The Measurement of House Prices: A Review of the Sale Price Appraisal Ratio Method

Jan de Haan,[a] **Erna van der Wal**[b] **and Paul de Vries**[c] [d]

14 August 2008

Abstract: The so-called sale price appraisal ratio (SPAR) method has been applied in a number of countries to construct house price indexes. This paper reviews the statistical and index number properties of the SPAR approach. Three types of SPAR indexes are distinguished: a weighted index, which aims at tracking the price change of the stock of owner-occupied houses, and two unweighted indexes, an arithmetic one and a geometric one. We also discuss stratified versions of these indexes and show how stratification can be used to estimate an expenditure-based Laspeyres-type price index. Empirical results for the Netherlands are given and compared with the repeat-sales index published by the Dutch land registry until January 2008.

Keywords: Government appraisal data, House price indexes.

JEL Classification: C43, R31.

[a] Corresponding author, address: Statistics Netherlands, Division of Macro-Economic Statistics and Dissemination, P.O. Box 24500, 2490 HA Den Haag, The Netherlands; email: jhhn@cbs.nl.

[b] Statistics Netherlands, Den Haag.

[c] OTB Research Institute for Housing, Urban and Mobility Studies, Delft University of Technology.

1. Introduction

The measurement of house prices is a challenging area for statistical agencies. They are often faced with severe data limitations and there are several difficult conceptual issues involved. Some of those measurement issues, for example adjusting for quality changes, are also encountered in the compilation of producer price indexes and the Consumer Price Index (CPI). But a number of other issues, while not entirely unique to housing, may be less familiar to index practitioners. One such issue is the distinction between a price index for the stock of houses – or more generally for a stock of assets – and a price index for houses that are acquired. The latter index can be separated into indexes for newly-built houses and second-hand houses, similar to what is being done in the CPI for durable goods like cars.

Statistics Netherlands participates in a Eurostat pilot study on the treatment of owner-occupied housing in the Harmonised Index of Consumer Prices (HICP). The aim is to compile price indexes for both newly-built and second-hand (re-sold) houses and ultimately incorporate them into the HICP using the 'net acquisitions approach'.[1] This approach means that second-hand houses exchanged between households are excluded. The number of second-hand dwellings purchased from outside the household sector is typically quite small in most countries, as will be the weight in the HICP. To meet user needs, Eurostat has decided to broaden the scope of the pilot study. Makaronidis and Hayes (2006) note that "While the net acquisitions approach is maintained, the intention is now to produce wider housing price indices, ones which will serve the needs of the HICP but also the needs of those users whose main interest is in dwelling price indices per se". In particular, a self-standing price index should now be compiled for all houses purchased, including second-hand houses exchanged between households.

In the present paper, however, we focus primarily on the compilation of an index that measures the price change of the total stock of owner-occupied houses. This has an advantage in its own right: a price index for the housing stock is presumably what the general public understands by a house price index. As most existing houses are not sold during a certain short time period, it is the market value of the property that should be

---

[1] The user-cost approach is an alternative approach for incorporating owner-occupied housing into a CPI. A third method, the rental equivalence approach, is applied in the Dutch national CPI. For a discussion of the different approaches, see Diewert (2003).

observed. By definition, the market value equals the sale price in case a house is sold. Market values must be estimated for houses that have not been sold. The market value, whether observed or estimated, is usually called the 'price'.

There are a number of methods for constructing house price indexes. Some use simple summary measures like the mean or median price. More sophisticated methods such as hedonic and repeat-sales approaches, employ regression analysis. Each method has its own advantages and weaknesses, which are documented extensively elsewhere; see e.g. Hansen (2006) for the Australian case. A major practical drawback of hedonic methods is that the characteristics of the sampled houses must be available to adjust for quality (mix) changes. In that respect the repeat-sales method has a distinct advantage;[2] only price changes of houses that were sold more than once enter the regression so that matched pairs of houses are compared over time.

Bourassa et al. (2006) present an alternative approach to measuring house price changes, referred to as the sale price appraisal ratio (SPAR) method. The SPAR method has been used in New Zealand, Sweden and Denmark for quite some time,[3] and is now also being used in the Netherlands. Like the repeat-sales approach, the SPAR method is based on matched pairs but in contrast uses (nearly) all price data that is available for the current period. Since the majority of the houses sold during the current period have not been sold during an earlier base period, base period sale prices are generally lacking. Those prices are therefore estimated using official valuations or appraisals. According to Bourassa et al. (2006) "the advantages and the relatively limited drawbacks of the SPAR method make it an ideal candidate for use by government agencies in developing house price indexes". Unfortunately the authors present the SPAR method – or rather an estimator of the period-to-period change in the index – without discussing it from an index number or statistical perspective. Several questions arise. Is the SPAR method aiming at a price index of the housing stock? Given the target, what assumptions are needed to derive the SPAR estimator, and what are its statistical properties? This paper addresses these and related questions.

---

[2] The repeat-sales technique was originally developed by Bailey et al. (1963) and has been further refined by Case and Shiller (1987), among others. See also Thibodeau (1997).

[3] The Swedish and Danish indexes are compiled by the national statistical agencies as part of their real estate statistics program. In New Zealand the SPAR index is produced and published by Quotable Value, a valuation and property information company.

In a broad sense two types of SPAR indexes can be distinguished, a weighted one and an equally weighted or unweighted one. The former is called a value-weighted index in the housing economics literature because the price indexes of the individual dwellings are implicitly weighted by their base period prices (values). In Section 3 we show that this index can, under a number of simplifying assumptions, be interpreted as an estimator of the price index (value index) for the stock of houses in a certain base period. In order to fully understand this index we will start by presenting in Section 2 a simple, though infeasible, estimator. Section 4 goes into the unweighted SPAR method. Here we discuss the choice between taking the arithmetic or the geometric mean of the individual price indexes. To gain further insight, Section 5 compares the unweighted geometric SPAR index with the hedonic time dummy variable index. Section 6 focuses on the lack of consistency in aggregation of the SPAR price indexes for sub-sets of the housing stock. Section 7 contains a discussion of the treatment of new and disappearing houses. In Section 8 we provide an alternative, expenditure-based interpretation of the value-weighted SPAR index and outline how a stratified value-weighted SPAR index can be constructed that most likely meets Eurostat requirements. Section 9 provides an overview of the data we used to compile weighted and unweighted SPAR price index numbers for the Netherlands. In Section 10 empirical results are presented, including a comparison with the repeat-sales price index that has been published by the Dutch land registry office until January 2008. Section 11 concludes.

2. A Simple Estimator of the Value Index of a Fixed Housing Stock

In general the choice of index number formula is an important issue. In price indexes like the CPI the price relatives must be weighted according to the economic importance of the goods (their value shares), and a symmetric treatment of the comparison period t and the base period 0 is preferable. For the construction of an index that aims at tracking the value of the stock of houses those issues are not equally important. Here, the target more or less automatically determines the index number formula to be used, the more so because every house is a unique asset, which implies that quantities do not play a role. This is particularly true if, as we assume for now, the housing stock is fixed over time. We will return to this issue in Section 7 where the assumption of a fixed stock of houses will be relaxed.

Let p_i^s denote the price (market value) of house i in period s ($s= 0,t$) and let U^0 be the total population of owner-occupied houses or housing stock, which is assumed fixed, in the base period with size N^0. Hence, the average price in period s ($s= 0,t$) is defined as $\bar{p}^s(U^0) = \sum_{i \in U^0} p_i^s / N^0$. The price index becomes

$$P_W^t = \frac{\sum_{i \in U^0} p_i^t}{\sum_{i \in U^0} p_i^0} = \frac{\sum_{i \in U^0} p_i^0\left(\frac{p_i^t}{p_i^0}\right)}{\sum_{i \in U^0} p_i^0} = \frac{\sum_{i \in U^0} p_i^t / N^0}{\sum_{i \in U^0} p_i^0 / N^0} = \frac{\bar{p}^t(U^0)}{\bar{p}^0(U^0)}. \tag{1}$$

The second expression on the right-hand side of equation (1) shows that P_W^t is a value-weighted index. In the standard index number literature price index formulas are usually named after their 'inventors', and P_W^t is called a Dutot index. One might alternatively refer to (1) as a Laspeyres price index in which all quantities are equal to 1 because each house is a unique property.

