How to Better Measure Hedonic Residential Property Price Indexes

Mick Silver

Paper presented at the 15th Meeting of the Ottawa Group, 10–12 May 2017, Elves, Germany, hosted by Deutsche Bundesbank.

The context

- **Macroeconomists and central banks need** to identify house price bubbles. Timely, proper measurement.
- **Other purposes include requirement of separation of land prices from structures** – Diewert, Huang, and Burnett-Isaacs, last session. Not the concern this paper.
- Eurostat (2013) ***Handbook on RPPIs****:* Chapter on hedonic methods by de Haan and Diewert (2013)
- G20 Data Gaps Initiative-2, IMF's SDSS plus, and Financial Soundness Indicators
- **Literature**: huge on hedonics; emerging property price indexes; practice. Many here.

The hard problem: requires a constant quality property price index

- Indexes of average prices tainted by changes in the quality-mix of properties transacted
- Matched models breaks down: infrequent transactions of heterogeneous items. Secondary source data
- Three approaches:
  - Repeat sales
  - Sales price appraisal ratio (SPAR)
  - **Hedonic regression**
- Commercial property price indexes even harder
- Erwin Diewert and Chihiro Shimizu; Inês Gonçalves Raposo and Rui Evangelista; and Barra Casey - later session.

Three main ways to compile a hedonic property price index: a practical paper

- Time dummy method:
- Imputation method
- Characteristics method

- *Many variants of each method: includes:*
- *which period the characteristics held constant,* ***superlative***
- *which functional form/aggregators/average of characteristics) linear or* ***semi-logarithmic*** *and* ***arithmetic or*** *geometric for characteristics; and*
- *single or* ***double imputation****.*

Time dummy approach

- **A semi-logarithmic form** is usually appropriate for a hedonic price index, with reference to the constant, β_0 , given as

$$\ln p_i^{0,t} = \beta_0 + \sum_{k=1}^{K} z_{k,i}^{0,t} \ln \beta_k + \sum_{t=1}^{T} \delta^t D_i^t + \varepsilon_i^t$$

- Rolling window advantageous if thin market, but effectively smooths and lags
- Weights can be introduced by WLS (Diewert (2005) but the paper warns of leverage effects.

Hedonic characteristics approach

- Constant period 0 average characteristics

$$\frac{\prod_{k=0}^{K}\left(\bar{z}_k^0\right)^{\hat{\beta}_k^t}}{\prod_{k=0}^{K}\left(\bar{z}_k^0\right)^{\hat{\beta}_k^0}}=\frac{\exp\left(\sum_{k=0}^{K}\bar{z}_k^0 \ln \hat{\beta}_k^t\right)}{\exp\left(\sum_{k=0}^{K}\bar{z}_k^0 \ln \hat{\beta}_k^0\right)}$$

- Constant period *t* average characteristics

$$\frac{\prod_{k=0}^{K}\left(\bar{z}_k^t\right)^{\hat{\beta}_k^t}}{\prod_{k=0}^{K}\left(\bar{z}_k^t\right)^{\hat{\beta}_k^0}}=\frac{\exp\left(\sum_{k=0}^{K}\bar{z}_k^t \ln \hat{\beta}_k^t\right)}{\exp\left(\sum_{k=0}^{K}\bar{z}_k^t \ln \hat{\beta}_k^0\right)}$$

Hedonic imputation indexes: geomeans; double imputation

- Constant period 0 characteristics

$$\frac{\prod_{i \in N^0} \left(\hat{p}^t_{i|z_i^0} \right)^{\frac{1}{N^0}}}{\prod_{i \in N^0} \left(\hat{p}^0_{i|z_i^0} \right)^{\frac{1}{N^0}}} = \frac{\exp\left(\sum_{i \in N^0} \ln \hat{p}^t_{i|z_i^0} \right)}{\exp\left(\sum_{i \in N^0} \ln \hat{p}^0_{i|z_i^0} \right)}$$

- Constant period *t* characteristics

$$\frac{\prod_{i \in N^t} \left(\hat{p}^t_{i|z_i^t} \right)^{\frac{1}{N^0}}}{\prod_{i \in N^t} \left(\hat{p}^0_{i|z_i^t} \right)^{\frac{1}{N^0}}} = \frac{\exp\left(\sum_{i \in N^t} \ln \hat{p}^t_{i|z_i^t} \right)}{\exp\left(\sum_{i \in N^t} \ln \hat{p}^0_{i|z_i^t} \right)}$$

