A Multi-Criteria Evaluation of House Price Indexes

Andy Krause^ – Zillow Group Reid Johnson – Zillow Group

2024-01-28

Abstract

This work refines the current typology

Goals of this paper:

- Refine the current typology of HPIs based on their fundamental trend extraction technique
- Test different classes, including by sample
- Add some ML options to the mix
- Create multi-criteria evaluation framework
- Present data and code (open source) for researchers

Typology

Class - Method - Implementation

- Aggregate
  - Raw (median or mean)
  - Aggregate ($/sf)
- Mathematical/Statistical
  - Control Method
  - Extraction Method
- Imputation
  - Partial (Using Sales a la Hill)
  - Full. (ZHVI)

Each has an implementation method as well. * Chaining * Index creation * Statistical estimators – base, robust, QR * Algorithmic choice – LM vs NN for Math, Any AVM for Imputation * Coefficient extraction – Chaining, Single statistical model, Post Hoc explainable method for ML

Outline

Show aggregate index on median prices.

Why do we create indexes?
* Primary: We don't believe that aggregate indexes control for the sample of homes that sell. * Secondary: There is noise in the observed data, how do we better generalize (bigger problem for small samples)

What questions are we trying to answer?

- What was the median home sale price? – agg
- How have prices/values moved for any given home? – index —- But which sample of homes? —— All sold homes? – making comparable adjustments —— All homes? – tracking stock —— Some special subset of homes – tracking a portfolio —— A specific, individual home — a porfolio of 1

To the extent that samples differ between what sells and what exists, the choice of HPI method should be driven by the question you are attemping to answer

Continuum

- Pure observed sold price changes – Agg
- Observed with controls – $ normalized by some sample features (ex. PPSF)
- All homes that sold twice – RT – Also looks to control for possible unobserved features in the data
- All homes that sold – In the respective periods: Hed Ind/Chained, NN – During the entire time period: Hed Imp w/sales
- All homes – Hedonic Imputation w/Universe

Lit Review

- Taxonomy
- Comparisons / Evaluations * Revisions (Silverstein 2014;) * Robustness * Samnple – RT sample vs all sample

Methods Tests

Eval Methods

Analysis

Index comparison

- Start with 10 year index comparison of the approaches. Talk through combination of sample and algorithm on results.

Volatility

Concept of a Series

Revision

Accuracy (two kinds)

Robustness checks

Time 5 and 20 year indices

Space Submarkets and Areas

Conclusion

- Evaluate indexes based on your goal

•