The question arises how to estimate P_W^t when not all data is available. Suppose that in period 0 and period t independent random samples have been drawn from the population U^0, say S^0 and S^t with size n^0 and n^t, for which we know the transaction or sale prices in both periods. Since (1) is a ratio of average prices (market values) in the population, the ratio of the sample averages $\bar{p}^s(S^s) = \sum_{i \in S^s} p_i^s / n^s$ ($s= 0, t$) is a natural estimator:

$$\widehat{P}_W^t = \frac{\sum_{i \in S^t} p_i^t / n^t}{\sum_{i \in S^0} p_i^0 / n^0} = \frac{\bar{p}^t(S^t)}{\bar{p}^0(S^0)}. \tag{2}$$

The use of (2) is problematic for three reasons. First, $\widehat{P}_W^t$ most likely has a high standard error. Second, $\widehat{P}_W^t$ is not unbiased; its expected value $E(\widehat{P}_W^t)$ differs from the target P_W^t. Third, there may be very few matched pairs and the problem of dealing with changes in the average quality of the sampled houses arises. Below we will discuss those problems in some detail.

In the Appendix approximating formulas for the bias and sampling variance of $\widehat{P}_W^t$ are derived based on Taylor series expansions. Still assuming that S^0 and S^t are independent random samples we obtain

$$E(\widehat{P}_W^t) - P_W^t \cong P_W^t \frac{\text{var}(\bar{p}^0(S^0))}{[\bar{p}^0(U^0)]^2}(> 0) \tag{3}$$

and

$$\text{var}(\widehat{P}_W^t) \cong (P_W^t)^2 \left[\frac{\text{var}(\bar{p}^t(S^t))}{[\bar{p}^t(U^0)]^2} + \frac{\text{var}(\bar{p}^0(S^0))}{[\bar{p}^0(U^0)]^2} \right]. \tag{4}$$

In practice the values of (3) and (4) may well turn out to be large. Nevertheless $\widehat{P}_W^t$ is actually used in some countries.[4]

As a consequence of independent sampling the set of houses belonging to both S^0 and S^t, i.e. the set $S^t \cap S^0$ of matched pairs, might be small or even empty, so that both p_i^0 and p_i^t will be known for a few houses only or for none at all. This situation of 'missing prices' is generally viewed as a problem of changes in the quality mix: if the set of sampled houses differs between two periods then the average quality will differ also. The conventional way of looking at this problem is that the price index should be adjusted for compositional change or changes in the quality mix, for example by using hedonic regression. Yet, above all things it is a sampling problem, or at least it can be interpreted as such: in the housing stock U^0 – where new and disappearing houses are excluded by assumption – changes in the quality mix do not occur. Put differently, the quality mix bias can be viewed as a part of the bias given by expression (3): the problem diminishes with an increasing sample size and vanishes if the whole population would be observed.

Suppose we knew the prices p_i^t in the comparison period t as well as the base period prices p_i^0 for all houses in the sample S^t. This assumption is totally unrealistic, but it helps to gain a better understanding of the SPAR method. Now we can estimate P_W^t using price observations from S^t only:

$$\breve{P}_W^t = \frac{\sum_{i \in S^t} p_i^t / n^t}{\sum_{i \in S^t} p_i^0 / n^t} = \frac{\bar{p}^t(S^t)}{\bar{p}^0(S^t)}, \tag{5}$$

where $\bar{p}^0(S^t) = \sum_{i \in S^t} p_i^0 / n^t$. Obviously, the problem of quality mix change no longer exists. Note that the set S^t constantly changes over time. Unlike most samples used for

---

[4] Before the introduction in May 2005 of the repeat-sales price index the Dutch land registry published this type of index for flats and single-family dwellings. The Australian Bureau of Statistics (ABS) applies the 'naive' estimator combined with detailed stratification to control for compositional change in its price index for established houses (ABS, 2005).

official price indexes, S^t is not a panel that is kept fixed as much as possible. As we do not know a priori which houses will be sampled, the base period prices for all $i \in U^0$ must be available. The bias and variance of $\breve{P}_W^t$ can now be approximated by (see again the Appendix)

$$E(\breve{P}_W^t) - P_W^t \cong P_W^t \left[\frac{\text{var}(\overline{p}^0(S^t))}{[\overline{p}^0(U^0)]^2} - \frac{\text{cov}(\overline{p}^t(S^t), \overline{p}^0(S^t))}{\overline{p}^t(U^0)\overline{p}^0(U^0)} \right] \tag{6}$$

and

$$\text{var}(\breve{P}_W^t) \cong (P_W^t)^2 \left[\frac{\text{var}(\overline{p}^t(S^t))}{[\overline{p}^t(U^0)]^2} + \frac{\text{var}(\overline{p}^0(S^t))}{[\overline{p}^0(U^0)]^2} - 2\frac{\text{cov}(\overline{p}^t(S^t), \overline{p}^0(S^t))}{\overline{p}^t(U^0)\overline{p}^0(U^0)} \right]. \tag{7}$$

It should be clear that the potentially large bias and variance are reduced considerably in this case. In each period $\breve{P}_W^t$ is based on a single sample (S^t) and since p_i^0 and p_i^t will generally be highly positively correlated, the covariance term $\text{cov}(\overline{p}^t(S^t), \overline{p}^0(S^t))$ most likely has a large positive value. The bias associated with ratio estimators such as (5) is called small sample bias. It decreases as the sample is enlarged (that is, the estimator is consistent) and can be neglected in practice, provided that the sample size is sufficiently large.

The assumption of known base period sale prices is one of the biggest practical problems of $\breve{P}_W^t$. In reality those prices will be observed for a limited number of houses belonging to S^t. A solution might be to estimate the 'missing' base period prices using a model-based approach that incorporates auxiliary data. The SPAR method is based on this idea.

3. The Value-Weighted SPAR Method

Replacing the true but largely unobserved base period prices (market values) p_i^0 in the denominator of (5) by model-based estimates $\hat{p}_i^0$ yields an alternative estimator:

$$\hat{P}_W^t = \frac{\sum_{i \in S^t} p_i^t / n^t}{\sum_{i \in S^t} \hat{p}_i^0 / n^t} = \frac{\overline{p}^t(S^t)}{\overline{\hat{p}}^0(S^t)}, \tag{8}$$

in which $\bar{\hat{p}}^0(S^t) = \sum_{i \in S^t} \hat{p}_i^0 / n^t$. While for some sampled houses the base period sale prices may be available, for reasons of consistency and practicality it seems preferable to use the estimated prices for those houses as well.

