application in china. *Journal of Housing Economics*, 25(Supplement C):20 – 38.

Hedeker, D. and Gibbons, R. D. (2006). *Longitudinal Data Analysis.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.

Kain, J. F. and Quigley, J. M. (1970). Measuring the value of housing quality. *Journal of the American Statistical Association*, 65(330):532–548.

Nagaraja, C. H., Brown, L. D., and Zhao, L. H. (2011). An autoregressive approach to house price modeling. *Ann. Appl. Stat.*, 5(1):124–149.

OECD, ILO, IMF, UNECE, and the World Bank (2013). *Handbook on Residential Property Prices Indices (RPPIs).* INTERNATIONAL MONETARY FUND.

Shiller, R. J. (1991). Arithmetic repeat sales price estimators. *Journal of Housing Economics*, 1(1):110 – 126.

Silver, M. and Graf, B. (2014). *Commercial Property Price Indexes: Problems of Sparse Data, Spatial Spillovers, and Weighting.* IMF Working Papers. INTERNATIONAL MONETARY FUND.

Wang, F. T. and Zorn, P. M. (1997). Estimating house price growth with repeat sales data: What's the aim of the game? *Journal of Housing Economics*, 6(2):93–118.

Thank you !