Variation across price segments and locations: A comprehensive quantile regression analysis of the Sydney housing market

Sofie R. Waltl
University of Graz and University of California at Berkeley

Motivation

Standard house price indices measure *average* movements of *average* houses in *average* locations belonging to an *average* price segment. Such procedures obscure huge variety of price development patterns across price segments and geographical areas, which may be of interest for policy makers, urban planners, home owners, and investors. This paper uses quantile regression techniques to reveal this kind of variation. A hedonic imputation method is developed to compute quality-adjusted price segment- and location-specific house price indices. As location and price levels are highly correlated, it is important to analyze both aspects simultaneously. As a benchmark, results are compared to quantile time-dummy indices. Both approaches are applied to house sales in Sydney between 2001 and 2014. Calculations are based on more than 400,000 observations.
The paper is forthcoming as Waltl (2016).

Methodology

The methodology aims to find a compromise between an analysis that is as comprehensive as possible, and an analysis that generates output that is easy to explain and present, and follow the traditional house price index literature. Hence, I focus on three locations/clusters (inner city, metro-residential, and outer-suburban) and three price segments (top, middle, and bottom).

Time-dummy approach
Hedonic models for quantile-level ϑ are estimated using characteristics X and time-region-dummies D,

$$Q_{\log P_{itr}}(\vartheta|X,D) = X_{itr}\beta(\vartheta) + \sum_{t=1}^{T}\sum_{r=1}^{R} D_{itr}\delta_{tr}(\vartheta).$$

An index number for period t, region r, and segment ϑ normalized to t^* is given by

$$P_t^r(\vartheta) = \frac{\exp(\bar{x}\hat{\beta} + \hat{\delta}_{rt})}{\exp(\bar{x}\hat{\beta} + \hat{\delta}_{rt^*})} = \frac{\exp(\hat{\delta}_{rt}(\vartheta))}{\exp(\hat{\delta}_{rt^*}(\vartheta))},$$

where $\bar{x}$ is an average set of characteristics. Cluster-specific indices are constructed using

$$P_t^{C_i}(\vartheta) = \sqrt[n_i]{\prod_{r=1}^{R} \left(P_t^r(\vartheta)\right)^{1_{C_i}(r)}}.$$

Imputation approach
Estimate separate hedonic models for each year t and quantile-level ϑ including a geographical spline $f(long, lat)$

$$Q_{\log P_{it}}(\vartheta|X, long, lat) = X_{it}\beta_t(\vartheta) + f_t^{\vartheta}(long_{it}, lat_{it}).$$

Predict prices for houses, i.e., sets of characteristics $z_{it} = (X_{it}^{\top}, long_{it}, lat_{it})$, for period s and quantile level ϑ

$$\hat{p}_{is}^{\vartheta}(z_{it}) = \exp\left(X_{it}^{\top}\hat{\beta}_s(\vartheta) + \hat{f}_s^{\vartheta}(long_{it}, lat_{it})\right).$$