A requirement that must be met for every price index is that it takes on the value of 1 if all prices remain unchanged between the base period and the comparison period. According to the axiomatic approach to index number theory the index then satisfies the identity test (ILO et al., 2004). This test should not only hold for the target index P_W^t but for estimators thereof as well. Notice that estimator (5) satisfies the identity test. To ensure that the model-based estimator (8) also satisfies this test, in other words to attach a value of 1 to $\hat{P}_W^t$ in the base period ($\hat{P}_W^0 = 1$), we impose the following restriction:

$$\frac{\sum_{i \in S^0} p_i^0}{\sum_{i \in S^0} \hat{p}_i^0} = 1. \tag{9}$$

Suppose next that the base period prices are estimated using a regression model. For simplicity we restrict ourselves here to the linear model $p_i^0 = \alpha^0 + \sum_k \beta_k^0 x_{ik} + \varepsilon_i^0$ (for all $i \in U^0$), where x_{ik} denotes the k-th explanatory variable ($k = 1,\ldots,K$) and β_k^0 the corresponding parameter; ε_i^0 is a disturbance term with $E(\varepsilon_i^0) = 0$. This model can be estimated by Ordinary Least Squares (OLS) regression on the data of the base period sample S^0, yielding estimated parameters (coefficients) $\hat{\alpha}^0(S^0)$ and $\hat{\beta}_k^0(S^0)$. In case of OLS, restriction (9) will be satisfied since the residuals $\hat{p}_i^0 - p_i^0$ sum to zero using a model that includes an intercept term (α^0). The predicted base period prices for $i \in S^t$ are $\hat{p}_i^0 = \hat{\alpha}^0(S^0) + \sum_k \hat{\beta}_k^0(S^0) x_{ik}$. Thus, the average estimated base period price for the period t sample is $\bar{\hat{p}}^0(S^t) = \hat{\alpha}^0(S^0) + \sum_k \hat{\beta}_k^0(S^0) \bar{x}_k(S^t)$, where $\bar{x}_k(S^t) = \sum_{i \in S^t} x_{ik} / n^t$ denotes the average value of x_{ik} in sample S^t. Substituting this result into (8) gives

$$\hat{P}_{W(A)}^t = \frac{\bar{p}^t(S^t)}{\hat{\alpha}^0(S^0) + \sum_k \hat{\beta}_k^0(S^0) \bar{x}_k(S^t)}. \tag{10}$$

Estimator $\hat{P}_{W(A)}^t$ has some drawbacks too. We have assumed that the regression model holds for all $i \in U^0$. However, the OLS estimators are unbiased, conditional on the data pertaining to the base-period sample S^0 from which the model is estimated. Predicting (base period) prices for houses belonging to the current-period sample S^t may give rise to biased results. Another problem has to do with the choice of regression

model. If we wish to have a model that explains house prices as accurate as possible, attributes of the houses would be needed. Since we explicitly want to avoid hedonics, other auxiliary variables must be found. Government appraisals might be a satisfactory alternative because we expect them to have substantial power to predict market prices. Suppose that appraisals a_i^0 are available for all houses belonging to U^0. Then we could postulate the following descriptive model: $p_i^0 = \gamma + \mu a_i^0 + \varepsilon_i^0$. Estimating this model on the data of S^0 using OLS produces predicted base period prices $\hat{p}_i^0 = \hat{\gamma}(S^0) + \hat{\mu}(S^0) a_i^0$. Substituting those into expression (8) for all $i \in S^t$ and denoting the average appraisal value in S^t by $\bar{a}^0(S^t) = \sum_{i \in S^t} a_i^0 / n^t$ gives

$$\hat{P}_{W(B)}^t = \frac{\bar{p}^t(S^t)}{\hat{\gamma}(S^0) + \hat{\mu}(S^0)\bar{a}^0(S^t)}.$$

(11)

An objection one might raise against $\hat{P}_{W(B)}^t$ is that the regression line should pass through the origin because if $a_i^0 \to 0$ we also expect $p_i^0 \to 0$.[5] The model should thus be re-formulated as $p_i^0 = \beta a_i^0 + \varepsilon_i^0$ for $i \in U^0$. Estimating this simplified model on the data of S^0 produces a regression coefficient $\hat{\beta}(S^0)$. Substituting $\hat{p}_i^0 = \hat{\beta}(S^0) a_i^0$ for $i \in S^t$ into (8) yields

$$\hat{P}_{W(C)}^t = \frac{\bar{p}^t(S^t)}{\hat{\beta}(S^0)\bar{a}^0(S^t)}.$$

(12)

Of course the problem mentioned earlier still holds: the OLS estimator is unbiased for $i \in S^0$, and predicting the base period prices for $i \in S^t$ may give rise to biased results. Moreover, because the simplified model does not contain an intercept term, restriction (9) is not met. In order to obtain a price index that equals unity in the base period, we simply divide $\hat{P}_{W(C)}^t$ by its starting value $\hat{P}_{W(C)}^0$. This normalisation leads to an estimator in which regression coefficient $\hat{\beta}(S^0)$ no longer plays a role:

$$\hat{P}_{W(S)}^t = \frac{\bar{p}^t(S^t)/\bar{a}^0(S^t)}{\bar{p}^0(S^0)/\bar{a}^0(S^0)} = \left[\frac{\bar{a}^0(S^0)}{\bar{p}^0(S^0)}\right] \frac{\sum_{i \in S^t} a_i^0 \left(\frac{p_i^t}{a_i^0}\right)}{\sum_{i \in S^t} a_i^0}.$$

(13)

---

[5] One could also argue that there is a lower bound to the market value of a dwelling so that this restriction is undesirable. In that case (11) would be a feasible, though not necessarily very accurate, estimator that satisfies the identity test.

$\hat{P}^t_{W(S)}$ given by (13) is what Bourassa et al. (2006) call the *value-weighted SPAR index*. We have used regression modelling to stress that the SPAR method is essentially a model-based approach. Alternatively, $\hat{P}^t_{W(S)}$ follows directly from $\breve{P}^t_W$, given by (5), if we assume an identical, non-stochastic ratio b between the market value of a house and its appraisal in the base period ($p^0_i / a^0_i = b$) and subsequently normalised the result. Of course $\hat{P}^t_{W(S)}$ reduces to $\breve{P}^t_W$ when this assumption would exactly hold. There is a simple relationship between $\hat{P}^t_{W(S)}$ and the 'naïve' estimator $\widehat{P}^t_W$, given by expression (2):

$$\hat{P}^t_{W(S)} = \left[\frac{\overline{a}^0(S^0)}{\overline{a}^0(S^t)} \right] \widehat{P}^t_W . \tag{14}$$

Equation (14) makes clear that the value-weighted SPAR index measures a larger price change than the ratio of sample mean prices ($\hat{P}^t_{W(S)} > \widehat{P}^t_W$) if the current period sample contains relatively many cheap houses as compared to the base period sample.

Bourassa et al. (2006) present the SPAR index as a pseudo chain index. Recast in our notation they define $\hat{P}^t_{W(S)}$ by

$$\hat{P}^t_{W(S)} = \prod_{\tau=1}^{t} \frac{\hat{P}^\tau_{W(S)}}{\hat{P}^{\tau-1}_{W(S)}} = \prod_{\tau=1}^{t} \frac{\overline{p}^\tau(S^\tau) / \overline{a}^0(S^\tau)}{\overline{p}^{\tau-1}(S^{\tau-1}) / \overline{a}^0(S^{\tau-1})} . \tag{15}$$

The presentation of a direct index, which directly relates period *t* to the base period 0, as a chain index may divert one's attention from what the SPAR method is actually about. But it does show that the measured period-to-period price change

$$\frac{\hat{P}^\tau_{W(S)}}{\hat{P}^{\tau-1}_{W(S)}} = \frac{\overline{p}^\tau(S^\tau)/\overline{a}^0(S^\tau)}{\overline{p}^{\tau-1}(S^{\tau-1})/\overline{a}^0(S^{\tau-1})} = \left[\frac{\overline{a}^0(S^{\tau-1})}{\overline{a}^0(S^\tau)} \right] \frac{\overline{p}^\tau(S^\tau)}{\overline{p}^{\tau-1}(S^{\tau-1})} = \left[\frac{\overline{a}^0(S^{\tau-1})}{\overline{a}^0(S^\tau)} \right] \breve{P}^{\tau-1,\tau}_W , \tag{16}$$

with $\breve{P}^{\tau-1,\tau}_W = \overline{p}^\tau(S^\tau) / \overline{p}^{\tau-1}(S^{\tau-1})$, has a similar structure as expression (14). Thus, if the current period sample contains relatively many cheap houses compared with the sample in the preceding period, then the price change measured by the value-weighted SPAR method will be larger than the price change measured by the ratio of sample means. Yet there is a subtle difference in the interpretation of (14) and (16). The direct SPAR index given by (14) or (13) is essentially based on matched pairs of houses, which is the basic principle underlying the SPAR method, and thus adjusts for compositional change. The period-to-period links, on the other hand, cannot be interpreted in this way and are thus affected by compositional change.

