CPPI Data Sources and Transaction/Appraisal Based Indices: What to do.

Mick Silver

Paper Presented at the Workshop on Commercial Property Price Indices, Central Bank of the Republic Of Turkey, Cappadocia, Turkey, 7-8 May 2018

The context

- **Macroeconomists and central banks need** to identify property price bubbles, the factors that drive them, instruments that contain them, and to analyze their relation to recessions.
- **Needs of Monetary Policy Committees:** Timely, proper measurement.
- **Guidelines on measurement:**

o **On RPPIs:** Eurostat *et al*. (2103) *Handbook on Residential Property Price Indices.* http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/methodology/hps/rppi_handbook

o **On CPPIs:** *Eurostat (2017), Commercial Property Price Indicators: Sources, Methods and Issues.* *http://ec.europa.eu/eurostat/web/products-statistical-reports/-/KS-FT-16-001*

Transaction-price CPPIs: Challenges: compilation of CPPIs highly problematic

- *Heterogeneous*
  - Within and between: office, retail, industrial, apartments, hotels
  - Quality considerations – capital renovations and depreciation
- *Sparse data*
  - Irregular transactions complicate compilation of fixed-quality price indexes over time
  - Transactions not representative of the population of commercial properties: weight by stock

Problem of limited transactions on heterogeneous properties is worse when measurement really matters – as we go into and during recessions

Appraisal price index

- Appraisal price indices are changes in the estimated amount for which properties should exchange in an arm's length transaction.
- Distinguish 'Appraisal-based CPPIs' from *Sale Price Appraisal Ratio* (SPAR) or transaction-based appraisal-regressor indices.
- Appraisal-based indices are designed as measures of change in commercial property investment performance: include the change in the capital value of the property assets.
- They avoid the problem of limited transaction data?

Appraisal/Valuation data

- In practice, appraisals are usually annual; quarterly data are interpolated or (stale) estimates by the manager/owner of the property used largely based on the last formal appraisal.
- Guidelines to professional appraisers are to base their appraisal on the transactions of similar properties currently in the market, introducing circularity in the argument that appraisals solve the problem of sparse data.
- Criteria for appraisal and timeliness varies within and between countries.
- Reliability of index depends on reliability of appraisers: judgmental; “stale” element; bias
- Sample usually of large properties in a portfolio whose composition changes over time.
- Information on capital expenditures and depreciation are used in appraisal-based indices, as a means for quality adjustment between appraisals. These variables are inappropriately defined for the needs of CPPIs.
- Problems include smoothing and lagged representation of commercial property price inflation. Shimizu, Nishimura, Geltner, Fisher..

CPPIs: Way forward

- Concern that existing appraisal indexes will be used with limited understanding of their deficiencies
- Development of hedonic transaction methods more robust to sparse data
- Segmentation of commercial markets into those prone to problems of sparse data
- More applied studies: papers at this conference

Three main ways to compile a *hedonic* property price index: a practical paper (Silver, 2018, EURONA)

- 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

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

- Constant **period** ***t*** average characteristics

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

Hedonic imputation indexes: geo-means; 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

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

Table 1, Equivalences of hedonic approaches		
Hedonic regression: functional form	Characteristics approach: form of average of characteristics	Imputation approach: Form of average of predicted prices
Linear	Arithmetic mean	Arithmetic mean
Log-linear	Arithmetic mean	Geometric mean
Log-log	Geometric mean	Geometric mean

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

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

- Axiomatic property
- Hill and Melser (2008); Hill (2013); de Haan and Diewert (2013); Rambaldi and Fletcher (2014); Silver (2017)

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?

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

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); Silver (2018); 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

- O**nly 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

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

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

For prices

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

How to estimate hedonic imputation CPPIs in three lines of code: STATA

*Using data lnp location size - stacked by quarter

>regress lnp i.location##c.size if quarter>0 & quarter<5

> predict lnp5 if quarter==5

> summarize lnp5

* Include i.quarter if time dummy

Regression diagnostics: STATA

*Cook's distance

> **predict d, cooksd**

> **list size d if d>4/n**

*Shapiro-Wilk W test for normality

> **predict r, resid**

> **swilk r**

*Breusch-Pagan test

> **estat hettest**

*(Ramsey RESET test) for omitted variables.

> **ovtest**

*Multicollinearity *variance inflation factor*

> **vif**

* SE Robust: other bias corrections: hc2 hc3

> **vce (robust)**