Equivalences: Characteristics and imputation approaches give the same results

- Linear hedonic and arithmetic aggregator (for characteristics)
- Log-linear (semi-log) and arithmetic aggregator
- Log-log (double-log) and geometric aggregator

$$\frac{\prod_{k=0}^{K}\left(\hat{\beta}_k^t\right)^{\bar{z}_k^0}}{\prod_{k=0}^{K}\left(\hat{\beta}_k^0\right)^{z_k^0}}=\frac{\exp\left(\sum_{k=0}^{K}\bar{z}_k^0\ln\hat{\beta}_k^t\right)}{\exp\left(\sum_{k=0}^{K}\bar{z}_k^0\ln\hat{\beta}_k^0\right)}=\frac{\exp\left(\frac{1}{N^0}\sum_{k=0}^{K}\sum_{i\in N^0}z_{i,k}^0\ln\hat{\beta}_k^t\right)}{\exp\left(\frac{1}{N^0}\sum_{k=0}^{K}\sum_{i\in N^0}z_{i,k}^0\ln\hat{\beta}_k^0\right)}=\frac{\exp\left(\frac{1}{N^0}\sum_{i\in N^0}\sum_{k=0}^{K}z_{k,i}^0\ln\hat{\beta}_k^t\right)}{\exp\left(\frac{1}{N^0}\sum_{i\in N^0}\sum_{k=0}^{K}z_{k,i}^0\ln\hat{\beta}_k^0\right)}=\frac{\prod_{i\in N^0}\left(\hat{p}_{i|z_i^0}^t\right)^{\frac{1}{N^0}}}{\prod_{i\in N^0}\left(\hat{p}_{i|z_i^0}^0\right)^{\frac{1}{N^0}}}$$

- Axiomatic property
- Hill and Melser (2008); Hill (2013); de haan and Diewert (2013); Rambaldi

Weights – A question:

- Why not weight each transaction's price change by its relative period 0 (period t) values?

$$\frac{\prod_{i\in N^0}\left(\hat{p}^t_{i|z_i^0}\right)^{\frac{\hat{p}^0_{i|z_i^0}}{\sum \hat{p}^0_{i|z_i^0}}}}{\prod_{i\in N^0}\left(\hat{p}^0_{i|z_i^0}\right)^{\frac{\hat{p}^0_{i|z_i^0}}{\sum \hat{p}^0_{i|z_i^0}}}} = \prod_{i\in N^0}\left(\frac{\hat{p}^t_{i|z_i^0}}{\hat{p}^0_{i|z_i^0}}\right)^{\frac{\hat{p}^0_{i|z_i^0}}{\sum_{i\in N^0}\hat{p}^0_{i|z_i^0}}}$$

A second question

Why not weight each transaction using "quasi-superlative" index number formula?

$$\frac{\prod_{i \in N^0} \left(\hat{p}^t_{i|z_i^0} \right)^{\hat{w}_i^\tau}}{\prod_{i \in N^0} \left(\hat{p}^0_{i|z_i^0} \right)^{w_i^\tau}} = \prod_{i \in N^0} \left(\frac{\hat{p}^t_{i|z_i^0}}{\hat{p}^0_{i|z_i^0}} \right)^{\hat{w}_i^\tau}$$

$$\text{where } \hat{w}_i^\tau = \frac{1}{2} \left(\frac{\hat{p}^0_{i|z_i^0}}{\sum_{i \in N^0} \hat{p}^0_{i|z_i^0}} + \frac{\hat{p}^t_{i|z_i^0}}{\sum_{i \in N^0} \hat{p}^t_{i|z_i^0}} \right)$$

And a third..