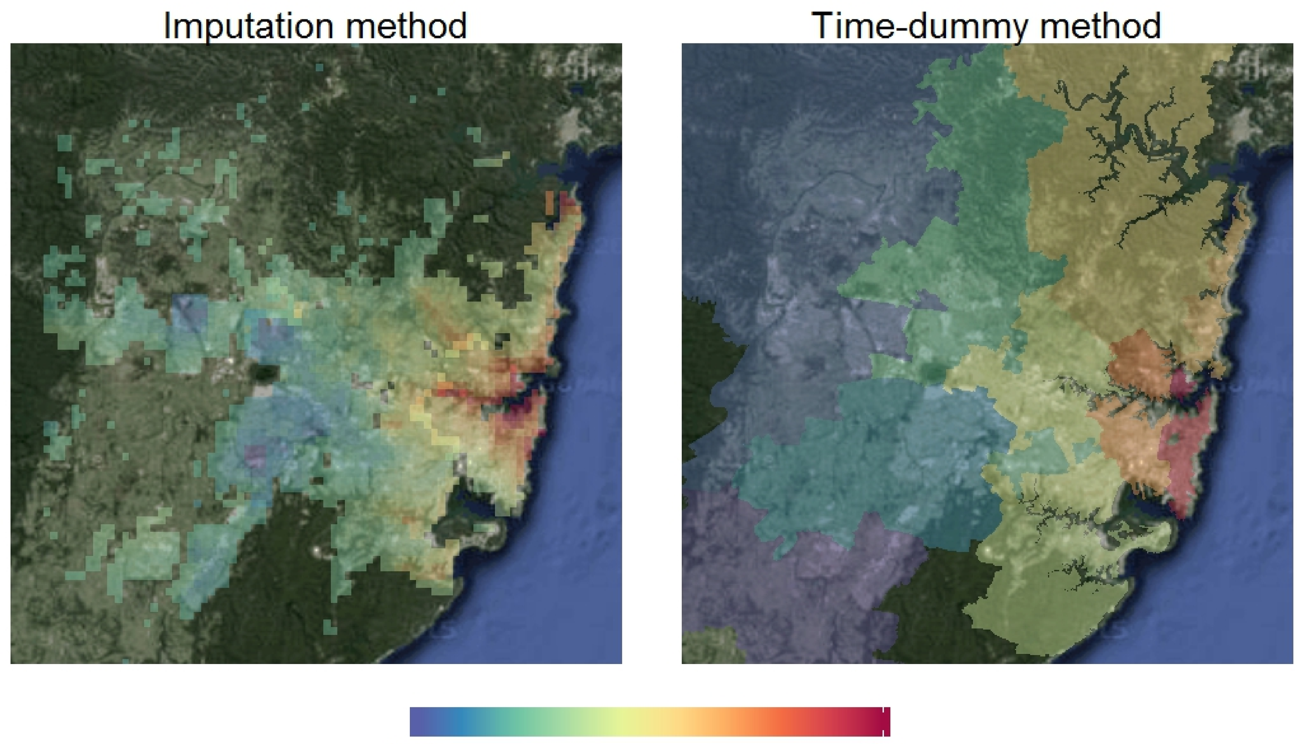
Calculate cluster-specific double-imputation Törnqvist indices:

$$\text{Geometric Laspeyres:}\quad P_{t,s}^{L,c} = \prod_{i=1}^{H_{tc}} \left(\frac{\hat{p}_{is}(z_{itc})}{\hat{p}_{it}(z_{itc})}\right)^{1/H_{tc}},$$

$$\text{Geometric Paasche:}\quad P_{t,s}^{P,c} = \prod_{h=1}^{H_{sc}} \left(\frac{\hat{p}_{is}(z_{isc})}{\hat{p}_{it}(z_{isc})}\right)^{1/H_{sc}},$$

