# An Evaluation of the Hedonic Methods Used by European Countries to Compute their Official House Price Indices

**Robert J. Hill*, Michael Scholz*, Chihiro Shimizu**, and Miriam Steurer***
*Department of Economics, University of Graz, Austria
robert.hill@unii-graz.at, michael.scholz@uni-graz.at, miriam steurer@uni-graz.at
**Nihon University, Tokyo, shimizu.chihiro@nihon-u.ac.jp

5 May 2017

## Abstract

Since 2012 Eurostat requires the national statistical institutes (NSIs) in all European Union (EU) countries to compute official House Price Indices (HPIs) at a quarterly frequency. Furthermore, Eurostat recommends computing the HPI using a hedonic method. Most NSIs have followed this advice, although they differ in their choice of hedonic method as follows:
(i) Repricing: used by Austria, Finland, Hungary, Italy, Latvia, Luxembourg, Norway, Slovenia;
(ii) Average characteristics: used by Romania, Spain;
(iii) Hedonic imputation: used by Germany, UK;
(iv) Rolling time dummy (RTD): used by Croatia, Cyprus, France, Ireland, Portugal.
We evaluate the theoretical and empirical properties of these hedonic methods. Our empirical comparisons use detailed micro-level data sets for Sydney and Tokyo, containing about 867 000 actual housing transactions. Our main findings are that the results are generally quite robust to the choice of hedonic method. However, the repricing method starts to become unreliable over longer time horizons. This is a concern given that it is the mostly widely used hedonic method by NSIs in Europe. We recommend that NSIs using repricing switch to one of the other three hedonic methods, or if repricing is used, the reference shadow prices should be updated every five years. (JEL: C43, E31, R31)

**Keywords:** Housing market; Price index; Repricing; Average characteristics; Hedonic imputation; Rolling time dummy

**Preliminary draft to be presented at 15th Ottawa Group Meeting, 10-12 May 2017, Eltville am Rhein, Germany**

Acknowledgement: Hill and Steurer were members of the Expert Team advising Eurostat on the treatment of Owner Occupied Housing (OOH) in the Harmonized Index of Consumer Prices (HICP). The idea for this paper developed out of this project.

Note: An Appendix listing the sources for the methods used by each country will be included in the next draft.

# 1 Introduction

The fundamental role played by housing in the broader economy has been demonstrated by the global financial crisis of 2007-2011, which began in the US housing market. It is essential therefore that governments, central banks and market participants are kept well informed of trends in house prices.

In Europe, Eurostat – the statistical institute of the European Union (EU)– has required since 2012 (see Eurostat 2016) that the national statistical institutes (NSIs) in all EU member countries compute official house price indices (HPIs).

House price indices, however, can be highly sensitive to the method of construction, and this sensitivity can be a source of confusion amongst users (see Silver 2011). In a European context it is also important that the HPIs of different countries are reasonably comparable, especially in the Eurozone where the HPIs are needed by the European Central Bank for its decisions on monetary policy, financial regulation, and the monitoring of financial stability.

The problem is that every house is different both in terms of its physical characteristics and its location. House price indices need to take account of these quality differences. Otherwise the price index will confound price changes and quality differences. The importance of these measurement problems has been recently recognized by the international community. The European Commission, Eurostat, the UN, ILO, OECD, World Bank and IMF together commissioned a Handbook on Residential Property Price Indices that was completed in 2013 (see European Commission et al., 2013).

Hedonic methods – which express house prices as a function of a vector of characteristics – are ideally suited for constructing quality-adjusted house price indices (see Diewert 2011 and Hill 2013). Eurostat recommends that the HPI should be computed using a hedonic approach, but has not provided guidance to NSIs as to which hedonic method should be used. As a result, different countries have adopted different methods. In total six different methods are being used: (i) Repricing: used by Austria, Finland, Hungary, Italy, Latvia, Luxembourg, Norway, Slovenia;

(ii) Average characteristics: used by Romania, Spain;

(iii) Hedonic imputation: used by Germany, UK;

(iv) Rolling time dummy (RTD): used by Croatia, Cyprus, France, Ireland, Portugal; (v) Mix adjusted (or stratified) median: used by Bulgaria, Estonia, Lithuania, Poland, Slovakia; (vi) Sales Price Appraisal Ratio (SPAR): used by Denmark, the Netherlands, and Sweden.

The first four methods are hedonic. Method (v) by averaging medians across strata provides some partial quality adjustment, although to the same extent as a hedonic method. Method (vi) combines actual prices with expert valuations (see de Haan, Wal, Steege, and Vries 2006).

For each method, the taxonomy can be further refined, in that two countries using the same basic method in some cases differ slightly in the way it is formulated. For example, with regard to the RTD method, some countries use a two quarter rolling window while others use a five quarter window.

Our objective here is to evaluate the theoretical and empirical properties of the methods (i), (ii), (iii), (iv) and (v) used by NSIs in Europe to compute their HPIs. We do not consider method (vi) – the SPAR method – since for our data sets we do not have access to any expert valuations.

It is shown that the underlying structures of the repricing, average characteristics, and hedonic imputation methods share some common features. The RTD method is somewhat different in its approach.

Empirically we compare the hedonic methods and mix-adjusted medians using detailed micro-level data for Sydney and Tokyo. These data sets were chosen for two reasons. First, they together contain about 867 000 actual housing transactions. Second, they cover quite long time spans. The Sydney data covers 11 years, while the Tokyo data covers 30 years. When comparing hedonic methods, it is important to have a sufficiently long time series, since problems of drift or bias may only emerge over these kinds of time horizons.

The empirical comparisons have two main objectives. The first it to establish how sensitive the HPI is to the choice of hedonic method. The second is to see whether any of the hedonic methods (computed on a quarterly basis) behave in anomalous ways, particularly over longer time horizons (e.g., 10+ years). This is potentially a concern especially for the widely used repricing method, which extrapolates to later periods using the estimated characteristic shadow prices of the base period.

To understand how these hedonic methods perform in practice it is important that they are compared using real housing data sets, rather than just simulated data. The large number of observations and long-time horizons make the Sydney and Tokyo data sets ideal for empirically testing the performance of hedonic methods. Also by evaluating EU methods using non-EU data we provide an independent check on method selection.

Eurostat recommends that each NSI compute separate hedonic indices for houses and apartments. We consider how these indices can be aggregated to obtain an overall HPI. Furthermore, an NSI may want to compute separate hedonic price indices for new housing. Indices for new housing are needed for the owner occupied housing price index (OOHI) which is being used on an experimental basis in the Harmonized Index of Consumer Prices (HICP) (see Eurostat 2016). For the case of Australia, we are able to compute separate indices for houses and apartments, bit not for new dwellings (since we do not have age as a characteristic). For Tokyo, by contrast, we are able to compute a price index for new apartments, but we are not able to compute any price indices for houses.

The remainder of the paper is structured as follows. Section 2 explains the theoretical properties of the hedonic methods used by NSIs in Europe to compute their HPIs. The problem of combining price indices across strata is considered in section 3. The hedonic methods are compared empirically using data for Sydney in section 4, and using data for Tokyo in section 5. Our main findings are then summarized in the conclusion in section 6.

# 2 Some Alternative Methods for Constructing Hedonic House Price Indices

All the methods considered here are formulated to be compatible with Eurostat guidelines. In other contexts, these methods could be structured in different ways.

## 2.1 Repricing method

The repricing method is currently the most widely used hedonic method for computing the HPI in Europe. It is used by the NSIs of Austria, Finland, Hungary, Italy, Latvia, Luxembourg, Norway, and Slovenia.

The repricing method begins by estimating a semilog hedonic model using only the data of the first year in the data set. The first year is called year 1.

$$\ln p_{(1,q),h} = \sum_{c=1}^{C} \beta_{1,c} z_{(1,q),h,c} + \varepsilon_{(1,q),h}, \tag{1}$$

where $h$ denotes a dwelling sold in year 1, $(1,q)$ is the quarter in year 1 in which the dwelling was sold, $c = 1,\ldots,C$ indexes the characteristics of dwellings available in the data set (such as number of bedrooms, and land area), and $\varepsilon$ is a random error term.

The objective of estimating the hedonic model in (4) is to compute the shadow prices $\beta_{1,c}$ on the characteristics. These shadow prices are computed using the whole year's data.

As it is typically applied in the HPI, the repricing method compares period $(t,q-1)$ with period $(t,q)$ using the base year's shadow price vector $\hat{\beta}_1$.

The repricing price index formula consists of two components: a quality unadjusted price index (QUPI) and a quality adjustment factor (QAF). The QUPI is the ratio of the geometric mean prices in both periods $(t,q-1)$ and $(t,q)$, computed as follows:

$$QUPI_{(t,q),(t,q-1)} = \frac{\tilde{p}_{(t,q)}}{\tilde{p}_{(t,q-1)}},$$

where $\tilde{p}_{(t,q-1)}$ and $\tilde{p}_{(t,q)}$ denote, respectively, the geometric mean price of dwellings sold in year $t$-quarter $q-1$ and year $t$-quarter $q$.

$$\tilde{p}_{(t,q-1)} = \prod_{h}^{H_{(t,q-1)}} \left(p_{(t,q-1),h}\right)^{1/H_{(t,q-1)}}, \quad \tilde{p}_{(t,q)} = \prod_{h}^{H_{(t,q)}} \left(p_{(t,q),h}\right)^{1/H_{(t,q)}},$$

where $H_{(t,q-1)}$ and $H_{(t,q)}$ denote the number of properties sold in $(t,q-1)$ and $(t,q)$ respectively. Arithmetic means could be used instead. However, geometric means have the advantage of being less sensitive to outliers, which is important since the house price distribution is typically highly right skewed.

The next step is to compute a quality adjustment factor (QAF). This is done by using shadow prices of year 1 as a point of reference to compare quality of the average dwelling sold in periods $(t, q-1)$, and $(t, q)$. The formula of the quality adjustment factor is as follows:

$$QAF_{(t,q-1),(t,q)} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q-1),c})},$$

where

$$\bar{z}_{(t,q-1),c} = \frac{1}{H_{(t,q-1)}} \sum_{h=1}^{H_{(t,q-1)}} z_{(t,q-1),h,c}, \quad \bar{z}_{(t,q),c} = \frac{1}{H_{(t,q)}} \sum_{h=1}^{H_{(t,q)}} z_{(t,q),h,c},$$

denote the average basket of characteristics of periods $(t, q-1)$ and $(t, q)$, respectively, computed using the arithmetic mean formula. In the case of dummy variables, such as postcodes, the average measures the proportion of transactions that feature that postcode. For example, if 1 percent of the transactions occur in postcode 1, then teh average basket for postcode 1 equals 0.01.

The quality adjustment factor can be rewritten as follows:

$$QAF_{(t,q-1),(t,q)} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{1,c})} \Bigg/ \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q-1),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{1,c})} = \frac{Q^L_{1,(t,q)}}{Q^L_{1,(t,q-1)}},$$

where $Q^L_{j,k}$ denotes a Laspeyres quantity index between periods $j$ and $k$. Hence the quality adjustment factor is the ratio of two Laspeyres quantity indices.

The repricing price index is now obtained by dividing the quality-unadjusted index (QUPI) by the quality adjustment factor (QAF) as follows:

$$\frac{P_{(t,q)}}{P_{(t,q-1)}} = \frac{QUPI_{(t,q),(t,q-1)}}{QAF_{(t,q-1),(t,q)}} = \frac{\tilde{p}_{(t,q)}}{\tilde{p}_{(t,q-1)}} \Bigg/ \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q-1),c})} = \frac{\tilde{p}_{(t,q)}}{\tilde{p}_{(t,q-1)}} \Bigg/ \frac{Q^L_{1,(t,q)}}{Q^L_{1,(t,q-1)}}. \tag{2}$$

More generally, relative to the first quarter in the data set (1,1), the price index for period $(t, q)$ is calculated as follows:

$$\frac{P_{(t,q)}}{P_{(1,1)}} = \frac{\tilde{p}_{(t,q)}}{\tilde{p}_{(1,1)}} \Bigg/ \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{1,c} \bar{z}_{(1,1),c})} = \frac{\tilde{p}_{(t,q)}}{\tilde{p}_{(1,1)}} \Bigg/ \frac{Q^L_{1,(t,q)}}{Q^L_{1,(1,1)}}. \tag{3}$$

An interesting feature of the repricing method is that it only requires the hedonic model to be estimated once (in the base year). This is perhaps one reason why it has proved popular with NSIs.

The base year under the repricing method could be updated at regular time intervals. However, as far as we are aware, none of the NSIs in the EU using the repricing method have done this yet. In the empirical comparisons in sections 3 and 4, we will consider two versions of the repricing method. The first never updates the base year, while the second updates it every five years.