- Why is it only quasi-superlative?
- Use of period 0 and period *t* transactions requires:

$$\prod_{i \in S(0\neg t)} \left(\frac{\hat{p}^{t}_{i|z_i^0}}{\hat{p}^{0}_{i|z_i^0}} \right)^{\hat{w}_i^{\tau} \times \frac{v_{0\neg t}}{V}} \times \prod_{i \in S(t\neg 0)} \left(\frac{\hat{p}^{t}_{i|z_i^t}}{\hat{p}^{0}_{i|z_i^t}} \right)^{\hat{w}_i^{\tau} \times \frac{v_{t\neg 0}}{V}} \times \prod_{i \in S(0\cap t)} \left(\frac{p^{t}_{i|z_i^{0\cap t}}}{p^{0}_{i|z_i^{0\cap t}}} \right)^{w_i^{\tau} \times \frac{v_{0\cap t}}{V}}$$

- Feenstra (1995), Ioannidis and Silver (1999), Silver and Heravi (2005), Diewert (2005), Diewert, Heravi, Silver (2009), de Haan (2009), de Haan and Gong (2013), Rambaldi and Rao (2013) and on stock vs transaction weights, Mehrhoff and Triebskorn (2016).

And differs from

$$\sqrt{\prod_{i \in S(0 \neg t)} \left[\frac{\hat{p}^{t}_{i|z_i^0}}{\hat{p}^{0}_{i|z_i^0}} \right]^{w_i^0} \times \prod_{S(t \neg 0)} \left[\frac{\hat{p}^{t}_{i|z_i^t}}{\hat{p}^{0}_{i|z_i^t}} \right]^{w_i^t}}$$

- Hill and Melser (2008)
- Akin to a Fisher: Laspeyres and Paasche cross
- Substitution effect; use of predicted vs. raw weights.

What the paper does..

- **Equivalences:** finds equivalences for reasonable forms of the imputation and characteristics approaches. Cuts down on choice by consolidating approaches and the many types of each. Validates them - axiomatic.
- **Weights:** shows how weights can be introduced at lower level - for price changes of *individual properties* within a strata.
- **Substitution effects:** shows how substitution effects can be included via a "quasi" superlative formulation – redefines a superlative index.
- **Re-visits the theory** on superlative hedonic RPPIs.

Also, ..

- In the practical context of thin markets – **sparse data -** and vagrancies of regular hedonic estimation

- **Only estimates a reference period hedonic regression** – with regular re-linking.

$$\prod_{i \in N^t} \left[\frac{p^t_{i|z^t_i}}{\hat{p}^0_{i|z^t_i}} \right]^{\left(\frac{\hat{p}^0_{i|z^t_i}}{\sum_{i \in N^t} \hat{p}^0_{i|z^t_i}} + \frac{p^t_{i|z^t_i}}{\sum_{i \in N^t} p^t_{i|z^t_i}} \right)/2} = \exp\left(\sum_{i \in N^t} \left(\hat{w}^0_i + w^t_i \right) / 2 \left[\ln p^t_{i|z^t_i} - \ln \hat{p}^0_{i|z^t_i} \right] \right)$$

  - Sample selectivity bias but limited substitution bias
  - **Use an extended reference period** for thin markets – sparse data - with regular re-linking, re-estimation.

But needs double imputation workarounds

For weights

$$w_i^{*t} = \frac{p_i^t \left(\frac{\hat{p}^0_{i|z_i^0}}{p_i^0} \right)}{\sum_{i \in N^t} p_i^t \left(\frac{\hat{p}^0_{i|z_i^0}}{p_i^0} \right)} = w_i^t e_i^0$$

$$w_i^{*\tau} = 0.5\left(\hat{w}_i^0 + w_i^{*t} \right)$$

For prices

$$\hat{p}^{**t}_{i|z_i^t} = p^t_{i|z_i^t} \left(\frac{\hat{p}^0_{i|z_i^t}}{p_i^0} \right) \simeq \hat{p}^t_{i|z_i^t}$$

Use an indirect volume measure

Value index/volume index=implicit price index

$$\prod_{i \in N^0} \left(\frac{\hat{p}^0_{i|z_i^t}}{\hat{p}^0_{i|z_i^0}} \right)^{\hat{w}_i^\tau} = \frac{\prod_{i \in N^0} \left(\hat{p}^0_{i|z_i^t} \right)^{\hat{w}_i^\tau}}{\prod_{i \in N^0} \left(\hat{p}^0_{i|z_i^0} \right)^{\hat{w}_i^\tau}}$$

The end