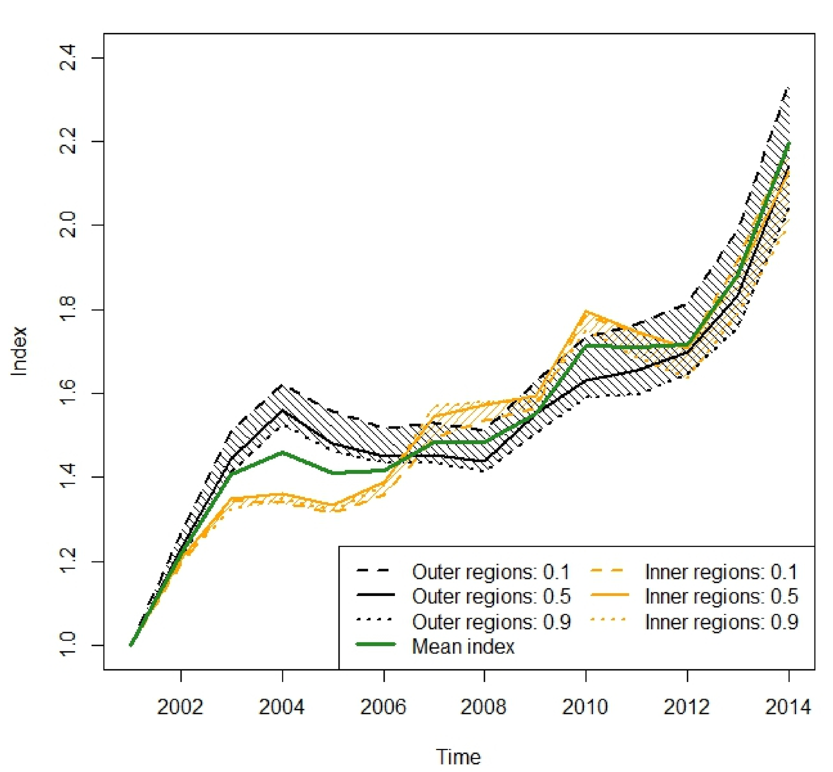
$$\text{Törnqvist:}\quad P_{t,s}^{T,c} = \sqrt{P_{t,s}^{L,c} \cdot P_{t,s}^{P,c}}.$$

Figure 1: Locational effects



Estimated locational effects for the middle segment in 2007. The left panel shows the smoothly estimated effect used for imputation indices. The right panel shows the effect relying on regional dummy variables as used in the time-dummy method. The figures clearly indicate that region dummies do not capture the full set of locational variation.