## 2.2 Average characteristics method

The average characteristics method and the hedonic imputation method both begin by estimating the following semilog hedonic model separately for each period $(t, q)$:

$$\ln p_{(t,q),h} = \sum_{c=1}^{C} \beta_{t,q),c} z_{(t,q),h,c} + \varepsilon_{(t,q),h}, \tag{4}$$

where $h$ indexes the dwelling transactions in period $(t, q)$, $p_{(t,q),h}$ the transaction price, and $z_{(t,q),h,c}$ is the level of characteristic $c$ in dwelling $h$. Unlike under the repricing method, the estimated shadow prices on the characteristics, $\beta_{(t,q),c}$, are specific to period $(t, q)$ and are updated every period.

The next step is to construct an average basket of characteristics. The hedonic method then measures the change in the imputed price of the average dwelling over time. The version used by European NSIs computes an average basket $\bar{z}_{t,c}$ based on a whole year's data calculated using the arithmetic mean formula. The price index between two adjacent quarters in the same year therefore is now calculated as follows:

$$\frac{P_{(t,q)}}{P_{(t,q-1)}} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q),c} \bar{z}_{t-1,c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q-1),c} \bar{z}_{t-1,c})}$$

$$= \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q),c} \bar{z}_{t-1,c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{t-1,c} \bar{z}_{t-1,c})} \Bigg/ \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q-1),c} \bar{z}_{t-1,c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{t-1,c} \bar{z}_{t-1,c})} = \frac{P^L_{t-1,(t,q)}}{P^L_{t-1,(t,q-1)}},$$

where $P^L_{j,k}$ denotes a Laspeyres price index between periods $j$ and $k$. Hence the quality adjustment factor is the ratio of two Laspeyres price indices.

Once a year, the average characteristic basket is updated. This can be done at the end of the year, once all the data for that year are available. The price index between the fourth quarter in one year and the first quarter in the next year therefore is calculated as follows:

$$\frac{P_{(t+1,1)}}{P_{(t,4)}} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t+1,1),c} \bar{z}_{t,c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,4),c} \bar{z}_{t,c})}$$

$$= \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(t+1,1),c}\bar{z}_{t,c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{t,c}\bar{z}_{t,c})} \Bigg/ \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(t,4),c}\bar{z}_{t,c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{t,c}\bar{z}_{t,c})} = \frac{P^L_{t,(t+1,1)}}{P^L_{t,(t,4)}}.$$

Again the quality adjustment factor is the ratio of two Laspeyres price indices.

Relative to the first quarter in the data set (1,1), the price index for period $(t+1,1)$ is calculated as follows:

$$\frac{P_{(t+1,1)}}{P_{(1,1)}} = \frac{P^L_{0,(1,2)}}{P^L_{0,(1,1)}} \times \frac{P^L_{0,(1,3)}}{P^L_{0,(1,2)}} \times \frac{P^L_{0,(1,4)}}{P^L_{0,(1,3)}} \times \frac{P^L_{1,(2,1)}}{P^L_{1,(1,4)}} \times \cdots \times \frac{P^L_{t-1,(t,4)}}{P^L_{t-1(t,3)}} \times \frac{P^L_{t,(t+1,1)}}{P^L_{t,(t,4)}}.$$

It turns out that the repricing method can be represented as a fixed base average characteristics method. Suppose the hedonic model is estimated for a single quarter, as is the case for the average characteristics and hedonic imputation methods. The imputed errors from the semilog hedonic model for quarter $s$ can be written as follows:

$$\hat{\varepsilon}_{sh} = \ln p_{sh} - \sum_{c=1}^{C}\hat{\beta}_{s,c}z_{shc}.$$

By construction under OLS, $\sum_{h=1}^{H_1}\hat{\varepsilon}_{sh} = 0$. Hence,

$$\sum_{h=1}^{H_s}[\ln p_{sh} - \sum_{c=1}^{C}\hat{\beta}_{s,c}z_{shc}] = 0,$$

which in turn implies that the geometric mean price takes the following form:

$$\tilde{p}_s = \exp(\sum_{c=1}^{C}\hat{\beta}_{s,c}\bar{z}_{sc}).$$

Substituting this expression into the repricing formula (with shadow prices estimated using only the first quarter not the first year) yields the following:

$$\frac{\bar{p}_{(t,q)}}{\bar{p}_{(t,q-1)}} \Bigg/ \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(1,1),c}\bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{(1,1),c}\bar{z}_{(t,q-1),c})} = \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(t,q),c}\bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{(t,q-1),c}\bar{z}_{(t,q-1),c})} \Bigg/ \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(1,1),c}\bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{(1,1),c}\bar{z}_{(t,q-1),c})}$$

$$= \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(t,q),c}\bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{1,1,c}\bar{z}_{t,q,c})} \Bigg/ \frac{\exp(\sum_{c=1}^{C}\hat{\beta}_{(1,1),c}\bar{z}_{(t,q-1),c})}{\exp(\sum_{c=1}^{C}\hat{\beta}_{(t,q-1),c}\bar{z}_{(t,q-1),c})} = \frac{P^P_{(1,1),(t,q)}}{P^P_{(1,1),(t,q-1)}},$$

where $P^P_{j,k}$ denotes a Paasche price index between periods $j$ and $k$. Hence the repricing method can also be interpreted as an average characteristics method that uses the Paasche price index formula.

## 2.3 Hedonic imputation method

The underlying rationale of the hedonic imputation method is to use the hedonic model to ask counterfactual questions regarding what a particular dwelling would have sold for if it had sold in a different period from the one it actually sold in. In this way it is possible to construct price relatives for each dwelling measuring how its price has changed over time. These price relatives can then be averaged to obtained the overall price index.

Here we will present two slightly different variants of the hedonic imputation method. The first is used by the UK NSI and the second by the German NSI.

The common feature of both versions is that they use the estimated hedonic model to impute prices for each dwelling. For example, let $\hat{p}_{(t,q),h}(z_{t-1,h})$ denote an imputed price in period $(t, q)$ for dwelling $h$ which was actually sold one year earlier in period $(t-1, q)$.

The UK version is a chained Lowe index where the reference basket is all the dwellings sold in the previous year. When comparing two quarters in the same year (here $t$), the formula is as follows:

$$\frac{P_{(t,q)}}{P_{(t,q-1)}} = \prod_{h=1}^{H_{t-1}} \left[ \frac{\hat{p}_{(t,q),h}(z_{t-1,h})}{\hat{p}_{(t,q-1),h}(z_{t-1,h})} \right]^{1/H_{t-1}},$$

where $H_{t-1}$ denotes the number of properties sold in year $t-1$. When the 4th quarter is compared with the 1st quarter of the next year the references basket is updated as follows:

$$\frac{P_{(t+1,1)}}{P_{(t,4)}} = \prod_{h=1}^{H_t} \left[ \frac{\hat{p}_{(t+1,1),h}(z_{t,h})}{\hat{p}_{(t,4),h}(z_{t,h})} \right]^{1/H_t}.$$

When the underlying hedonic model is semilog, the UK method is in fact identical to the average characteristics method described above. This duality between the average characteristics method and the hedonic imputation method is explored in more detail in Hill and Melser (2008). In the case of the UK method, the duality can be demonstrated as follows:

$$\frac{P_{(t,q)}}{P_{(t,q-1)}} = \prod_{h=1}^{H_{t-1}} \left[ \frac{\hat{p}_{(t,q),h}(z_{t-1,h})}{\hat{p}_{(t,q-1),h}(z_{t-1,h})} \right]^{1/H_{t-1}} = \prod_{h=1}^{H_{t-1}} \left[ \frac{\sum_{c=1}^{C} \exp(\hat{\beta}_{(t,q)} z_{t-1,h})}{\sum_{c=1}^{C} \exp(\hat{\beta}_{(t,q-1)} z_{t-1,h})} \right]^{1/H_{t-1}}$$

$$= \frac{\frac{1}{H_{t-1}} \sum_{c=1}^{C} \sum_{h=1}^{H_{t-1}} \exp(\hat{\beta}_{(t,q)} z_{t-1,h})}{\frac{1}{H_{t-1}} \sum_{c=1}^{C} \sum_{h=1}^{H_{t-1}} \exp(\hat{\beta}_{(t,q-1)} z_{t-1,h})} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q),c} \bar{z}_{t-1,c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q-1),c} \bar{z}_{t-1,c})} = \frac{P^{L}_{t-1,(t,q)}}{P^{L}_{t-1,(t,q-1)}}.$$

In an analogous way it can be shown that

$$\frac{P_{(t+1,1)}}{P_{(t,4)}} = \prod_{h=1}^{H_t} \left[ \frac{\hat{p}_{(t+1,1),h}(z_{t,h})}{\hat{p}_{(t,4),h}(z_{t,h})} \right]^{1/H_t} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t+1,1),c} \bar{z}_{t,c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,4),c} \bar{z}_{t,c})} = \frac{P^{L}_{t,(t+1,1)}}{P^{L}_{t,(t,4)}}.$$

The German version by contrast uses a Törnqvist-type formula (i.e., the geometric mean of geometric-Laspeyres and geometric-Paasche-type formulas) defined as follows:

$$\text{Geometric Laspeyres (GL)}: \quad \frac{P_{(t,q)}}{P_{(t,q-1)}} = \left[ \prod_{h=1}^{H_{(t,q-1)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q-1),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q-1),h})} \right]^{1/H_{(t,q-1)}};$$

$$\text{Geometric Paasche (GL)}: \quad \frac{P_{(t,q)}}{P_{(t,q-1)}} = \left[ \prod_{h=1}^{H_{(t,q)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q),h})} \right]^{1/H_{(t,q)}}$$

.

$$\text{Törnqvist}: \quad \frac{P_{(t,q)}}{P(t,q-1)} = \left\{ \left[ \prod_{h=1}^{H_{(t,q-1)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q-1),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q-1),h})} \right]^{1/H_{(t,q-1)}} \left[ \prod_{h=1}^{H_{(t,q)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q),h})} \right]^{1/H_{(t,q)}} \right\}^{1/2}.$$

Here it makes no difference whether we are comparing two quarters in the same year or the last quarter in one year with the first quarter in the next.

When the underlying hedonic model is semilog, the geometric-Laspeyres (GL), geometric-Paasche (GP), and Törnqvist hedonic imputation indices can be represented as average characteristic methods as follows:

$$\text{GL}: \quad \left[ \prod_{h=1}^{H_{(t,q-1)}} \frac{\hat{p}_{(t,q,h)}(z_{(t,q-1),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q-1),h})} \right]^{1/H_{(t,q-1)}} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q),c} \bar{z}_{(t,q-1),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q-1),c} \bar{z}_{(t,q-1),c})} = P^{L}_{(t,q-1),(t,q)};$$

$$\text{GP}: \quad \frac{P_{(t,q)}}{P_{(t,q-1)}} = \left[ \prod_{h=1}^{H_{(t,q)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q),h})} \right]^{1/H_{(t,q)}} = \frac{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q),c} \bar{z}_{(t,q),c})}{\exp(\sum_{c=1}^{C} \hat{\beta}_{(t,q-1),c} \bar{z}_{(t,q),c})} = P^{P}_{(t,q-1),(t,q)};$$

$$\text{Törnqvist}: \quad \left\{ \left[ \prod_{h=1}^{H_{(t,q-1)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q-1),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q-1),h})} \right]^{1/H_{(t,q-1)}} \left[ \prod_{h=1}^{H_{(t,q)}} \frac{\hat{p}_{(t,q),h}(z_{(t,q),h})}{\hat{p}_{(t,q-1),h}(z_{(t,q),h})} \right]^{1/H_{(t,q)}} \right\}^{1/2}$$

$$= \left\{ \frac{\exp[\sum_{c=1}^{C} \hat{\beta}_{(t,q),c}(\bar{z}_{(t,q-1),c} + \bar{z}_{(t,q),c})]}{\exp[\sum_{c=1}^{C} \hat{\beta}_{(t,q-1),c}(\bar{z}_{(t,q-1),c} + \bar{z}_{(t,q),c})]} \right\}^{1/2} = \left( P^{L}_{(t,q-1),(t,q)} \times P^{P}_{(t,q-1),(t,q)} \right)^{1/2} = P^{F}_{(t,q-1),(t,q)},$$

where $P^F_{j,k}$ denotes a Fisher price index comparison between periods $j$ and $k$.

Relative to the first quarter in the data set (1,1), the price index for period $(t+1,1)$ is calculated as follows:

$$\frac{P_{(t+1,1)}}{P_{(1,1)}} = P^F_{(1,1),(1,2)} \times P^F_{(1,2),(1,3)} \times \cdots \times P^F_{(t+1,1),(t,4)}.$$

In summary, the UK method is essentially equivalent to the average characteristic method used by Romania and Spain. While its method can also be represented as an average characteristics method, Germany is the only country in the EU that uses the Törnqvist formula to construct its HPI.

## 2.4 Rolling time dummy (RTD) method

The Rolling Time Dummy (RTD) method was proposed by Shimizu, Takatsuji, Ono and Nishimura (2010) (see also O'Hanlon 2011). Even though it has only been developed recently, it is already being used by a number of NSIs in Europe.