The value-weighted SPAR index $\hat{P}_{W(S)}^t$ is a ratio of two ratios and the question arises whether it is an approximately unbiased estimator of the population index P_W^t. In the Appendix an approximating expression for the expected value of $\hat{P}_{W(S)}^t$ is derived. Again assuming that S^0 and S^t are independently drawn random samples from U^0, the following expression for the bias of $\hat{P}_{W(S)}^t$ holds:

$$E(\hat{P}_{W(S)}^t) - P_W^t \cong 2P_W^t [cv(\bar{p}^0(S^0))]^2 (1-\rho) , \tag{17}$$

where $\rho = \text{cov}(\bar{p}^0(S^0), \bar{a}^0(S^0)) / [\text{var}\,\bar{p}^0(S^0)\,\text{var}\,\bar{a}^0(S^0)]^{1/2}$ (with $-1 \le \rho \le 1$) denotes the correlation coefficient of the average price and the average appraisal in period 0 and $cv(\bar{p}^0(S^0)) = [\text{var}(\bar{p}^0(S^0))]^{1/2} / \bar{p}(U^0)$ the coefficient of variation of $\bar{p}^0(S^0)$. Provided that the appraisals are reasonably accurate reflections of the sale prices (market values), ρ will be close to 1.[6] Note that this type of bias is of a technical nature. It is implicitly assumed that the relation between appraisals and the true market values in the base 0 is the same for houses sold in period t as that for houses sold in period 0, which cannot be tested as the base period market values for the sample S^t are unknown. The magnitude of the bias of $\hat{P}_{W(S)}^t$ depends on the value of $cv(\bar{p}^0(S^0))$ and hence on the sample size. As an illustration we give a numerical example. Taking $P_W^t = 1.01$ (a price increase of 1%), $\rho = 0.95$ and $cv(\bar{p}^0(S^0)) = 0.05$ the bias amounts to 0.00025, i.e. only 0.025%, according to (17). If the coefficient of variation would be twice as high (other things equal), the bias becomes 0.1%, which is not negligible at an appreciation rate of 1%.

De Haan (2007a) explains how the variance of $\hat{P}_{W(S)}^t$ can be approximated using Taylor linearization. Since we are dealing with a ratio of two ratios, the formula for the variance is more complex than the formula presented in Section 2 and will not be shown here. Not surprisingly, the standard error depends on the size of the samples S^t and S^0. In practice the 'samples' contain all sales as registered by the Dutch land registry. An important feature is that the standard error of the index may vary considerably over time as a result of the changing sample size. In periods with few sales the standard error of the SPAR index could be substantial, especially for a small country like the Netherlands with less than four million owner-occupied houses, and even more so for SPAR indexes for sub-populations.

---

[6] This correlation coefficient appears to be equal to the root of the coefficient of determination R-squared obtained from a univariate regression, including an intercept term, that 'explains' the base period sale price using the appraisal as the explanatory variable.

4. Unweighted SPAR Indexes

This section addresses the estimation of equally weighted or unweighted averages of the individual price (market value) indexes p_i^t / p_i^0 using the SPAR approach. At first sight the unweighted arithmetic average

$$P_{UA}^t = \frac{\sum_{i \in U^0} \left(\frac{p_i^t}{p_i^0} \right)}{N^0} \tag{18}$$

seems an obvious target index, but the unweighted geometric average

$$P_{UG}^t = \prod_{i \in U^0} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{1}{N^0}} \tag{19}$$

could be used also. In the index number literature (18) and (19) are usually referred to as a Carli and Jevons price index, respectively. On what grounds can we make a choice between both population indexes? A useful criterion, based on the axiomatic approach to index number theory, is to what extent they satisfy a set of axioms or tests, including the identity test mentioned earlier. See e.g. Balk (1995) for a review of the axiomatic approach. The Jevons index (19) satisfies all 'reasonable' tests (ILO, 2004). The Carli index (18), on the other hand, violates the time reversal test. This test says that if the data for periods 0 and t are interchanged, then the resulting price index should equal the reciprocal of the original price index. The time reversal test is a fundamental one since it is difficult to accept an index that gives a different picture if the ordering of time is reversed. Notice that the Dutot price index (1) does satisfy this test.

The fact that P_{UG}^t satisfies the time reversal test and P_{UA}^t does not, can formally be represented by

$$\prod_{i \in U^0} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{1}{N^0}} \prod_{i \in U^0} \left(\frac{p_i^0}{p_i^t} \right)^{\frac{1}{N^0}} = 1 \tag{20}$$

and

$$\frac{\sum_{i \in U^0} \left(\frac{p_i^t}{p_i^0} \right)}{N^0} \frac{\sum_{i \in U^0} \left(\frac{p_i^0}{p_i^t} \right)}{N^0} \geq 1. \tag{21}$$

Jensen's Inequality implies $P_{UA}^t \geq P_{UG}^t$, where equality only holds if the individual price indexes p_i^t / p_i^0 are identical for all $i \in U^0$. Based on Jensen's Inequality we also have $\sum_{i \in U^0} (p_i^0 / p_i^t) / N^0 \geq \prod_{i \in U^0} (p_i^0 / p_i^t)^{1/N^0}$. It follows that the product, given by (21), of P_{UA}^t and the unweighted arithmetic price index in which the prices of periods 0 and t are interchanged, must be greater than or equal to 1. So in general P_{UA}^t has an upward bias. For direct indexes the magnitude of the bias could be limited, but a chained variant of P_{UA}^t will exhibit upward drift, especially when prices oscillate. From the point of view of the axiomatic approach to index number theory the geometric index P_{UG}^t is preferred to the arithmetic index P_{UA}^t.

The derivation of the *unweighted arithmetic and geometric SPAR indexes* runs in a way similar to that of the value-weighted SPAR index. For a random sample S^t the following estimators of (18) and (19) can be defined (compare (5)):

$$\breve{P}_{UA}^t = \frac{\sum_{i \in S^t} \left(\dfrac{p_i^t}{p_i^0} \right)}{n^t} \tag{22}$$

and

$$\breve{P}_{UG}^t = \prod_{i \in S^t} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{1}{n^t}}. \tag{23}$$

$\breve{P}_{UA}^t$ is an unbiased estimator of P_{UA}^t. Due to the geometric structure this does not hold for $\breve{P}_{UG}^t$, which can be regarded as a disadvantage. On the other hand, an advantage of geometric means is that they are often less sensitive to outliers. Note that the bias of the sample geometric mean is usually negligible (see Balk, 2005). $\breve{P}_{UA}^t$ and $\breve{P}_{UG}^t$ are pseudo estimators since the base period prices are largely unknown. Assuming an identical non-stochastic ratio b between the market values and the appraisals during the base period for all houses, we obtain the unweighted SPAR price indexes by substituting $p_i^0 = a_i^0 b$ into (22) and (23) and subsequently dividing the outcomes by their base period values. This leads to

$$\hat{P}_{UA(S)}^t = \frac{\dfrac{1}{n^t} \sum_{i \in S^t} \left(\dfrac{p_i^t}{a_i^0} \right)}{\dfrac{1}{n^0} \sum_{i \in S^0} \left(\dfrac{p_i^0}{a_i^0} \right)} \tag{24}$$

and

$$\hat{P}^t_{UG(S)} = \frac{\prod_{i \in S^t} \left(\dfrac{p^t_i}{a^0_i} \right)^{\frac{1}{n^t}}}{\prod_{i \in S^0} \left(\dfrac{p^0_i}{a^0_i} \right)^{\frac{1}{n^0}}} = \left[\frac{\tilde{a}^0(S^0)}{\tilde{a}^0(S^t)} \right] \frac{\tilde{p}^t(S^t)}{\tilde{p}^0(S^0)} . \tag{25}$$

Due to the normalisation both types of unweighted SPAR price indexes satisfy the time reversal test.[7] So we are faced with the somewhat paradoxical situation that the unweighted arithmetic (Carli) target index (18) violates the time reversal test whereas the unweighted arithmetic SPAR index (24), which is an estimator of that population index, does satisfy this test.[8] We would nevertheless prefer the geometric SPAR index (25) as the target to be estimated – the unweighted geometric (Jevons) index (19) – does satisfy the time reversal test.