Figure 2: Price indices

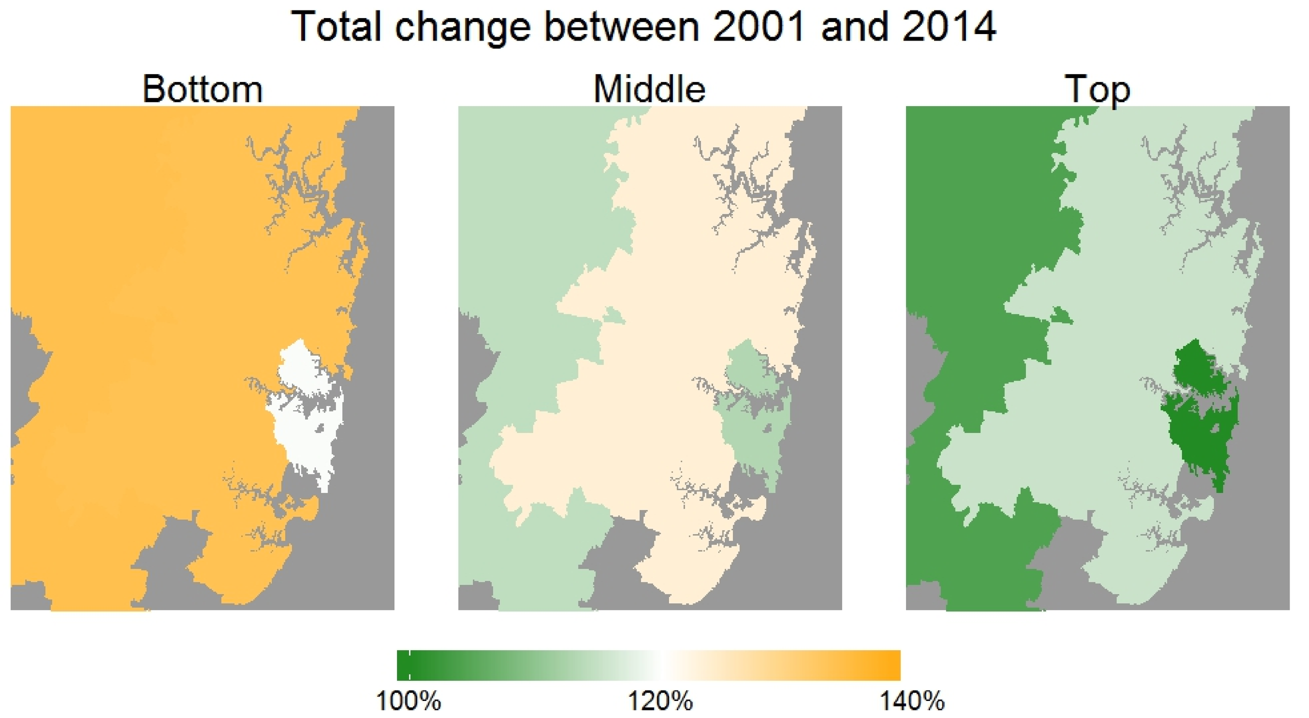


The figure compares cluster-segment-specific imputation indices for the inner and outer cluster, and three price segments to a standard hedonic mean index.

Results and Conclusions

- ✻ The paper's empirical findings are inline with a growing literature suggesting that even within urban areas house price appreciation rates vary substantially (see Guerrieri et al., 2013; Case and Mayer, 1996). It contributes to the yet sparse literature on how to measure such kind of variation (see Coulson and McMillen, 2007; McMillen, 2014).
- ✻ Overall appreciation rates differ strongly across locations and price segments: Prices of suburban homes belonging to the bottom segment increased by roughly 130% between 2001 and 2014 whereas prices of inner-city homes belonging to the top segment increased by roughly 100%.
- ✻ The housing boom peaking in 2004 and declining prices thereafter were mainly driven by price developments of suburban low-priced houses.
- ✻ Results from standard indices and indices for the central segment / metro cluster are very similar, which supports the statement that standard indices measure average movements of average house prices in average locations.
- ✻ Standard indices detect two periods of constant prices, which are driven by different patterns (2007–2008, 2010–2012).
- ✻ Standard errors in general are very low indicating stable results. They are slightly lower for the imputation index.
- ✻ Imputation and time-dummy indices generally report similar results. Strongest deviations are found for inner-city indices, which also shows the highest degree of locational variation. The imputation index measures location more precisely and is hence expected to be more reliable.
- ✻ From a methodological point of view, the imputation approach is superior to the time-dummy approach as it naturally reduces a potential omitted variables bias and uses a more precise technique to control for locational effects. Both approaches rely on location- and segment-specific coefficients. The imputation approach is even more flexible as coefficients may also vary over time. The time-dummy approach becomes impracticable for long periods of time and a large number of sub-regions.

Figure 3: Overall price change



The figure depicts the overall price change between 2001 and 2014 for the bottom (left), middle (center), and top (right) segment. Price increases were largest for suburban, low-priced and lowest for inner-city, high-priced houses.

References

Case, K. E. and Mayer, C. J. (1996). Housing price dynamics within a metropolitan area. *Regional Science and Urban Economics*, 26(3-4):387–407.
Coulson, N. E. and McMillen, D. P. (2007). The dynamics of intraurban quantile house price indexes. *Urban Studies*, 44(8):1517–1537.
Guerrieri, V., Hartley, D., and Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, 100:45–60.
McMillen, D. P. (2014). Local quantile house price indices. *University of Illinois Working Paper*.
Waltl, S. R. (2016). Variation across price segments and locations: A comprehensive quantile regression analysis of the Sydney housing market. *Real Estate Economics*, doi:10.1111/1540-6229.12177.

Acknowledgments

This project has benefited from funding from the Austrian National Bank (Jubiläumsfondsprojekt 14947) and the Council of the University of Graz (JungforscherInnenfonds 2014). I thank Australian Property Monitors for supplying the data on house prices and characteristics. This poster has been prepared for the AREUEA Doctoral Session during the AREUEA-ASSA conference in Chicago in January 2017. I thank AREUEA for providing me with a travel grant.

Contact

sofie.waltl@gmail.com
www.sofiewaltl.com