RTD is a variant on the widely used time-dummy hedonic method. The relationship between time-dummy and hedonic imputations methods has been explored by Diewert, Heravi, and Silver (2009) and de Haan (2010).

When discussing the RTD method we will use here a slightly different notation than we have used thus far in this paper. We refer simply to periods denoted by $s$ and $t$, without distinguishing which year and quarter they are in.

The RTD method begins by estimating the following hedonic model over a time window of $k+1$ periods starting with period $s$:

$$\ln p_{uh} = \sum_{c=1}^{C} \beta_{(s,s+k),c} z_{uhc} + \sum_{i=s+1}^{s+k} \delta_i D_{ih} + \varepsilon_{uh}, \tag{5}$$

where $h$ now indices the dwelling transactions in periods $s, \ldots, s+k$, and $D_{ih}$ is a dummy variable that equals 1 when $u = i$ is the period in which the dwelling sold, and zero otherwise. Now the characteristic shadow prices $\beta_{(s,s+k),c}$ are common to periods $s, \ldots, s+k$. The RTD method then moves the window forward one period, and re-estimates the model.

The RTD method derives the price index comparing period $t+k-1$ to period

$t + k$ as follows:

$$\frac{P_{t+k}}{P_{t+k-1}} = \frac{\exp(\hat{\delta}^t_{t+k})}{\exp(\hat{\delta}^t_{t+k-1})}. \tag{6}$$

A superscript $t$ is included on the estimated $\delta$ coefficients to indicate that they obtained from the hedonic model with period $t$ as the base. The hedonic model with period $t$ as the base is only used to compute the change in house prices from period $t + k - 1$ to period $t + k$. The window is then rolled forward one period and the hedonic model is re-estimated. The change in house prices from period $t + k$ to period $t + k + 1$ is now computed as follows:

$$\frac{P_{t+k+1}}{P_{t+k}} = \frac{\exp(\hat{\delta}^{t+1}_{t+k+1})}{\exp(\hat{\delta}^{t+1}_{t+k})}, \tag{7}$$

where now the base period in the hedonic model is period $t + 1$. The price index over multiple periods is computed by chaining these bilateral comparisons together as follows:

$$\frac{P_{t+k+1}}{P_t} = \left(\frac{\exp(\hat{\delta}^{t-k}_{t+1})}{\exp(\hat{\delta}^{t-k}_{t})}\right)\left(\frac{\exp(\hat{\delta}^{t-k+1}_{t+2})}{\exp(\hat{\delta}^{t-k+1}_{t+1})}\right)\cdots\left(\frac{\exp(\hat{\delta}^{t+1}_{t+k+1})}{\exp(\hat{\delta}^{t+1}_{t+k})}\right). \tag{8}$$

NSIs in Europe using the RTD method differ in the length of their rolling windows: France=2,Cyprus=4,Ireland=5, Portugal=2, Croatia=4.

A trade-off exists when choosing the window length. A shorter window makes the price index more current, but less robust. When data points are scarce, RTD4Q and RTD5Q (i.e., 4 or 5 quarter windows) are recommended over RTD2Q (i.e., a 2 quarter window).

An important feature of the RTD method is that once a price change $P_{t+k}t + k - 1$ has been computed, it is never revised. Hence when data for a new period $t + k + 1$ becomes available, the price indices $P_t$, $P_{t+1}$, $\ldots$, $P_{t+k}$ are already fixed. The sole objective when estimating the hedonic model inclusive of data from period $t + k + 1$ is to compute $P_{t+k+1}$, irrespective of how many periods are included in the hedonic model.

More generally, this property of never being revised is shared by all the hedonic price indices considered here. It is an important requirement of the HPI.

## 2.5 Mix adjusted methods

Mix-adjusted methods are also sometimes referred to as stratified median methods. The first step in computing a mix-adjusted median index is to split the data set into strata. A first split would be between houses and apartments. Each stratum can be further subdivided based on location, for example by postcode. When information on the physical characteristics of dwelling are available, splits can also be done based say on size (for example floor area less than 80 square meters and greater than 80 square meters). In the empirical application, after splitting houses and apartments, we then focus only on locational stratification based on postcodes or their equivalent.

Once the strata have been constructed, the median price for each stratum is computed. These medians are then averaged separately for houses and apartments, typically using the arithmetic mean formula. One issue that arises is whether the arithmetic mean formula should be weighted by the number of transactions in each stratum.

A mix-adjusted method lies somewhere in between a simple median method and a quality-adjusted hedonic method. Averaging medians across strata reduces the noise in the index resulting from compositional changes in the median dwelling over time. While in principle more strata should imply better quality adjustment, this approach soon runs into the problem when the classification becomes finer that some of the strata may be empty in some periods (i.e., there are no transactions with that particular mix of characteristics). This imposes limits on how far mix-adjusted methods can take the quality-adjustment process.

# 3 Combining Price Indices Across Strata

In our empirical analysis, irrespective of which method is used, we estimate separate house price indices for houses and apartments in Sydney. How then should these price indices be aggregated to obtain an overall price index for Sydney? In the Eurostat terminology, houses and apartments are referred to as strata. Eurostat recommends strata should be aggregated annually using a chained Laspeyres-type formula.

In our context, the aggregation formula comparing quarter 1 in year $t$ with quarter $q$ in that same year, therefore, takes the following form:

$$\frac{P_{(t,q)}}{P_{(t,1)}} = \sum_{n=1}^{N} w_{t-1}^{n} \left( \frac{p_{t,q}^{n}}{p_{t,1}^{n}} \right). \tag{9}$$

The term $w_{t-1}^{n}$ is the expenditure share of strata $n$ in year $t-1$ (the most recent year for which complete data are available). It is calculated as follows (where $Exp_t^n$ here denotes total expenditure on strata $n$ in year $t$):

$$w_t^n = \frac{Exp_t^n}{\sum_{n=1}^{N} Exp_t^n}$$

It should be noted that while the price indices are computed at a quarterly frequency, the expenditure weights are only updated once a year, and are computed using a whole year of data. This helps to increase the stability of the index.

The formula in (9) does not deal with the first quarter each year. The price index for the first quarter of year $t$ is computed by comparing it with the first quarter of year $t-1$ using year $t-1$'s weights:

$$\frac{P_{(t,1)}}{P_{(t-1,1)}} = \sum_{n=1}^{N} w_{t-1}^n \left( \frac{p_{t,1}^n}{p_{t-1,1}^n} \right).$$

Given this structure, the overall price index comparing year 1-quarter 1 with year $t$-quarter $q$ is computed as follows:

$$\frac{P_{(t,q)}}{P_{(1,1)}} = \prod_{\tau=1}^{t} \left[ \sum_{n=1}^{N} w_{\tau-1}^n \left( \frac{p_{\tau,1}^n}{p_{\tau-1,1}^n} \right) \right] \sum_{n=1}^{N} w_{t-1}^n \left( \frac{p_{t,q}^n}{p_{t,1}^n} \right).$$

# 4 Results for Sydney (2003-2014)

## 4.1 The Sydney data set

We use a data set obtained from Australian Property Monitors that consists of prices and characteristics of houses and apartments sold in Sydney (Australia) for the years 2002–2014. Results are presented for the years 2003-2014. For some methods, data for 2002 are needed to compute the reference baskets used in 2003.

For each house we have the following characteristics: the actual sale price, time of sale, postcode, property type (i.e., detached or semi), number of bedrooms, number of bathrooms, land area, and postcode. There are 202 postcodes in the data set.

For apartments we have the same set of characteristics. However, we drop the land area characteristic for apartments in our hedonic analysis since it refers to the whole strata, and we do not have any information on the number of apartments in the building.

For a robust analysis it was necessary to remove some outliers. This is because there is a concentration of data entry errors in the tails, caused for example by the inclusion of erroneous extra zeroes. These extreme observations can distort the results. Complete data on all our hedonic characteristics are available for 380 414 house transactions. For apartments the corresponding figure is 250 005.

## 4.2 Summary of methods to be considered

The methods that will be compared * of which the first nine are hedonic) are listed below:

1. Repricing (no updating of base year)
2. Repricing (base year updated every five years)
3. Average characteristics
4. Double imputation Geometric-Laspeyres
5. Double imputation Geometric-Paasche
6. Double imputation Törnqvist
7. RTD (2 quarters)
8. RTD (4 quarters)
9. RTD (5 quarters)
10. Mix adjusted

In the case of Sydney, price indices will be computed for houses, apartments.. For Tokyo, almost all the data pertains to apartments. Hence hedonic price indices will only be computed for apartments. Age of dwellings is available for Tokyo but not for Sydney. So we will also compute price indices for new apartments in Tokyo. This focus on new dwellings is important in the context of the experimental owner-occupied housing price index (OOHI) being computed by countries in Europe for possible inclusion at a later date in the harmonized index of consumer prices (HICP).

## 4.3 House and apartment price indices for Sydney

The house price indices for Sydney generated by the various methods discussed above are shown in Table 1. Four of the series are graphed in Figure 1. It is clear from

Table 1 and Figure 1 that the house price index is quite robust to the choice of method. Over the whole sample period, depending on the choice of hedonic method, house prices rose by between 73.7 and 78.1 percent. The two repricing methods – RP1 (which uses shadow prices from 2003) and RP2 (which updates the shadow prices every five years) – generate the lowest increase in house prices. This may be attributable to a small amount of substitution bias, in which buyers substitute slightly over time towards postcodes that have become relatively more expensive. Interestingly, rebasing the shadow prices every five years does not seem to help much in reducing this drift in this case.

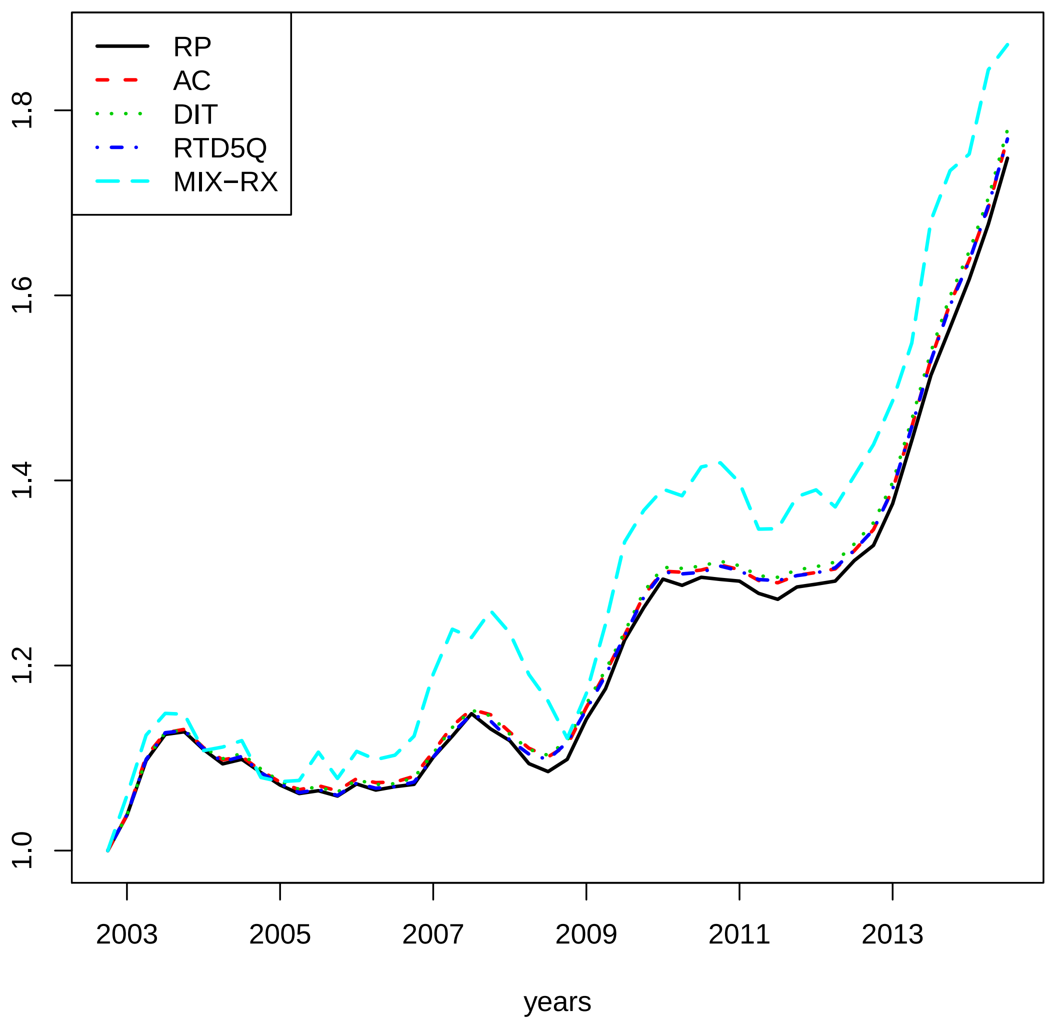
Also shown in Table 1 are mix-adjusted results computed in two different ways. MIX-PC stratifies houses by postcodes of which there are 202. MIX-RX stratifies by Residex region of which there are 16.[1] The MIX-PC stratification is hence much finer than its MIX-RX counterpart. It is not surprising therefore that the MIX-PC index is less erratic and closer to the hedonic indices. The MIX-PC index rises by 82 percent while MIX-RX rises by 87 percent. The concern is not just that it rises faster than the hedonic indices, but also that it is more volatile.