In the second expression on the right-hand side of (25), $\tilde{a}^0(S^0) = \prod_{i \in S^0} (a^0_i)^{1/n^0}$ and $\tilde{a}^0(S^t) = \prod_{i \in S^t} (a^0_i)^{1/n^t}$ denote the unweighted geometric average appraisals in the base and current period samples; $\tilde{p}^0(S^0) = \prod_{i \in S^0} (p^0_i)^{1/n^0}$ and $\tilde{p}^t(S^t) = \prod_{i \in S^t} (p^t_i)^{1/n^t}$ are the corresponding geometric average prices. Equation (25) has a similar structure as equation (14). Just like in (14), the bracketed factor in (25) can be interpreted as a factor which adjusts the ratio of mean prices $\tilde{p}^t(S^t) / \tilde{p}^0(S^0)$ for any differences in the quality mix of the sampled houses. This topic will be addressed again in Section 5 where we make a comparison with a bilateral hedonic time dummy index.

5. The Unweighted Geometric SPAR Index and Hedonic Regression

Suppose the semi-log hedonic regression model holds for all $i \in U^0$:

$$\ln p^0_i = \alpha^0 + \sum_{k=1}^{K} \beta^0_k x_{ik} + \varepsilon^0_i , \tag{26}$$

---

[7] The value-weighted SPAR index (13) also satisfies this test.

[8] This type of 'paradox' is not an unknown phenomenon in the price index literature. In case of a fixed set of goods that have been sampled proportionally to base period quantities, for example, the Dutot price index, which does satisfy the time reversal test, is an approximately unbiased estimator of the population Laspeyres price index, which violates this test. See Balk (2005).

13

where x_{ik} is the k-th characteristic of house i.[9] A semi-log specification usually leads to more satisfactory results than a completely linear one and has been successfully applied in much empirical work. Assuming we know the characteristics of all houses, an (OLS) regression of the semi-log model on the data of sample S^0 yields predicted base period prices $\hat{p}_i^0 = \exp[\hat{\alpha}^0(S^0)]\exp[\sum_{k=1}^{K}\hat{\beta}_k^0(S^0)x_{ik}]$.[10] Suppose the appraisals were obtained using regression model (26),[11] so that $a_i^0 = \hat{p}_i^0 = \exp[\hat{\alpha}^0(S^0)]\exp[\sum_{k=1}^{K}\hat{\beta}_k^0(S^0)x_{ik}]$ for all $i \in U^0$; hence, $\tilde{a}^0(S^0) = \prod_{i \in S^0}(a_i^0)^{1/n^0} = \exp[\hat{\alpha}^0(S^0)]\exp[\sum_{k=1}^{K}\hat{\beta}_k^0(S^0)\bar{x}_k(S^0)]$ and $\tilde{a}^0(S^t) = \prod_{i \in S^t}(a_i^0)^{1/n^t} = \exp[\hat{\alpha}^0(S^0)]\exp[\sum_{k=1}^{K}\hat{\beta}_k^0(S^0)\bar{x}_k(S^t)]$, in which $\bar{x}_k(S^0)$ and $\bar{x}_k(S^t)$ denote the unweighted arithmetic averages of attribute k in S^0 and S^t. We thus find $\tilde{a}^0(S^0)/\tilde{a}^0(S^t) = \exp\{\sum_{k=1}^{K}\hat{\beta}_k^0(S^0)[\bar{x}_k(S^0) - \bar{x}_k(S^t)]\}$, which makes clear that the ratio of the mean appraisals in both samples can be viewed as a factor that controls for changes in the quality mix. Substituting this mix-adjustment factor into expression (25) yields

$$\hat{P}_{UG(S)}^t = \exp\left[\sum_{k=1}^{K}\hat{\beta}_k^0(S^0)[\bar{x}_k(S^0) - \bar{x}_k(S^t)]\right]\frac{\tilde{p}^t(S^t)}{\tilde{p}^0(S^0)}. \tag{27}$$

Notice that $\hat{P}_{UG(S)}^t = \tilde{p}^t(S^t)/\tilde{p}^t(S^0)$ if $\bar{x}_k(S^t) = \bar{x}_k(S^0)$ for all k. Notice further that (27) would not change if we had assumed $a_i^0 = b\hat{p}_i^0$ ($b \neq 1$) instead of $a_i^0 = \hat{p}_i^0$. Expressions similar to (27) are well known from the hedonic price index literature; see e.g. Triplett (2004) and De Haan (2007b). In particular, an expression can be obtained which almost coincides with (27) using a weighted version of the time dummy variable hedonic approach. In this case the semi-log hedonic model is estimated on the pooled data of S^0 and S^t. The time dummy variable model reads

---

[9] We are using the same notation for the parameters and the disturbance term as in the linear model of Section 3. This should not lead to confusion.

[10] The predictors are not unbiased because the expected value of a nonlinear function (the exponent) is not equal to the nonlinear function of the expected value. See for example Goldberger (1968), who provides a correction term. In practice the bias can often be neglected.

[11] According to Bourassa et al. (2006), footnote 3, in New Zealand the appraisals are actually based on hedonic regression modelling but we do not know if the semi-logarithmic or some other specification has been used. The geometric formula is not applied in New Zealand to estimate SPAR indexes. Currently the value-weighted variant is applied. Prior to September 2004 the unweighted arithmetic variant was used. For more information see the Reserve Bank's website: www.rbnz.govt.nz/keygraphs/1697975.html. Note that Bourassa et al. (2006) do not discuss the geometric SPAR index.

$$\ln p_i^s = \alpha + \delta D_i + \sum_{k=1}^{K} \beta_k x_{ik} + \varepsilon_i^s \quad (s = 0, t), \tag{28}$$

where D_i is a dummy variable that takes on the value of 1 for $i \in S^t$ and the value of 0 for $i \in S^0$. The time dummy model is restrictive since the parameters are assumed fixed over time. Furthermore, we assume that (28) is estimated using Weighted Least Squares (WLS) regression with weights w_i^0 for $i \in S^0$ and w_i^t for $i \in S^t$. The predicted prices are $\widetilde{p}_i^0 = \exp(\widetilde{\alpha}) \exp[\sum_{k=1}^{K} \widetilde{\beta}_k x_{ik}]$ and $\widetilde{p}_i^t = \exp(\widetilde{\alpha}) \exp(\widetilde{\delta}) \exp[\sum_{k=1}^{K} \widetilde{\beta}_k x_{ik}]$. Because the time dummy model controls for changes in the characteristics, the exponent of the time dummy coefficient, $\exp(\widetilde{\delta}) = \widetilde{p}_i^t / \widetilde{p}_i^0$, is a quality-adjusted price index.