The results for apartments in Sydney shown in Figure 1 are also reasonably robust to the choice of method, when we restrict the comparison to the hedonic methods actually used by NSIs to compute the HPI. The measured cumulative rise in apartment prices for the hedonic methods ranges between 68.1 and 72.6 percent. The mix-adjusted MIX-RX index, by contrast, rises by 80 percent.

However, the double imputation Paasche (DIP) and Laspeyres (DIL) indices – shown in Table 2 but excluded from Figure 1 – exhibit clear evidence of drift. According to DIP, prices rise by only 65.3 percent while according to DIL prices rise by 78.1 percent. It is fortunate therefore that none of the NSIs are using either DIP or DIL. The German NSI though is using double imputation Törnqvist (DIT), which is the geometric mean

[1]The Residex regions (with their constituent postcodes listed in brackets) are as follows: Inner Sydney (2000 to 2020), Eastern Suburbs (2021 to 2036), Inner West (2037 to 2059), Lower North Shore (2060 to 2069), Upper North Shore (2070 to 2087), Mosman-Cremorne (2088 to 2091), Manly-Warringah (2092 to 2109), North Western (2110 to 2126), Western Suburbs (2127 to 2145), Parramatta Hills (2146 to 2159), Fairfield-Liverpool (2160 to 2189), Canterbury-Bankstown (2190 to 2200), St George (2201 to 2223), Cronulla-Sutherland (2224 to 2249), Campbelltown (2552 to 2570), Penrith-Windsor (2740 to 2777).

Figure 1: Price Indices for Houses in Sydney (2003Q1=1)




Note: RP = Repricing; AC = Average characteristics; DIT = Double imputation Törnqvist; RTD5Q = Rolling time dummy with five quarter window; MIX-RX = Mix adjusted stratified by Residex region.

Table 1: Price Indices for Houses in Sydney (2003Q1=1)

| | RP1 | RP2 | AC | DIL | DIP | DIT | RTD2Q | RTD4Q | RTD5Q | MIX-PC | MIX-RX |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2003Q1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 2003Q2 | 1.038 | 1.038 | 1.039 | 1.038 | 1.038 | 1.038 | 1.038 | 1.039 | 1.039 | 1.033 | 1.059 |
| 2003Q3 | 1.097 | 1.097 | 1.102 | 1.098 | 1.099 | 1.099 | 1.099 | 1.098 | 1.098 | 1.111 | 1.125 |
| 2003Q4 | 1.126 | 1.126 | 1.127 | 1.128 | 1.127 | 1.127 | 1.127 | 1.127 | 1.127 | 1.129 | 1.148 |
| 2004Q1 | 1.128 | 1.128 | 1.131 | 1.131 | 1.129 | 1.130 | 1.131 | 1.129 | 1.130 | 1.146 | 1.147 |
| 2004Q2 | 1.109 | 1.109 | 1.111 | 1.113 | 1.110 | 1.111 | 1.112 | 1.110 | 1.110 | 1.091 | 1.108 |
| 2004Q3 | 1.094 | 1.094 | 1.098 | 1.100 | 1.099 | 1.099 | 1.099 | 1.096 | 1.096 | 1.083 | 1.112 |
| 2004Q4 | 1.099 | 1.099 | 1.102 | 1.106 | 1.104 | 1.105 | 1.105 | 1.102 | 1.102 | 1.097 | 1.119 |
| 2005Q1 | 1.084 | 1.084 | 1.086 | 1.090 | 1.087 | 1.088 | 1.088 | 1.085 | 1.084 | 1.074 | 1.079 |
| 2005Q2 | 1.071 | 1.071 | 1.074 | 1.076 | 1.074 | 1.075 | 1.075 | 1.072 | 1.072 | 1.044 | 1.074 |
| 2005Q3 | 1.062 | 1.062 | 1.066 | 1.068 | 1.065 | 1.066 | 1.066 | 1.063 | 1.063 | 1.047 | 1.076 |
| 2005Q4 | 1.065 | 1.065 | 1.070 | 1.070 | 1.067 | 1.069 | 1.069 | 1.067 | 1.066 | 1.072 | 1.106 |
| 2006Q1 | 1.059 | 1.059 | 1.065 | 1.065 | 1.062 | 1.063 | 1.064 | 1.060 | 1.060 | 1.057 | 1.078 |
| 2006Q2 | 1.072 | 1.072 | 1.078 | 1.077 | 1.075 | 1.076 | 1.077 | 1.074 | 1.073 | 1.078 | 1.107 |
| 2006Q3 | 1.065 | 1.065 | 1.073 | 1.073 | 1.070 | 1.071 | 1.072 | 1.068 | 1.067 | 1.092 | 1.098 |
| 2006Q4 | 1.069 | 1.069 | 1.074 | 1.073 | 1.071 | 1.072 | 1.072 | 1.070 | 1.069 | 1.094 | 1.103 |
| 2007Q1 | 1.072 | 1.072 | 1.080 | 1.079 | 1.078 | 1.079 | 1.079 | 1.075 | 1.074 | 1.109 | 1.124 |
| 2007Q2 | 1.101 | 1.101 | 1.107 | 1.105 | 1.105 | 1.105 | 1.105 | 1.102 | 1.101 | 1.139 | 1.191 |
| 2007Q3 | 1.124 | 1.124 | 1.135 | 1.133 | 1.133 | 1.133 | 1.133 | 1.129 | 1.127 | 1.204 | 1.239 |
| 2007Q4 | 1.148 | 1.148 | 1.153 | 1.152 | 1.152 | 1.152 | 1.152 | 1.148 | 1.147 | 1.205 | 1.230 |
| 2008Q1 | 1.131 | 1.131 | 1.147 | 1.144 | 1.147 | 1.146 | 1.146 | 1.142 | 1.140 | 1.211 | 1.259 |
| 2008Q2 | 1.119 | 1.112 | 1.128 | 1.124 | 1.127 | 1.126 | 1.126 | 1.121 | 1.120 | 1.203 | 1.235 |
| 2008Q3 | 1.094 | 1.096 | 1.111 | 1.108 | 1.111 | 1.109 | 1.109 | 1.106 | 1.104 | 1.160 | 1.190 |
| 2008Q4 | 1.085 | 1.090 | 1.101 | 1.101 | 1.105 | 1.103 | 1.103 | 1.100 | 1.098 | 1.137 | 1.162 |
| 2009Q1 | 1.099 | 1.110 | 1.114 | 1.114 | 1.124 | 1.119 | 1.119 | 1.118 | 1.117 | 1.137 | 1.121 |
| 2009Q2 | 1.142 | 1.142 | 1.155 | 1.153 | 1.164 | 1.158 | 1.159 | 1.155 | 1.152 | 1.180 | 1.170 |
| 2009Q3 | 1.175 | 1.179 | 1.192 | 1.190 | 1.200 | 1.195 | 1.196 | 1.191 | 1.189 | 1.241 | 1.245 |
| 2009Q4 | 1.228 | 1.217 | 1.233 | 1.230 | 1.242 | 1.236 | 1.236 | 1.235 | 1.232 | 1.302 | 1.334 |
| 2010Q1 | 1.263 | 1.258 | 1.276 | 1.274 | 1.285 | 1.280 | 1.280 | 1.277 | 1.274 | 1.366 | 1.368 |
| 2010Q2 | 1.293 | 1.287 | 1.302 | 1.301 | 1.311 | 1.306 | 1.307 | 1.304 | 1.301 | 1.399 | 1.391 |
| 2010Q3 | 1.287 | 1.285 | 1.301 | 1.300 | 1.310 | 1.305 | 1.306 | 1.302 | 1.299 | 1.373 | 1.383 |
| 2010Q4 | 1.295 | 1.285 | 1.303 | 1.303 | 1.311 | 1.307 | 1.308 | 1.304 | 1.301 | 1.365 | 1.415 |
| 2011Q1 | 1.293 | 1.290 | 1.309 | 1.309 | 1.317 | 1.313 | 1.313 | 1.310 | 1.307 | 1.389 | 1.419 |
| 2011Q2 | 1.291 | 1.287 | 1.304 | 1.305 | 1.311 | 1.308 | 1.309 | 1.305 | 1.302 | 1.359 | 1.398 |
| 2011Q3 | 1.278 | 1.280 | 1.292 | 1.293 | 1.301 | 1.297 | 1.298 | 1.296 | 1.293 | 1.321 | 1.347 |
| 2011Q4 | 1.272 | 1.279 | 1.289 | 1.291 | 1.300 | 1.295 | 1.296 | 1.294 | 1.292 | 1.339 | 1.348 |
| 2012Q1 | 1.285 | 1.277 | 1.297 | 1.300 | 1.307 | 1.303 | 1.304 | 1.300 | 1.297 | 1.343 | 1.382 |
| 2012Q2 | 1.288 | 1.279 | 1.301 | 1.303 | 1.310 | 1.307 | 1.307 | 1.303 | 1.300 | 1.347 | 1.390 |
| 2012Q3 | 1.291 | 1.289 | 1.304 | 1.309 | 1.315 | 1.312 | 1.313 | 1.308 | 1.306 | 1.342 | 1.371 |
| 2012Q4 | 1.313 | 1.301 | 1.324 | 1.327 | 1.335 | 1.331 | 1.332 | 1.328 | 1.325 | 1.377 | 1.405 |
| 2013Q1 | 1.330 | 1.324 | 1.347 | 1.352 | 1.357 | 1.354 | 1.355 | 1.350 | 1.347 | 1.408 | 1.439 |
| 2013Q2 | 1.375 | 1.366 | 1.390 | 1.395 | 1.400 | 1.398 | 1.398 | 1.394 | 1.390 | 1.453 | 1.486 |
| 2013Q3 | 1.443 | 1.433 | 1.458 | 1.463 | 1.469 | 1.466 | 1.466 | 1.462 | 1.458 | 1.530 | 1.549 |
| 2013Q4 | 1.514 | 1.502 | 1.531 | 1.536 | 1.542 | 1.539 | 1.540 | 1.534 | 1.530 | 1.613 | 1.681 |
| 2014Q1 | 1.565 | 1.562 | 1.591 | 1.597 | 1.602 | 1.599 | 1.599 | 1.593 | 1.589 | 1.676 | 1.735 |
| 2014Q2 | 1.617 | 1.609 | 1.638 | 1.646 | 1.649 | 1.647 | 1.648 | 1.642 | 1.637 | 1.704 | 1.753 |
| 2014Q3 | 1.677 | 1.668 | 1.695 | 1.702 | 1.708 | 1.705 | 1.706 | 1.702 | 1.697 | 1.752 | 1.844 |
| 2014Q4 | 1.748 | 1.737 | 1.770 | 1.777 | 1.781 | 1.779 | 1.780 | 1.774 | 1.769 | 1.819 | 1.871 |

Note: The hedonic methods are as follows: RP1 = Repricing; RP2 = Repricing where the base period is updated every five years; AC = Average characteristics; Double imputation Laspeyres = DIL; Double imputation Paasche = DIP; Double imputation Törnqvist = DIT; RTD2Q = Rolling time dummy with a 2 quarter rolling window; RTD4Q and RTD5Q have 4 and 5 quarter rolling windows; MIX-PC = Mix adjusted stratified by postcode; MIX-RX = Mix adjusted stratified by Residex region.

of DIP and DIL. The results indicate that the drifts in DIP and DIL are offsetting, and hence DIT is unaffected by any drift problems.

Given the duality between average characteristic and hedonic imputation methods, we should also consider the implications of this finding for the former. The average characteristics method, which uses a Laspeyres type formula, is potentially also at risk of drift. However, the drift arises here when the average dwelling is updated each quarter based on the previous quarter's data. The average characteristics method used by NSIs only updates the average dwelling annually and computes it based on a whole year's data. This seems to be enough to prevent drift.

It is also noticeable that RP2 (repricing where the base period is updated every five years) generates very similar results to the average characteristics (AC) method. Given the duality that exists between the repricing and average characteristics methods this result is not so surprising. However, such similarity in the RP2 and AC results was not observed for houses in Table 1.

The results for the mix-adjusted methods are similar to those obtained for houses. The mix-adjusted indices rise more than the hedonic indices, and are more volatile. Again, as expected, of the two mix-adjusted indices, MIX-PC is less erratic and approximates better the hedonic indices.