Due to the inclusion of an intercept term and a time dummy variable in model (28), the weighted residuals $\ln p_i^s - \ln \widetilde{p}_i^s = \ln(p_i^s / \widetilde{p}_i^s)$ $(s = 0, t)$ sum to zero in each time period (De Haan, 2004a, 2004b, 2007b):

$$\sum_{i \in S^0} w_i^0 \ln\left(\frac{p_i^0}{\widetilde{p}_i^0}\right) = \sum_{i \in S^t} w_i^t \ln\left(\frac{p_i^t}{\widetilde{p}_i^t}\right) = 0. \tag{29}$$

Exponentiating yields

$$\frac{\prod_{i \in S^0}(p_i^0)^{w_i^0}}{\prod_{i \in S^0}(\widetilde{p}_i^0)^{w_i^0}} = \frac{\prod_{i \in S^t}(p_i^t)^{w_i^t}}{\prod_{i \in S^t}(\widetilde{p}_i^t)^{w_i^t}} (= 1). \tag{30}$$

After substituting the predicted prices into (30) and some rearranging we find

$$\prod_{i \in S^t}\left[\exp(\widetilde{\delta})\right]^{w_i^t} = \frac{\prod_{i \in S^t}(p_i^t)^{w_i^t}}{\prod_{i \in S^0}(p_i^0)^{w_i^0}} \frac{\prod_{i \in S^0}\left[\exp(\widetilde{\alpha})\right]^{w_i^0}}{\prod_{i \in S^t}\left[\exp(\widetilde{\alpha})\right]^{w_i^t}} \frac{\prod_{i \in S^0}\left[\exp(\sum_{k=1}^{K}\widetilde{\beta}_k x_{ik})\right]^{w_i^0}}{\prod_{i \in S^t}\left[\exp(\sum_{k=1}^{K}\widetilde{\beta}_k x_{ik})\right]^{w_i^t}}. \tag{31}$$

Suppose that a WLS-procedure is used where the observations are weighted by the reciprocal of the sample sizes in both periods, that is $w_i^0 = 1/n^0$ and $w_i^t = 1/n^t$. In a sense we are treating the two time periods symmetrically by proportionally reducing the impact of the largest sample in the pooled regression. Now (31) reduces to

$$\widetilde{P}_{TD}^t = \exp(\widetilde{\delta}) = \frac{\prod_{i \in S^t}(p_i^t)^{\frac{1}{n^t}} \prod_{i \in S^0}\left[\exp(\sum_{k=1}^{K}\widetilde{\beta}_k x_{ik})\right]^{\frac{1}{n^0}}}{\prod_{i \in S^0}(p_i^0)^{\frac{1}{n^0}} \prod_{i \in S^t}\left[\exp(\sum_{k=1}^{K}\widetilde{\beta}_k x_{ik})\right]^{\frac{1}{n^t}}}$$

$$= \exp\left[\sum_{k=1}^{K} \widetilde{\beta}_k [\bar{x}_k(S^0) - \bar{x}_k(S^t)]\right] \frac{\widetilde{p}(S^t)}{\widetilde{p}(S^0)} . \tag{32}$$

The time dummy price index $\widetilde{P}_{TD}^t$, given by (32), differs from expression (27) only in the regression coefficients. In (32) a kind of average value over periods 0 and t is taken, whereas in (27) the coefficients relate to period 0 only. As a consequence of the pooling of data the coefficients $\widetilde{\beta}_k$ will likely have smaller standard errors than the coefficients $\hat{\beta}_k^0(S^0)$. Thus if the appraisals would be estimated using the semi-log hedonic model (26), then the geometric SPAR index is presumably less accurate than the time dummy index $\widetilde{P}_{TD}^t$. One the other hand, the WLS procedure used in the pooled regression may cause heteroskedasticity which leads to standard errors of the coefficients that are larger than would be strictly necessary.

The bilateral time dummy index $\widetilde{P}_{TD}^t$ is, in a sense, unnecessarily inefficient. Instead of running a regression on the data of periods 0 and t only, it would make more sense to estimate a time dummy index on the pooled data of all periods and incorporate dummy variables for all periods into the model, including those in between 0 and t. Such a multilateral or multi-period time dummy index has three things in common with a repeat-sales index: both are based on the assumption of fixed parameters, they become more efficient as time passes and more and more data is being added to the sample, and they suffer from revision. Revision is the phenomenon that previously computed figures will be 'revised' when additional data become available.[12]

6. Consistency in Aggregation

So far we have only looked at indexes for all houses. Users may also desire indexes for certain well-defined sub-populations, for example indexes for different regions or types

---

[12] Repeat-sales indexes belong to the class of stochastic price indexes (Clements et al., 2006). Stochastic indexes are estimated using regression techniques and will be revised in the course of time: they violate temporal fixity (Hill, 2004). This is also one of Diewert's (1995) criticisms of the stochastic approach. In his words (p. 20): "[Stochastic] price indexes are not invariant to the number of periods T in the sample". Some authors (e.g. Clapp and Giacotto, 1999, and Butler et al., 2005) speak of revision bias when there is a systematic, or asymmetric, adjustment in one direction. For repeat-sales indexes downward adjustments seem to prevail. In an empirical study on Swedish data Clapham et al. (2004) show that the revisions for hedonic multi-period time dummy index numbers are smaller than those for repeat-sales index numbers and that they are not asymmetric.

of dwellings. Suppose that the total population U^0 is subdivided into $h = 1,\ldots,H$ non-overlapping sub-groups or strata U_h^0 with sizes N_h^0. The value-weighted index (1) and the unweighted arithmetic and geometric indexes (18) and (19) can now be re-written as

$$P_W^t = \sum_{h=1}^{H} \left[\frac{\sum_{i \in U_h^0} p_i^0}{\sum_{i \in U^0} p_i^0} \right] P_{W,h}^t ; \tag{33}$$

$$P_{UA}^t = \sum_{h=1}^{H} \left(\frac{N_h^0}{N^0} \right) P_{UA,h}^t ; \tag{34}$$

$$P_{UG}^t = \prod_{h=1}^{H} \left(P_{UG,h}^t \right)^{\frac{N_h^0}{N^0}} , \tag{35}$$

with stratum price indexes $P_{W,h}^t = \sum_{i \in U_h^0} p_i^t / \sum_{i \in U^0} p_i^0$, $P_{UA,h}^t = \sum_{i \in U_h^0} (p_i^t / p_i^0)/N_h^0$, and $P_{UG,h}^t = \prod_{i \in U_h^0} (p_i^t / p_i^0)^{1/N_h^0}$. In (33), (34) and (35) the aggregate indexes are presented as weighted arithmetic and geometric averages of the stratum indexes; the base period value shares $\sum_{i \in U_h^0} p_i^0 / \sum_{i \in U^0} p_i^0$ and the relative number of dwellings N_h^0 / N^0 serve as weights. The population indexes are *consistent in aggregation*: the aggregate indexes can be written as weighted averages of stratum indexes that have the same mathematical form.

Unlike the target indexes, the SPAR indexes are not consistent in aggregation. By post-stratifying the total samples S^0 and S^t according to the above scheme we find $h = 1,\ldots,H$ sub-samples S_h^0 and S_h^t. The value-weighted SPAR index for stratum h with n_h^s observations ($s = 0,t$) is given by $\hat{P}_{W(S),h}^t = [\sum_{i \in S_h^t} p_i^t / \sum_{i \in S_h^t} a_i^0]/[\sum_{i \in S_h^0} p_i^0 / \sum_{i \in S_h^0} a_i^0]$; $\hat{P}_{UA(S),h}^t = [\sum_{i \in S_h^t} (p_i^t / a_i^0)/n_h^t]/\sum_{i \in S_h^0} (p_i^0 / a_i^0)/n_h^0$ is the unweighted arithmetic variant and $\hat{P}_{UG(S),h}^t = \prod_{i \in S_h^t} (p_i^t / a_i^0)^{1/n_h^t} / \prod_{i \in S_h^0} (p_i^0 / a_i^0)^{1/n_h^0}$ the unweighted geometric one. In case of the unweighted indexes, for example, the aggregation problem implies that the aggregate indexes cannot unambiguously be written as weighted averages of stratum-specific SPAR indexes using the relative sub-sample sizes as weights. If the sampling fractions are identical for all h in periods 0 and t ($n_h^0 / n^0 = n_h^t / n^t = f_h$) we know for sure that $\hat{P}_{UA(S)}^t = \sum_{h=1}^{H} f_h \hat{P}_{UA(S),h}^t$ and $\hat{P}_{UG(S)}^t = \prod_{h=1}^{H} (\hat{P}_{UG(S),h}^t)^{f_h}$. In practice this situation will hardly ever occur. It therefore seems reasonable to estimate the aggregate indexes by

$$\widetilde{P}_{UA}^{t} = \sum_{h=1}^{H} \left(\frac{N_h^0}{N^0} \right) \hat{P}_{UA(S),h}^{t} \tag{36}$$

and

$$\widetilde{P}_{UG}^{t} = \prod_{h=1}^{H} \left(\hat{P}_{UG(S),h}^{t} \right)^{\frac{N_h^0}{N^0}}, \tag{37}$$

which are found by replacing the stratum population indexes in (34) and (35) by the stratum-specific SPAR indexes. An advantage is that the assumption of a fixed ratio between market values (prices) and appraisals can be relaxed and applied to the stratum level. A drawback might be that post-stratification could raise the standard error of the aggregate index in case of very small sub-sample sizes.

The problem is a little bit more complex for the value-weighted SPAR index. Again the SPAR index is not consistent in aggregation. To estimate an aggregate index in a similar way as (36) and (37) we must know the value shares $\sum_{i \in U_h^0} p_i^0 / \sum_{i \in U^0} p_i^0$. Those are of course unknown. Under the assumption of a fixed ratio between the market values and the appraisals for all $i \in U^0$ we could replace the market value shares by the appraisal-based shares $\sum_{i \in U_h^0} a_i^0 / \sum_{i \in U^0} a_i^0$ and estimate the aggregate index as

$$\widetilde{P}_{W}^{t} = \sum_{h=1}^{H} \left[\frac{\sum_{i \in U_h^0} a_i^0}{\sum_{i \in U^0} a_i^0} \right] \hat{P}_{W(S),h}^{t}. \tag{38}$$

Since the assumption will be violated, the appraisal-based shares are rough estimates of the true value shares. More importantly, the above-mentioned advantage of making this assumption at the stratum level instead of the aggregate level no longer holds.

7. Dynamic Population and Sampling

Apart from quality changes of existing houses, which will be touched upon at the end of this section, the total stock of owner-occupied houses is dynamic in two respects. Some houses will be demolished and disappear from the stock while newly-built houses and existing houses which are sold by, for example, local governments or 'corporations' to private households are added to the stock. The assumption of a fixed housing stock U^0 that has been maintained so far should be relaxed. The problem of new and disappearing

houses raises several issues. What is precisely meant by the stock of existing houses? If we know the answer, which index number formula should be used to measure the price change of the dynamic stock? And is the available sample appropriate for calculating a SPAR index which aims at estimating the population index?

Let us begin with the first question. Although the term 'stock of existing houses' is common in the literature (e.g. in Hwang and Quigley, 2004), it is a bit of a misnomer since any newly-built house exists and belongs to the target population we are interested in once it has been sold (for the first time). Consequently, a distinction between newly-built houses and existing houses is irrelevant for our main goal: any index that is meant to measure the changes in the market value of the stock of 'existing' houses should in principle take newly-built houses into account. There are nevertheless several problems that need attention.

The value index of a dynamic housing stock is $\sum_{i \in U^t} p_i^t / \sum_{i \in U^0} p_i^0$, where U^t denotes the total stock in period t. If U^t changes over time, this value index is affected by quantity changes and cannot be called a price index. Suppose first that some houses disappear but no new houses are added to the stock, so that the total stock shrinks. The natural way to construct a value-weighted price index would still be using P_W^t, given by (1), and imputing the 'missing prices' of period t for houses that disappear from U^0. The use of explicit hedonic imputations is preferable. Imputations are implicitly made if we apply the value-weighted SPAR index and view it as an estimator of (1). Given the lack of a hedonic benchmark index, the impact of such implicit imputations cannot be measured.

Suppose next that houses are added to the stock but that no houses disappear, so that the total stock expands. Assuming, as argued above, that the price index should also relate to new houses, the following target index might be considered:

$$P_P^t = \frac{\sum_{i \in U^t} p_i^t}{\sum_{i \in U^t} p_i^0} . \tag{39}$$

This index could be referred to as a Paasche price index. Four remarks are in place. Firstly, imputations are needed for the missing base period prices of the new houses, something that is done implicitly using the SPAR method. Secondly, when applying the SPAR method the base period appraisals must be known for the new houses; thus, in

each period t the (fictitious) appraised values that would have been attached to the new dwellings had they existed in the base period must be available. In the Netherlands appraisals for newly-built houses are calculated retrospectively as they are needed for income tax and local tax purposes. Thirdly, it is important to realize that S^t is a sample from the period t stock U^t rather than from the base period stock U^0. As a result, if the value-weighted SPAR index were based on S^t it would be aiming at (39) instead of (1). Fourthly, the period-to-period changes P_P^t / P_P^{t-1}, as well as the annual changes, will be affected by changes in the population. This is the dynamic-population counterpart to the period-to-period changes in the sample-based value-weighted SPAR index that is also affected by compositional changes, as we saw in Section 3.

A symmetric treatment of the periods 0 and t is preferable in a price index. From this point of view the Fisher price index $P_F^t = [P_W^t P_P^t]^{1/2}$ would be a sensible choice. The construction of a chain index might also be considered, particularly when the dynamics are substantial, i.e. when many houses appear or disappear. Unfortunately, short-term chain indexes may exhibit drift if the data exhibit a seasonal pattern. Moreover, the use of the SPAR method to estimate period-to-period chain Paasche or Fisher price indexes would require appraisals for each period t-1, which are of course not available. So this is not a feasible approach.

All in all the use of P_W^t seems both unavoidable and acceptable provided that the dynamics are limited. To estimate this index using the value-weighted SPAR method, houses that were added to the stock after the base period must be removed from sample S^t. This can easily be attained by coupling the sample to the base period housing stock data set. The same goes for unweighted SPAR indexes. If the dynamics are substantial the use of a chain index on an annual basis would be preferable. As from January 2008 an annual chain index can be constructed in the Netherlands because houses will then be appraised annually. Since the houses are not appraised in real time but retrospectively, there will inevitably be a one or two year time lag involved, but that does not seem too problematic.

In the preceding sections we have assumed that in each period the set of houses sold (excluding new houses) is a random sample from the base period housing stock. This is a strong assumption. Some types of houses, in the Netherlands apartments and relatively cheap single-family dwellings, are sold more frequently than others and have a higher probability to be included in the 'sample'. Furthermore, the land registry only

records the prices of second-hand houses; the prices of newly-built houses – which we would like to include in the base period sample – are not recorded. If random selection does not hold and different types of houses or different regions, etc. show different price changes, sample selection bias will occur. In contrast to the small sample bias given by (17), which stems from the non-linearity of the SPAR index, little can be said a priori about the sign and magnitude of sample selection bias. This problem could be mitigated by post-stratifying the sample as outlined in Section 6.

So far we have been assuming that the quality of each house remains constant. In reality the quality of a house may deteriorate due to wear and tear, especially if major repairs are omitted, but it may also improve due to renovations or extensions. While the SPAR method controls for changes in the quality mix of the sample, it does not control for such quality changes.[13] It has been suggested to us to adjust the prices for property improvements that require a buiding permit. In the Netherlands such an adjustment is unfortunately infeasible because permit data are available only at aggregate levels but not for individual houses. Our suggestion to construct an annually chained SPAR index is more promising. Assuming that the appraisals take account of major improvements, subsequent appraisals can possibly be used to construct annually chained SPAR price indexes that will to some extent control for quality changes. This is beyond the scope of the present paper and left for future research.