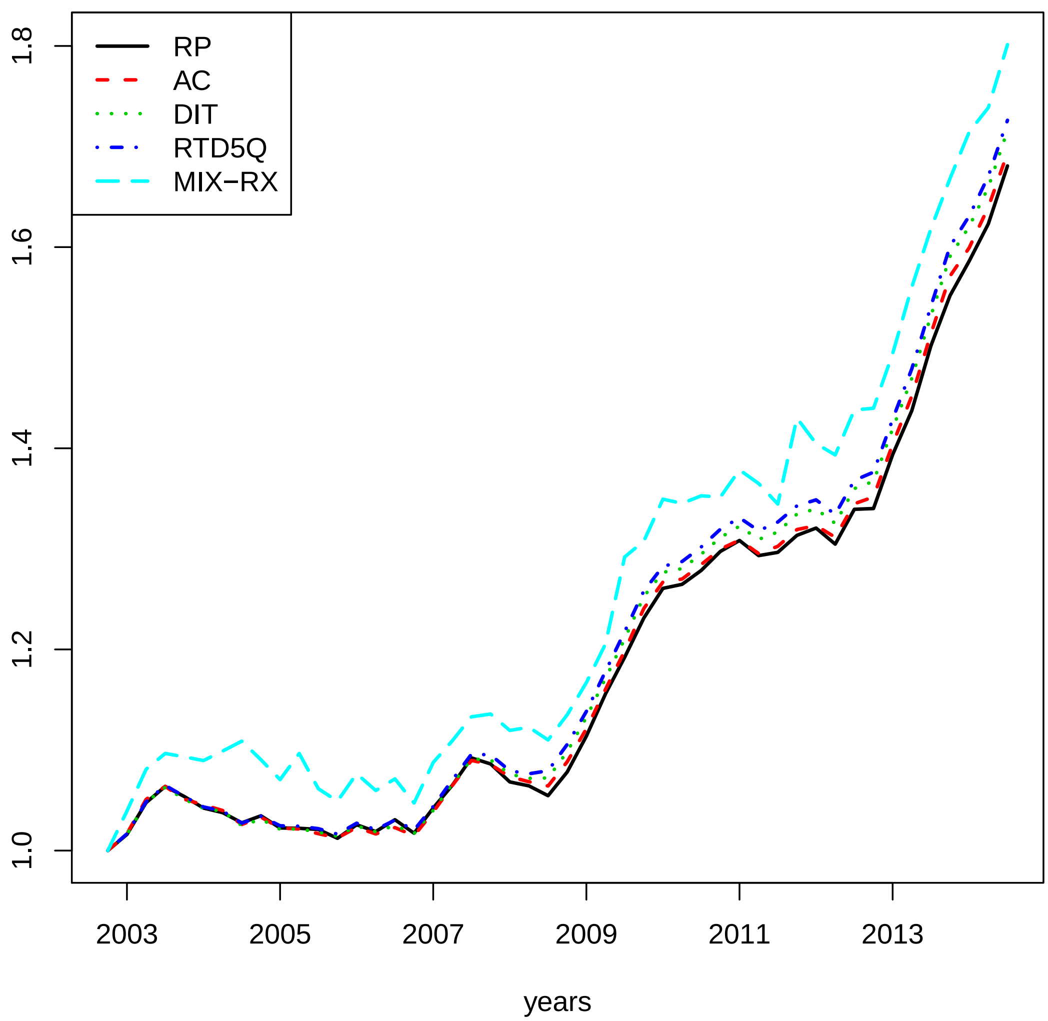
Overall, the results in Tables 1 and 2 should be reassuring to Eurostat. They indicate that the HPIs of different countries should be broadly comparable even when computed using different hedonic methods. For those countries using mix adjusted methods, it is important that the strata are sufficiently finely defined. Otherwise, like MIX-RX, the index will behave erratically.

# 5 Results for Tokyo (1986-2016)

## 5.1 The Tokyo data set

The Tokyo data set consists of 23 wards of the Tokyo metropolitan area (621 square kilometers), and the analysis period is approximately 30 years between January 1986 and June 2016. The data set covers previously-owned condominiums (apartments) published in Residential Information Weekly (or Shukan Jyutaku Joho in Japanese) published by RECRUIT, Co. This magazine provides information on the characteristics and asking prices of listed properties on a weekly basis. Moreover, Shukan Jutaku Joho

Figure 2: Price Indices for Apartments in Sydney (2003Q1=1)




Note: Note: RP = Repricing; AC = Average characteristics; DIT = Double imputation Törnqvist; RTD5Q = Rolling time dummy with five quarter window; MIX-RX = Mix adjusted stratified by Residex region.

Table 2: Price Indices for Apartments in Sydney (2003Q1=1)

| | RP1 | RP2 | AC | HIL | HIP | HIT | RTD2Q | RTD4Q | RTD5Q | MIX-PC | MIX-RX |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2003Q1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 2003Q2 | 1.016 | 1.016 | 1.017 | 1.018 | 1.015 | 1.016 | 1.016 | 1.016 | 1.017 | 1.019 | 1.039 |
| 2003Q3 | 1.048 | 1.048 | 1.051 | 1.046 | 1.050 | 1.048 | 1.048 | 1.049 | 1.049 | 1.050 | 1.081 |
| 2003Q4 | 1.064 | 1.064 | 1.064 | 1.060 | 1.066 | 1.063 | 1.063 | 1.065 | 1.065 | 1.078 | 1.097 |
| 2004Q1 | 1.054 | 1.054 | 1.051 | 1.047 | 1.056 | 1.051 | 1.051 | 1.054 | 1.054 | 1.061 | 1.093 |
| 2004Q2 | 1.042 | 1.042 | 1.045 | 1.037 | 1.047 | 1.042 | 1.042 | 1.044 | 1.043 | 1.075 | 1.089 |
| 2004Q3 | 1.038 | 1.038 | 1.040 | 1.032 | 1.046 | 1.039 | 1.039 | 1.040 | 1.039 | 1.072 | 1.099 |
| 2004Q4 | 1.028 | 1.028 | 1.026 | 1.017 | 1.033 | 1.025 | 1.026 | 1.028 | 1.027 | 1.065 | 1.109 |
| 2005Q1 | 1.035 | 1.035 | 1.033 | 1.023 | 1.040 | 1.031 | 1.031 | 1.034 | 1.034 | 1.067 | 1.090 |
| 2005Q2 | 1.022 | 1.022 | 1.022 | 1.013 | 1.029 | 1.021 | 1.021 | 1.025 | 1.025 | 1.041 | 1.071 |
| 2005Q3 | 1.022 | 1.022 | 1.021 | 1.013 | 1.030 | 1.022 | 1.022 | 1.025 | 1.024 | 1.054 | 1.097 |
| 2005Q4 | 1.021 | 1.021 | 1.017 | 1.009 | 1.027 | 1.018 | 1.018 | 1.022 | 1.022 | 1.040 | 1.061 |
| 2006Q1 | 1.012 | 1.012 | 1.013 | 1.004 | 1.024 | 1.014 | 1.014 | 1.018 | 1.016 | 1.064 | 1.049 |
| 2006Q2 | 1.026 | 1.026 | 1.023 | 1.015 | 1.034 | 1.025 | 1.024 | 1.028 | 1.027 | 1.049 | 1.076 |
| 2006Q3 | 1.019 | 1.019 | 1.017 | 1.008 | 1.029 | 1.019 | 1.019 | 1.022 | 1.021 | 1.034 | 1.060 |
| 2006Q4 | 1.031 | 1.031 | 1.023 | 1.012 | 1.038 | 1.025 | 1.025 | 1.031 | 1.030 | 1.054 | 1.071 |
| 2007Q1 | 1.017 | 1.017 | 1.015 | 1.002 | 1.032 | 1.017 | 1.017 | 1.022 | 1.020 | 1.026 | 1.047 |
| 2007Q2 | 1.042 | 1.042 | 1.038 | 1.025 | 1.055 | 1.040 | 1.040 | 1.045 | 1.044 | 1.063 | 1.088 |
| 2007Q3 | 1.065 | 1.065 | 1.065 | 1.050 | 1.083 | 1.066 | 1.066 | 1.071 | 1.070 | 1.098 | 1.109 |
| 2007Q4 | 1.092 | 1.092 | 1.090 | 1.073 | 1.109 | 1.091 | 1.091 | 1.097 | 1.096 | 1.132 | 1.133 |
| 2008Q1 | 1.086 | 1.086 | 1.086 | 1.070 | 1.110 | 1.090 | 1.090 | 1.096 | 1.095 | 1.109 | 1.136 |
| 2008Q2 | 1.068 | 1.070 | 1.073 | 1.054 | 1.098 | 1.076 | 1.076 | 1.080 | 1.079 | 1.084 | 1.119 |
| 2008Q3 | 1.064 | 1.066 | 1.068 | 1.050 | 1.094 | 1.072 | 1.072 | 1.076 | 1.076 | 1.072 | 1.123 |
| 2008Q4 | 1.055 | 1.068 | 1.064 | 1.048 | 1.097 | 1.072 | 1.072 | 1.079 | 1.079 | 1.088 | 1.110 |
| 2009Q1 | 1.078 | 1.095 | 1.088 | 1.073 | 1.125 | 1.098 | 1.098 | 1.105 | 1.106 | 1.106 | 1.135 |
| 2009Q2 | 1.114 | 1.126 | 1.121 | 1.105 | 1.161 | 1.132 | 1.132 | 1.138 | 1.138 | 1.133 | 1.167 |
| 2009Q3 | 1.156 | 1.166 | 1.160 | 1.142 | 1.202 | 1.171 | 1.170 | 1.178 | 1.178 | 1.170 | 1.205 |
| 2009Q4 | 1.192 | 1.198 | 1.199 | 1.177 | 1.245 | 1.210 | 1.210 | 1.218 | 1.218 | 1.239 | 1.292 |
| 2010Q1 | 1.231 | 1.242 | 1.241 | 1.218 | 1.289 | 1.253 | 1.252 | 1.259 | 1.258 | 1.271 | 1.308 |
| 2010Q2 | 1.261 | 1.265 | 1.267 | 1.241 | 1.313 | 1.277 | 1.275 | 1.283 | 1.283 | 1.296 | 1.349 |
| 2010Q3 | 1.265 | 1.269 | 1.270 | 1.244 | 1.318 | 1.280 | 1.279 | 1.287 | 1.287 | 1.296 | 1.345 |
| 2010Q4 | 1.279 | 1.282 | 1.284 | 1.259 | 1.332 | 1.295 | 1.294 | 1.302 | 1.302 | 1.311 | 1.353 |
| 2011Q1 | 1.297 | 1.300 | 1.300 | 1.274 | 1.349 | 1.311 | 1.310 | 1.319 | 1.319 | 1.324 | 1.351 |
| 2011Q2 | 1.308 | 1.315 | 1.309 | 1.283 | 1.363 | 1.323 | 1.321 | 1.330 | 1.331 | 1.316 | 1.379 |
| 2011Q3 | 1.293 | 1.300 | 1.295 | 1.268 | 1.352 | 1.309 | 1.308 | 1.317 | 1.318 | 1.320 | 1.365 |
| 2011Q4 | 1.297 | 1.313 | 1.303 | 1.274 | 1.361 | 1.317 | 1.315 | 1.325 | 1.327 | 1.307 | 1.345 |
| 2012Q1 | 1.313 | 1.316 | 1.319 | 1.289 | 1.382 | 1.335 | 1.333 | 1.342 | 1.343 | 1.358 | 1.430 |
| 2012Q2 | 1.321 | 1.325 | 1.323 | 1.293 | 1.388 | 1.340 | 1.339 | 1.348 | 1.349 | 1.361 | 1.405 |
| 2012Q3 | 1.305 | 1.314 | 1.311 | 1.277 | 1.375 | 1.325 | 1.324 | 1.333 | 1.334 | 1.346 | 1.393 |
| 2012Q4 | 1.339 | 1.341 | 1.345 | 1.311 | 1.411 | 1.360 | 1.359 | 1.367 | 1.368 | 1.388 | 1.438 |
| 2013Q1 | 1.340 | 1.353 | 1.352 | 1.317 | 1.420 | 1.367 | 1.366 | 1.375 | 1.376 | 1.412 | 1.440 |
| 2013Q2 | 1.394 | 1.404 | 1.404 | 1.367 | 1.474 | 1.419 | 1.418 | 1.427 | 1.429 | 1.462 | 1.494 |
| 2013Q3 | 1.437 | 1.453 | 1.452 | 1.415 | 1.526 | 1.469 | 1.469 | 1.478 | 1.479 | 1.500 | 1.560 |
| 2013Q4 | 1.502 | 1.513 | 1.514 | 1.476 | 1.589 | 1.532 | 1.531 | 1.540 | 1.541 | 1.584 | 1.618 |
| 2014Q1 | 1.552 | 1.573 | 1.571 | 1.533 | 1.651 | 1.591 | 1.590 | 1.600 | 1.601 | 1.627 | 1.668 |
| 2014Q2 | 1.586 | 1.603 | 1.599 | 1.561 | 1.682 | 1.621 | 1.620 | 1.630 | 1.631 | 1.669 | 1.714 |
| 2014Q3 | 1.623 | 1.641 | 1.640 | 1.598 | 1.723 | 1.660 | 1.659 | 1.670 | 1.671 | 1.675 | 1.739 |
| 2014Q4 | 1.681 | 1.693 | 1.694 | 1.653 | 1.781 | 1.716 | 1.715 | 1.726 | 1.726 | 1.767 | 1.801 |

Note: The hedonic methods are again as follows: RP1 = Repricing; RP2 = Repricing where the base period is updated every five years; AC = Average characteristics; Double imputation Laspeyres = DIL; Double imputation Paasche = DIP; Double imputation Törnqvist = DIT; RTD2Q = Rolling time dummy with a 2 quarter rolling window; RTD4Q and RTD5Q have 4 and 5 quarter rolling windows; MIX-PC = Mix adjusted stratified by postcode; MIX-RX = Mix adjusted stratified by Residex region.

provides time-series data on housing prices from the week they were first posted until the week they were removed as a result of successful transactions. We only use the price in the final week because this can be safely regarded as sufficiently close to the contract price.

The available housing characteristics are: floor area, age of building, travel time to nearest station, travel time to Tokyo central station, and the 23 wards (i.e., city codes). The hedonic model for Tokyo is estimated over 237 190 observations. Prior to that a few observations needed to be deleted since they were incomplete, or contained clear errors. The total number of deletions was less than 1 percent.

The functional form for our hedonic models is again semilog. The explanatory variables used are:

log of floor area

age (included as a quadratic)

time to nearest station

time to Tokyo central station (included as a quadratic)

ward dummy.

The hedonic methods considered are the same as in section 4.2. The reason for including quadratics for *age* and *time to Tokyo central station* is that the impact of these variables on log(price) may be nonlinear and even possibly nonmonotonic. For example, there may be an optimal time to Tokyo central station (i.e., one may not want to live too near and not too far way either). This quadratic specification can create problems with the repricing method, as is explained below.

## 5.2 Apartment price indices for Tokyo

The results for Tokyo for the years 1986 to 2016 for all apartments are shown in Table 6 and Figure 3. The general pattern that emerges is similar to that observed for Sydney, although there are some important differences.

Focusing first on the differences, two versions of the Repricing method without rebasing – RP1 and RP1 (lin) – are presented in Table 6. RP1 (lin) is much closer to the other methods than RP1. The difference between RP1 and RP1 (lin) is that the former uses the functional form discussed above that include *age* and *time to Tokyo* *central station* as quadratics. RP1 (lin) includes these variables as linear functions. The problem with RP1 is that while the quadratics fit the data well in 1987 – the first full year of the data set – this specification does not perform so well when it is applied to data in other years. The squared term in the quadratics can distort the results. The implication is that when the repricing method is used, quadratic terms in the hedonic model should be avoided. It is better to stick with a simpler more linear model. This problem was not observed for the Sydney data set because these variables were not included in the hedonic model.

A second difference is that there is no clear evidence of drift in the DIL and DIP results in Table 6, as compared with what was observed for Sydney apartments.

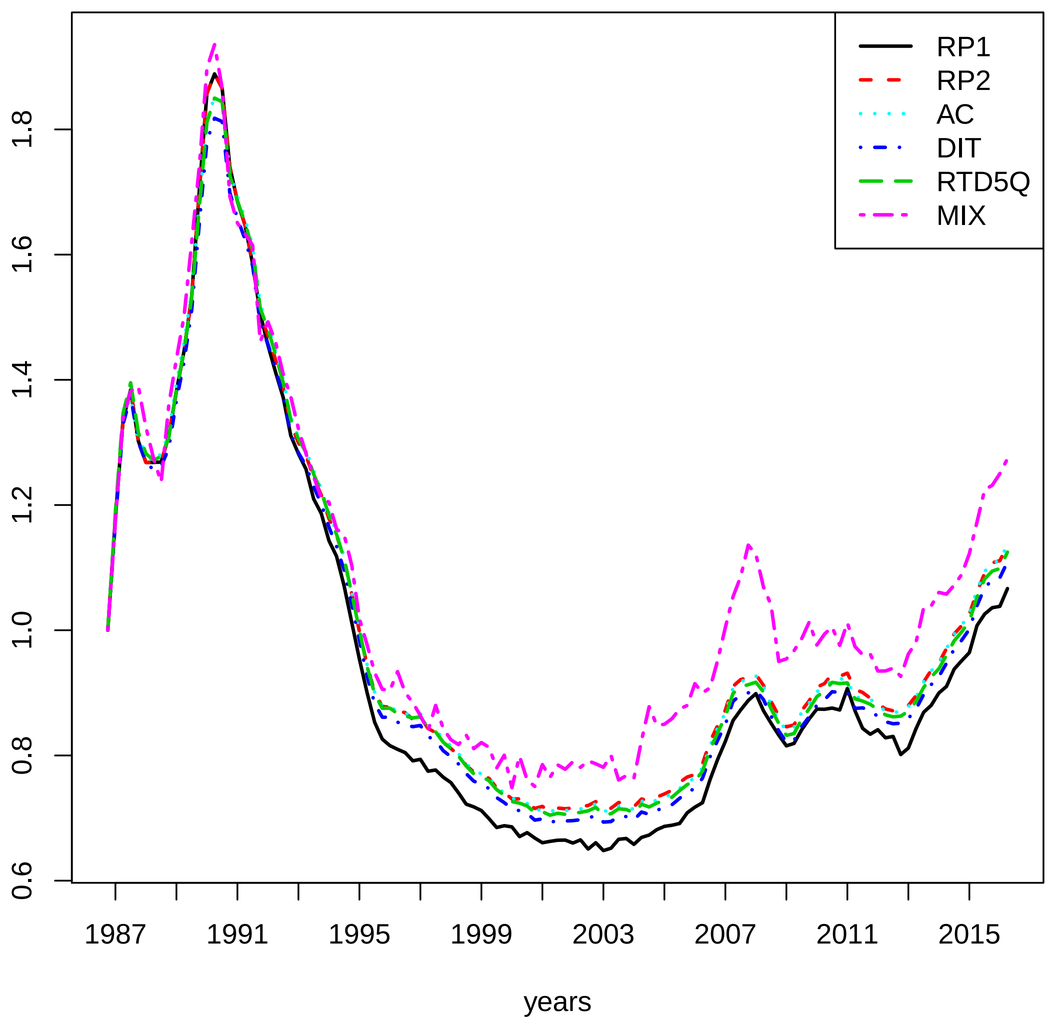
Over the whole sample period, the rise in house prices for all hedonic methods, excluding repricing without updating, ranges between 8.5 and 13.8 percent. The average masks a rollercoaster ride where prices first went way up and then way down before gradually returning to near their starting point.

Turning now to the similarities, the drift in the Repricing results is again downward, suggesting the same type of substitution towards wards that have become relatively more expensive. Having said that, when RP1 (lin) is used, the downward bias is not that large. According to RP1 (lin) prices rose by about 7 percent, as opposed to a 6 percent fall when RP1 is used.

It is noticeable that, as with Sydney apartments, RP2 (repricing where the base period is updated every five years) generates very similar results to the average characteristics (AC) method. Given the duality that exists between the repricing and average characteristics methods this result is not so surprising. However, such similarity in the RP2 and AC results was not observed for houses in Table 1.

The mix-adjusted results differ quite a bit from the hedonic differences. It rises by 27.4 percent, as compared with the hedonic range of 8.5 and 13.8 percent. It is worth noting that the mix-adjusted indices in all three Figures rise faster than their hedonic counterparts. One possible explanation for this finding is that the average quality of dwellings sold has increased over time.

Figure 3: Price Indices for Apartments in Tokyo (1986Q1=1)




Note: Note: RP1 = Repricing; RP2 = Repricing where the base period is updated every five years; AC = Average characteristics; DIT = Double imputation Törnqvist; RTD5Q = Rolling time dummy with five quarter window.

## 5.3 Price indices for new apartments in Tokyo

Estimating a price index for new apartments is problematic in the Tokyo data set due to the small sample size. We were forced to define a new build as any apartment that is less than three years old. We would have preferred less than two years old, but all the hedonic methods except RTD encountered problems in this case.

The new apartment price indices are shown in Figure 4. It can be seen that the index is much more sensitive to the choice of method than say in Figure 3. The mix-adjusted index is particularly badly affected by the small sample size. Faced with a small sample problem, we have greatest confidence in the RTD method, with a relatively long window (e.g., RTD5Q).

Figure 4 illustrates one of the problems with the acquisitions method for including OOH in the HICP. It is much harder to construct a reliable quality-adjusted house price index for new builds than it is to construct an index covering all housing transactions. Nevertheless, the acquisitions method as recommended by Eurostat requires, where possible, a price index specifically for new builds.

# 6 Conclusion

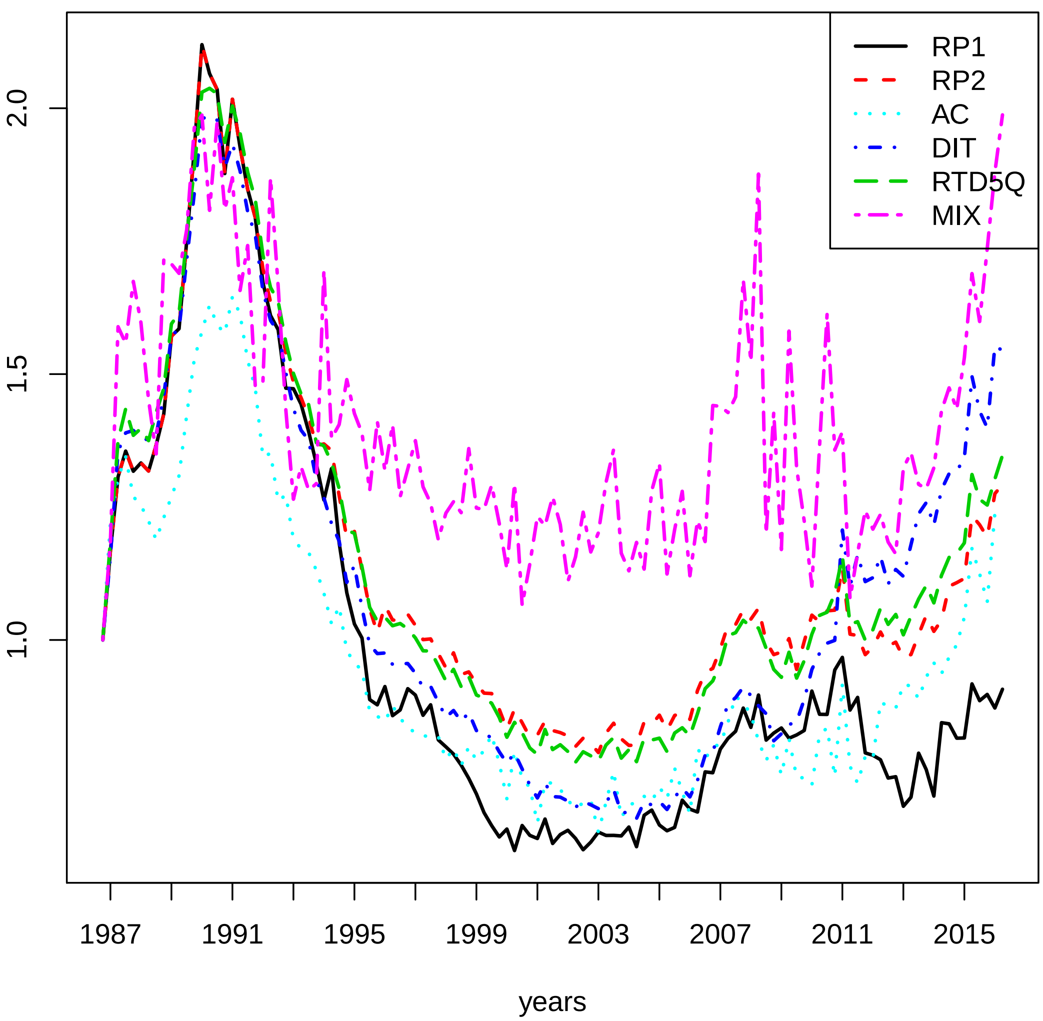
Our main findings are as follows:

- The price indices seem to be quite robust to the choice between the range of hedonic methods used by NSIs in Europe to compute their HPIs.

- The double imputation Paasche and Lasepyres (DIP and DIL) indices for apartments in Sydney are subject to drift. Evidence of a small amount of drift is also apparent in the Tokyo data. The results for Sydney apartments indicate that DIP and DIL should not be used. Fortunately, no NSIs are using either of these methods.

- The repricing method does nto perform that badly, even when the same reference shadow prices are used for many years. Nevertheless, in both our data sets it has a slight downward bias. It is recommended therefore that NSIs currently using repricing switch to one of the other methods (i.e., average characteristics, double

Figure 4: Price Indices for New Apartments in Tokyo (1986Q1=1)




Note: Note: RP1 = Repricing; RP2 = Repricing where the base period is updated every five years; AC = Average characteristics; DIT = Double imputation Törnqvist; RTD5Q = Rolling time dummy with five quarter window.

imputation Törnqvist-DIT, or RTD), or at the very least update the reference repricing shadow prices every five years.

- Where possible mix-adjusted indices should be avoided. In both our data sets, the mix-adjusted indices are somewhat erratic and rise faster than the hedonic indices. This suggests that average quality of transacted dwellings rose over time in both data sets.

- It is much more difficult (because there are less data) to construct a quality-adjusted price index for new builds than for all transacted dwellings. RTD4Q or RTD5Q are recommended when there is a shortage of data points.

Table 6: Price Indices for Apartments in Tokyo (1986Q4=1)

| | RP1 | RP1 (lin) | RP2 | AC | DIL | DIP | DIT | RTD2Q | RTD4Q | RTD5Q | MIX |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1986Q4 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 1987Q1 | 1.181 | 1.180 | 1.180 | 1.175 | 1.179 | 1.169 | 1.174 | 1.173 | 1.185 | 1.190 | 1.175 |
| 1987Q2 | 1.336 | 1.336 | 1.336 | 1.338 | 1.326 | 1.332 | 1.329 | 1.329 | 1.346 | 1.347 | 1.339 |
| 1987Q3 | 1.383 | 1.384 | 1.384 | 1.372 | 1.366 | 1.382 | 1.374 | 1.374 | 1.397 | 1.395 | 1.382 |
| 1987Q4 | 1.302 | 1.302 | 1.302 | 1.304 | 1.295 | 1.306 | 1.301 | 1.300 | 1.318 | 1.316 | 1.390 |
| 1988Q1 | 1.267 | 1.268 | 1.268 | 1.284 | 1.265 | 1.272 | 1.268 | 1.267 | 1.285 | 1.282 | 1.324 |
| 1988Q2 | 1.269 | 1.268 | 1.268 | 1.268 | 1.252 | 1.260 | 1.256 | 1.256 | 1.272 | 1.272 | 1.275 |
| 1988Q3 | 1.272 | 1.269 | 1.269 | 1.283 | 1.259 | 1.265 | 1.262 | 1.262 | 1.279 | 1.278 | 1.236 |
| 1988Q4 | 1.310 | 1.311 | 1.311 | 1.313 | 1.289 | 1.292 | 1.291 | 1.291 | 1.309 | 1.307 | 1.360 |
| 1989Q1 | 1.377 | 1.382 | 1.382 | 1.394 | 1.367 | 1.368 | 1.367 | 1.368 | 1.387 | 1.384 | 1.434 |
| 1989Q2 | 1.438 | 1.446 | 1.446 | 1.460 | 1.427 | 1.429 | 1.428 | 1.429 | 1.449 | 1.446 | 1.501 |
| 1989Q3 | 1.522 | 1.533 | 1.533 | 1.541 | 1.510 | 1.513 | 1.511 | 1.512 | 1.534 | 1.531 | 1.620 |
| 1989Q4 | 1.690 | 1.703 | 1.703 | 1.683 | 1.647 | 1.652 | 1.650 | 1.650 | 1.676 | 1.674 | 1.741 |
| 1990Q1 | 1.844 | 1.858 | 1.858 | 1.815 | 1.778 | 1.784 | 1.781 | 1.782 | 1.812 | 1.811 | 1.898 |
| 1990Q2 | 1.875 | 1.889 | 1.889 | 1.852 | 1.814 | 1.821 | 1.818 | 1.819 | 1.851 | 1.850 | 1.935 |
| 1990Q3 | 1.847 | 1.866 | 1.866 | 1.846 | 1.809 | 1.816 | 1.813 | 1.814 | 1.846 | 1.844 | 1.866 |
| 1990Q4 | 1.729 | 1.740 | 1.740 | 1.730 | 1.697 | 1.702 | 1.699 | 1.700 | 1.729 | 1.727 | 1.693 |
| 1991Q1 | 1.670 | 1.685 | 1.685 | 1.693 | 1.656 | 1.660 | 1.658 | 1.659 | 1.688 | 1.686 | 1.650 |
| 1991Q2 | 1.632 | 1.647 | 1.647 | 1.657 | 1.622 | 1.627 | 1.624 | 1.625 | 1.653 | 1.651 | 1.633 |
| 1991Q3 | 1.571 | 1.585 | 1.585 | 1.614 | 1.580 | 1.584 | 1.582 | 1.583 | 1.609 | 1.608 | 1.615 |
| 1991Q4 | 1.491 | 1.505 | 1.505 | 1.520 | 1.486 | 1.491 | 1.489 | 1.490 | 1.516 | 1.514 | 1.461 |
| 1992Q1 | 1.441 | 1.455 | 1.471 | 1.487 | 1.454 | 1.460 | 1.457 | 1.458 | 1.483 | 1.481 | 1.493 |
| 1992Q2 | 1.396 | 1.412 | 1.432 | 1.446 | 1.416 | 1.419 | 1.418 | 1.418 | 1.442 | 1.440 | 1.461 |
| 1992Q3 | 1.357 | 1.372 | 1.388 | 1.402 | 1.372 | 1.376 | 1.374 | 1.374 | 1.398 | 1.396 | 1.409 |
| 1992Q4 | 1.297 | 1.311 | 1.329 | 1.340 | 1.311 | 1.316 | 1.314 | 1.314 | 1.338 | 1.336 | 1.373 |
| 1993Q1 | 1.264 | 1.282 | 1.299 | 1.308 | 1.281 | 1.286 | 1.284 | 1.284 | 1.307 | 1.306 | 1.322 |
| 1993Q2 | 1.238 | 1.257 | 1.279 | 1.287 | 1.260 | 1.266 | 1.263 | 1.263 | 1.285 | 1.284 | 1.282 |
| 1993Q3 | 1.185 | 1.210 | 1.244 | 1.254 | 1.227 | 1.232 | 1.230 | 1.230 | 1.251 | 1.249 | 1.246 |
| 1993Q4 | 1.168 | 1.186 | 1.216 | 1.225 | 1.199 | 1.204 | 1.202 | 1.202 | 1.223 | 1.221 | 1.211 |
| 1994Q1 | 1.123 | 1.143 | 1.179 | 1.188 | 1.162 | 1.168 | 1.165 | 1.165 | 1.186 | 1.184 | 1.205 |
| 1994Q2 | 1.098 | 1.118 | 1.152 | 1.157 | 1.132 | 1.136 | 1.134 | 1.134 | 1.154 | 1.153 | 1.162 |
| 1994Q3 | 1.047 | 1.071 | 1.114 | 1.117 | 1.094 | 1.098 | 1.096 | 1.096 | 1.115 | 1.114 | 1.152 |
| 1994Q4 | 0.988 | 1.012 | 1.058 | 1.061 | 1.038 | 1.043 | 1.041 | 1.041 | 1.059 | 1.058 | 1.103 |
| 1995Q1 | 0.931 | 0.954 | 0.998 | 1.004 | 0.981 | 0.987 | 0.984 | 0.984 | 1.002 | 1.000 | 1.019 |
| 1995Q2 | 0.875 | 0.901 | 0.945 | 0.945 | 0.924 | 0.929 | 0.927 | 0.927 | 0.943 | 0.942 | 0.977 |
| 1995Q3 | 0.828 | 0.853 | 0.901 | 0.901 | 0.881 | 0.885 | 0.883 | 0.883 | 0.899 | 0.898 | 0.932 |
| 1995Q4 | 0.798 | 0.826 | 0.878 | 0.879 | 0.860 | 0.862 | 0.861 | 0.861 | 0.876 | 0.875 | 0.906 |
| 1996Q1 | 0.786 | 0.816 | 0.877 | 0.879 | 0.860 | 0.862 | 0.861 | 0.861 | 0.876 | 0.876 | 0.904 |
| 1996Q2 | 0.776 | 0.810 | 0.871 | 0.870 | 0.853 | 0.854 | 0.853 | 0.853 | 0.868 | 0.867 | 0.934 |
| 1996Q3 | 0.771 | 0.804 | 0.868 | 0.867 | 0.849 | 0.849 | 0.849 | 0.849 | 0.864 | 0.863 | 0.900 |
| 1996Q4 | 0.754 | 0.791 | 0.858 | 0.864 | 0.846 | 0.846 | 0.846 | 0.846 | 0.860 | 0.860 | 0.884 |
| 1997Q1 | 0.753 | 0.794 | 0.861 | 0.867 | 0.848 | 0.848 | 0.848 | 0.848 | 0.862 | 0.862 | 0.864 |
| 1997Q2 | 0.734 | 0.775 | 0.843 | 0.847 | 0.829 | 0.831 | 0.830 | 0.830 | 0.844 | 0.844 | 0.839 |
| 1997Q3 | 0.737 | 0.777 | 0.836 | 0.842 | 0.823 | 0.825 | 0.824 | 0.824 | 0.838 | 0.838 | 0.880 |
| 1997Q4 | 0.722 | 0.765 | 0.821 | 0.825 | 0.807 | 0.808 | 0.808 | 0.808 | 0.821 | 0.821 | 0.843 |
| 1998Q1 | 0.716 | 0.757 | 0.811 | 0.815 | 0.797 | 0.799 | 0.798 | 0.798 | 0.811 | 0.811 | 0.825 |
| 1998Q2 | 0.693 | 0.740 | 0.801 | 0.803 | 0.785 | 0.787 | 0.786 | 0.786 | 0.799 | 0.799 | 0.817 |
| 1998Q3 | 0.680 | 0.722 | 0.784 | 0.787 | 0.769 | 0.772 | 0.771 | 0.771 | 0.784 | 0.783 | 0.833 |
| 1998Q4 | 0.676 | 0.718 | 0.773 | 0.775 | 0.757 | 0.761 | 0.759 | 0.759 | 0.771 | 0.771 | 0.811 |
| 1999Q1 | 0.667 | 0.712 | 0.770 | 0.771 | 0.753 | 0.756 | 0.754 | 0.754 | 0.767 | 0.766 | 0.821 |
| 1999Q2 | 0.651 | 0.699 | 0.763 | 0.764 | 0.745 | 0.749 | 0.747 | 0.747 | 0.760 | 0.760 | 0.813 |
| 1999Q3 | 0.637 | 0.685 | 0.748 | 0.748 | 0.731 | 0.735 | 0.733 | 0.733 | 0.746 | 0.745 | 0.780 |
| 1999Q4 | 0.641 | 0.688 | 0.741 | 0.740 | 0.723 | 0.727 | 0.725 | 0.725 | 0.737 | 0.736 | 0.800 |
| 2000Q1 | 0.639 | 0.686 | 0.731 | 0.732 | 0.713 | 0.716 | 0.715 | 0.715 | 0.727 | 0.726 | 0.748 |
| 2000Q2 | 0.617 | 0.670 | 0.731 | 0.727 | 0.711 | 0.712 | 0.712 | 0.712 | 0.724 | 0.724 | 0.799 |
| 2000Q3 | 0.628 | 0.677 | 0.724 | 0.723 | 0.706 | 0.709 | 0.707 | 0.707 | 0.720 | 0.719 | 0.761 |
| 2000Q4 | 0.616 | 0.668 | 0.715 | 0.712 | 0.695 | 0.699 | 0.697 | 0.696 | 0.709 | 0.709 | 0.751 |
| 2001Q1 | 0.605 | 0.661 | 0.719 | 0.714 | 0.696 | 0.701 | 0.699 | 0.698 | 0.711 | 0.710 | 0.785 |
| 2001Q2 | 0.609 | 0.663 | 0.710 | 0.710 | 0.690 | 0.696 | 0.693 | 0.693 | 0.705 | 0.705 | 0.765 |
| 2001Q3 | 0.605 | 0.665 | 0.716 | 0.714 | 0.693 | 0.699 | 0.696 | 0.696 | 0.709 | 0.708 | 0.785 |

Table 6: Price Indices for Apartments in Tokyo (1986Q4=1) (continued)

| | RP1 | RP1 (lin) | RP2 | AC | DIL | DIP | DIT | RTD2Q | RTD4Q | RTD5Q | MIX |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2001Q4 | 0.607 | 0.665 | 0.715 | 0.712 | 0.692 | 0.698 | 0.695 | 0.695 | 0.707 | 0.706 | 0.778 |
| 2002Q1 | 0.599 | 0.660 | 0.716 | 0.713 | 0.693 | 0.699 | 0.696 | 0.696 | 0.709 | 0.708 | 0.790 |
| 2002Q2 | 0.603 | 0.665 | 0.718 | 0.715 | 0.694 | 0.700 | 0.697 | 0.697 | 0.710 | 0.709 | 0.781 |
| 2002Q3 | 0.588 | 0.651 | 0.720 | 0.718 | 0.696 | 0.703 | 0.699 | 0.699 | 0.713 | 0.712 | 0.791 |
| 2002Q4 | 0.594 | 0.661 | 0.726 | 0.723 | 0.701 | 0.708 | 0.705 | 0.704 | 0.718 | 0.717 | 0.787 |
| 2003Q1 | 0.583 | 0.648 | 0.715 | 0.711 | 0.689 | 0.698 | 0.694 | 0.693 | 0.706 | 0.706 | 0.781 |
| 2003Q2 | 0.591 | 0.652 | 0.716 | 0.712 | 0.689 | 0.700 | 0.694 | 0.694 | 0.708 | 0.707 | 0.800 |
| 2003Q3 | 0.594 | 0.666 | 0.725 | 0.719 | 0.697 | 0.709 | 0.703 | 0.703 | 0.716 | 0.715 | 0.761 |
| 2003Q4 | 0.600 | 0.667 | 0.723 | 0.720 | 0.698 | 0.707 | 0.702 | 0.702 | 0.715 | 0.714 | 0.768 |
| 2004Q1 | 0.587 | 0.658 | 0.718 | 0.713 | 0.693 | 0.701 | 0.697 | 0.697 | 0.709 | 0.708 | 0.763 |
| 2004Q2 | 0.608 | 0.669 | 0.730 | 0.727 | 0.706 | 0.714 | 0.710 | 0.710 | 0.723 | 0.722 | 0.824 |
| 2004Q3 | 0.608 | 0.673 | 0.728 | 0.722 | 0.701 | 0.711 | 0.706 | 0.706 | 0.719 | 0.718 | 0.878 |
| 2004Q4 | 0.617 | 0.681 | 0.734 | 0.729 | 0.707 | 0.717 | 0.712 | 0.712 | 0.725 | 0.724 | 0.848 |
| 2005Q1 | 0.622 | 0.687 | 0.738 | 0.734 | 0.712 | 0.721 | 0.717 | 0.717 | 0.729 | 0.728 | 0.850 |
| 2005Q2 | 0.624 | 0.689 | 0.744 | 0.737 | 0.717 | 0.726 | 0.721 | 0.721 | 0.734 | 0.733 | 0.859 |
| 2005Q3 | 0.623 | 0.691 | 0.757 | 0.747 | 0.726 | 0.737 | 0.732 | 0.731 | 0.745 | 0.744 | 0.874 |
| 2005Q4 | 0.638 | 0.709 | 0.766 | 0.757 | 0.734 | 0.748 | 0.741 | 0.741 | 0.754 | 0.753 | 0.880 |
| 2006Q1 | 0.646 | 0.717 | 0.769 | 0.764 | 0.741 | 0.754 | 0.747 | 0.747 | 0.760 | 0.759 | 0.915 |
| 2006Q2 | 0.645 | 0.724 | 0.787 | 0.780 | 0.757 | 0.771 | 0.764 | 0.764 | 0.776 | 0.775 | 0.900 |
| 2006Q3 | 0.683 | 0.762 | 0.822 | 0.816 | 0.791 | 0.805 | 0.798 | 0.798 | 0.811 | 0.810 | 0.908 |
| 2006Q4 | 0.708 | 0.794 | 0.848 | 0.844 | 0.818 | 0.831 | 0.825 | 0.824 | 0.838 | 0.837 | 0.952 |
| 2007Q1 | 0.744 | 0.823 | 0.873 | 0.868 | 0.842 | 0.857 | 0.850 | 0.849 | 0.864 | 0.862 | 1.005 |
| 2007Q2 | 0.776 | 0.856 | 0.910 | 0.908 | 0.878 | 0.895 | 0.887 | 0.886 | 0.901 | 0.899 | 1.053 |
| 2007Q3 | 0.791 | 0.873 | 0.921 | 0.918 | 0.887 | 0.906 | 0.897 | 0.896 | 0.911 | 0.910 | 1.085 |
| 2007Q4 | 0.810 | 0.888 | 0.925 | 0.922 | 0.891 | 0.909 | 0.900 | 0.899 | 0.914 | 0.913 | 1.136 |
| 2008Q1 | 0.826 | 0.899 | 0.929 | 0.927 | 0.895 | 0.913 | 0.904 | 0.903 | 0.918 | 0.917 | 1.121 |
| 2008Q2 | 0.797 | 0.871 | 0.912 | 0.910 | 0.881 | 0.898 | 0.889 | 0.889 | 0.903 | 0.901 | 1.069 |
| 2008Q3 | 0.765 | 0.851 | 0.886 | 0.880 | 0.852 | 0.874 | 0.863 | 0.862 | 0.875 | 0.874 | 1.040 |
| 2008Q4 | 0.749 | 0.832 | 0.863 | 0.856 | 0.827 | 0.852 | 0.839 | 0.839 | 0.852 | 0.851 | 0.950 |
| 2009Q1 | 0.734 | 0.815 | 0.846 | 0.840 | 0.806 | 0.837 | 0.821 | 0.821 | 0.833 | 0.832 | 0.954 |
| 2009Q2 | 0.738 | 0.819 | 0.849 | 0.845 | 0.808 | 0.840 | 0.824 | 0.824 | 0.835 | 0.835 | 0.967 |
| 2009Q3 | 0.754 | 0.841 | 0.872 | 0.869 | 0.830 | 0.862 | 0.846 | 0.846 | 0.858 | 0.858 | 0.986 |
| 2009Q4 | 0.778 | 0.858 | 0.887 | 0.884 | 0.847 | 0.878 | 0.862 | 0.862 | 0.875 | 0.875 | 1.013 |
| 2010Q1 | 0.782 | 0.874 | 0.909 | 0.902 | 0.865 | 0.896 | 0.880 | 0.881 | 0.894 | 0.894 | 0.976 |
| 2010Q2 | 0.778 | 0.874 | 0.915 | 0.909 | 0.874 | 0.904 | 0.889 | 0.889 | 0.903 | 0.903 | 0.994 |
| 2010Q3 | 0.767 | 0.876 | 0.930 | 0.922 | 0.885 | 0.919 | 0.902 | 0.902 | 0.917 | 0.917 | 1.005 |
| 2010Q4 | 0.774 | 0.873 | 0.927 | 0.923 | 0.883 | 0.918 | 0.901 | 0.901 | 0.915 | 0.915 | 0.976 |
| 2011Q1 | 0.809 | 0.907 | 0.931 | 0.921 | 0.884 | 0.918 | 0.901 | 0.902 | 0.916 | 0.916 | 1.012 |
| 2011Q2 | 0.771 | 0.871 | 0.906 | 0.892 | 0.856 | 0.895 | 0.875 | 0.878 | 0.891 | 0.890 | 0.974 |
| 2011Q3 | 0.732 | 0.843 | 0.901 | 0.893 | 0.853 | 0.899 | 0.876 | 0.876 | 0.888 | 0.887 | 0.961 |
| 2011Q4 | 0.733 | 0.834 | 0.891 | 0.890 | 0.850 | 0.894 | 0.871 | 0.871 | 0.883 | 0.882 | 0.962 |
| 2012Q1 | 0.739 | 0.841 | 0.882 | 0.881 | 0.842 | 0.883 | 0.862 | 0.862 | 0.874 | 0.873 | 0.935 |
| 2012Q2 | 0.720 | 0.828 | 0.874 | 0.871 | 0.834 | 0.874 | 0.854 | 0.854 | 0.865 | 0.865 | 0.935 |
| 2012Q3 | 0.723 | 0.830 | 0.871 | 0.868 | 0.830 | 0.872 | 0.851 | 0.850 | 0.862 | 0.862 | 0.939 |
| 2012Q4 | 0.693 | 0.801 | 0.872 | 0.870 | 0.831 | 0.873 | 0.851 | 0.851 | 0.863 | 0.863 | 0.926 |
| 2013Q1 | 0.708 | 0.812 | 0.880 | 0.878 | 0.838 | 0.880 | 0.859 | 0.859 | 0.871 | 0.871 | 0.962 |
| 2013Q2 | 0.735 | 0.843 | 0.895 | 0.893 | 0.852 | 0.897 | 0.874 | 0.874 | 0.887 | 0.886 | 0.981 |
| 2013Q3 | 0.758 | 0.869 | 0.917 | 0.918 | 0.875 | 0.920 | 0.897 | 0.897 | 0.910 | 0.909 | 1.035 |
| 2013Q4 | 0.777 | 0.880 | 0.936 | 0.935 | 0.892 | 0.935 | 0.913 | 0.913 | 0.927 | 0.926 | 1.039 |
| 2014Q1 | 0.791 | 0.900 | 0.948 | 0.949 | 0.905 | 0.948 | 0.926 | 0.926 | 0.940 | 0.938 | 1.060 |
| 2014Q2 | 0.795 | 0.910 | 0.971 | 0.970 | 0.926 | 0.969 | 0.947 | 0.947 | 0.961 | 0.960 | 1.058 |
| 2014Q3 | 0.815 | 0.938 | 0.994 | 0.993 | 0.949 | 0.991 | 0.970 | 0.970 | 0.984 | 0.983 | 1.072 |
| 2014Q4 | 0.840 | 0.951 | 1.007 | 1.008 | 0.962 | 1.007 | 0.984 | 0.984 | 0.999 | 0.998 | 1.089 |

Table 6: Price Indices for Apartments in Tokyo (1986Q4=1) (continued)

| | RP1 | RP1 (lin) | RP2 | AC | DIL | DIP | DIT | RTD2Q | RTD4Q | RTD5Q | MIX |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2015Q1 | 0.850 | 0.964 | 1.026 | 1.027 | 0.980 | 1.024 | 1.002 | 1.002 | 1.016 | 1.014 | 1.122 |
| 2015Q2 | 0.895 | 1.007 | 1.061 | 1.065 | 1.017 | 1.062 | 1.039 | 1.039 | 1.054 | 1.053 | 1.173 |
| 2015Q3 | 0.908 | 1.026 | 1.091 | 1.094 | 1.044 | 1.092 | 1.068 | 1.068 | 1.083 | 1.082 | 1.224 |
| 2015Q4 | 0.918 | 1.036 | 1.108 | 1.107 | 1.057 | 1.104 | 1.080 | 1.081 | 1.096 | 1.094 | 1.232 |
| 2016Q1 | 0.924 | 1.038 | 1.111 | 1.111 | 1.060 | 1.108 | 1.084 | 1.084 | 1.100 | 1.099 | 1.250 |
| 2016Q2 | 0.947 | 1.067 | 1.136 | 1.138 | 1.085 | 1.134 | 1.109 | 1.110 | 1.125 | 1.125 | 1.274 |

Note: The hedonic methods are again as follows: RP1 = Repricing; RP1 (lin) = Repricing using a simplified functional form where the quadratics in age and time to Tokyo central station are replaced with linear functions; RP2 = Repricing where the base period is updated every five years; AC = Average characteristics; Double imputation Laspeyres = DIL; Double imputation Paasche = DIP; Double imputation Törnqvist = DIT; RTD2Q = Rolling time dummy with a 2 quarter rolling window; RTD4Q and RTD5Q have 4 and 5 quarter rolling windows; MIX = Mix adjusted stratified by ward.

# References


de Haan, J., van der Wal, E., and P. de Vries (2008), "The Measurement of House Prices: A Review of the SPAR Method," Statistics Netherlands Working Paper.

de Haan J. (2010), "Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Re-Pricing Methods," *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*, 230(6), 772-791.

Diewert, W. E. (2011), Alternative Approaches to Measuring House Price Inflation, Economics Working Paper 2011-1, Vancouver School of Economics.

Diewert, W. E., S. Heravi, and M. Silver (2009), "Hedonic Imputation versus Time Dummy Hedonic Indexes," in W. E. Diewert, J. Greenlees, and C. Hulten (eds.), *Price Index Concepts and Measurement*, NBER Studies in Income and Wealth, University of Chicago Press, Chicago, 161–196.

European Commission, Eurostat, OECD, and World Bank (2013), *Handbook on Residential Property Price Indexes (RPPI)*, Publications Office of the European Union. Luxembourg: Eurostat.

Eurostat (2016), *Detailed Technical Manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices*, Eurostat, Luxembourg.

Hill, R. J. (2013), "Hedonic Price Indexes for Housing: A Survey, Evaluation and Taxonomy," *Journal of Economic Surveys*, 27(5), 879-914.

Hill, R. J. and D. Melser (2008), "Hedonic Imputation and the Price Index Problem: An Application to Housing," *Economic Inquiry*, 46(4), 593-609.

O'Hanlon, N. (2011), "Constructing a National House Price Index for Ireland," *Journal of the Statistical and Social Inquiry Society of Ireland* 40, 167-196.

Shimizu, C, H.Takatsuji, H.Ono and K. G. Nishimura (2010), "Structural and Temporal Changes in the Housing Market and Hedonic Housing Price Indices, *International Journal of Housing Markets and Analysis* 3(4), 351-368.

Silver, M. (2011), "House Price Indices: Does Measurement Matter?" *World Economics*, 12(3), 69-86.