8. An Expenditure-Based Interpretation of the Value-Weighted SPAR Index

Our focus has been on the construction of a price index for the total stock of houses as we feel that this should be a country's primary house price index. In the introduction we mentioned, however, that Eurostat prefers a net acquisitions approach to incorporating owner-occupied housing into the HICP. The weight reflects net expenditures on newly-built and re-sold, second-hand houses, thus excluding purchases within the household sector and also excluding purchases of land. Data problems in some (smaller) countries and the desire to serve the needs of users that are interested in house price growth per

---

[13] The same applies to the repeat-sales method. Hybrid methods, which combine repeat sales methods and hedonic regression, have been developed to overcome this problem (Wang and Zorn, 1997).

se, have led Eurostat to broaden the scope. More specifically, a 'Laspeyres-type' price index must be compiled for all houses purchased by private households during a certain base year in which house prices should include the price of land. Under an acquisitions or expenditure approach a value-weighted index is most appropriate.[14] Below we show that the value-weighted SPAR has a straightforward interpretation as an expenditure-based house price index. Since it is not a Laspeyres-type index we briefly outline how to construct a stratified value-weighted SPAR index that is most likely to meet Eurostat requirements.

In the land registry's data set all transactions are registered. So far the total set of houses sold in the current month t, denoted by S^t, has been viewed as a sample from the fixed base period housing stock U^0. From an expenditure point of view S^t is not a sample, however, and $\sum_{i \in S^t} p_i^t$ equals total current period expenditure on second-hand houses. Hence equation (5) describes the population Paasche price index for purchases of second-hand houses, denoted here by $P_{P(E)}^t$:

$$P_{P(E)}^t = \frac{\sum_{i \in S^t} p_i^t}{\sum_{i \in S^t} p_i^0}. \tag{40}$$

We still assume the stock of houses fixed over time. This is not necessary. If population data on new houses were available (which is not the case in the Netherlands), the prices could be included in (40).

Since the base period prices p_i^0 are generally unknown for $i \in S^t$, estimator (8) or what should now be called the imputation Paasche index

$$\hat{P}_{P(E)}^t = \frac{\sum_{i \in S^t} p_i^t}{\sum_{i \in S^t} \hat{p}_i^0} \tag{41}$$

---

[14] Not everyone shares this view. Hill and Melser (2007, footnote 7) note that: "a case can be made for giving equal weight to all house sales, rather than weighting by expenditure shares. The usual justification for weighting price indexes by expenditure shares is so that the basket of goods and services will be representative for the average household. Housing is unusual, however, in that most households only buy one house, and the distribution of house prices is significantly skewed to the right. Most households therefore are not buying from the right tail. Hence it may be preferable to give all houses equal weight in the index [....], since this will provide an index that is more representative for the median household."

measures what the Paasche index given by (40) would have been when all houses sold during the current period had also been sold during the base period. The usual approach to estimating such 'missing prices' is to rely on hedonic regression. The value-weighted SPAR, given by expression (13), is an alternative estimator of (40). Again writing the appraisals as $a_i^0 = b\hat{p}_i^0$, the value-weighted SPAR index can be re-written as

$$\hat{P}_{W(S)}^t = \left[\frac{\sum_{i \in S^0} \hat{p}_i^0}{\sum_{i \in S^0} p_i^0} \right] \frac{\sum_{i \in S^t} p_i^t}{\sum_{i \in S^t} \hat{p}_i^0} . \tag{42}$$

Note that index volatility cannot be caused by sampling variance when the SPAR index is interpreted in the above way because there is no sampling involved – this volatility is either real or due to 'model variance' or measurement error that results from using the appraisals as auxiliary information. The better the fit of a linear regression of the base period sale prices on appraisals, the smaller this model variance is likely to be. For the same reason we can no longer speak of sample selection bias. Stratification of the value-weighted SPAR index on this account would be meaningless.

Stratification still makes sense for two other reasons though: to relax the basic assumption underlying the SPAR method and to compute Laspeyres-type price indexes. The basic assumption is that of a uniform sale price appraisal ratio, apart from a random disturbance term. In the Netherlands this assumption might be violated at the aggregate level because of the autonomous way in which municipalities determine the appraisals. By stratififying according to region, and possibly also according to type of dwelling, we expect the basic assumption to hold approximately at the level of the strata. The number of strata to be distinguished is limited since many strata with few observations can make the aggregate index unstable.

Stratification is also helpful for estimating a price index pertaining to all second-hand houses purchased during a particular base year that fulfils Eurostat requirements. This 'gross acquisitions' index should be a Laspeyres-type price index, with the weights reflecting base year expenditures. Several European countries combine stratification and hedonic regression to control for compositional change while some other countries are planning to do so. Due to a lack of data on housing attributes the use of hedonics is infeasibele in the Netherlands, and we propose a combination of stratification with the use of the value-weighted SPAR index at the stratum level.

Following the notation used in Section 6 the stratum-specific SPAR index and the set of houses sold during the base period are denoted $\hat{P}_{W(S),h}^{t}$ and S_h^0, respectively, for strata $h=1,\ldots,H$. Let us begin by holding on to the appraisal reference month 0 as the base period. Then an expenditure-based 'Laspeyres-type' SPAR price index may be defined by

$$\hat{P}_{L(E)}^{t} = \sum_{h=1}^{H} \left[\frac{\displaystyle\sum_{i\in S_h^0} p_i^0}{\displaystyle\sum_{i\in S^0} p_i^0} \right] \hat{P}_{W(S),h}^{t} . \tag{43}$$

$\hat{P}_{L(E)}^{t}$ can be criticised for its hybrid nature: fixed base period expenditure weights are combined with Paasche indexes at the stratum level. On the other hand, it is not unusual in price index construction to make a distinction between upper level and lower level aggregation methods. In a CPI lower level indexes are often unweighted. The difference between $\hat{P}_{L(E)}^{t}$ and the stratified value-weighted SPAR index $\tilde{P}_{W}^{t}$ given by (38) is clear. The latter uses estimated value shares for the stock of housing in month 0 whereas $\hat{P}_{L(E)}^{t}$ uses expenditure shares for the houses bought in that month.

Eurostat requires a year instead of a month as the base period. The computation of expenditure shares on a yearly basis does not pose any problems since sale prices for the population are known. To calculate the stratum indexes we suggest normalising the initial current period stratum values by dividing them by their average monthly values across the whole base year. This is admittedly a crude approach, but it mimics what many statistical agencies actually do when rebasing their index.

9. Data

The Dutch housing stock consists of approximately 7.8 million houses in 2007, half of which are owner-occupied. It concerns mainly flats and single-family dwellings. In the Netherlands real estate transactions are registered by the land registry office. As the land registry does not register prices of newly-built houses, the sale price data only refer to purchases of second-hand houses. Our data set comprises monthly sale price data for the period January 1995 – December 2007 and a limited number of characteristics of the houses, in particular address (including postal code) and type of dwelling. On an annual basis some 200 thousand owner-occupied houses change hands.

Figure 1 reveals a seasonal pattern in the number of sales: in July and especially in December relatively many houses are sold, in January relatively few. It should be mentioned that the data are registered at the date the definitive contract of sale has been signed at a notary's office. However, on average the sale is agreed two months earlier, namely when a provisional contract is signed. From an economic point of view this is when the actual transaction takes place, so that the 'true' peaks in the number of sales occur in October rather than in December. Also, the price indexes presented in Section 8 – including the land registry's repeat-sales index – essentially lag two months behind. It is important to bear this in mind when interpreting the results and analysing the Dutch housing market.
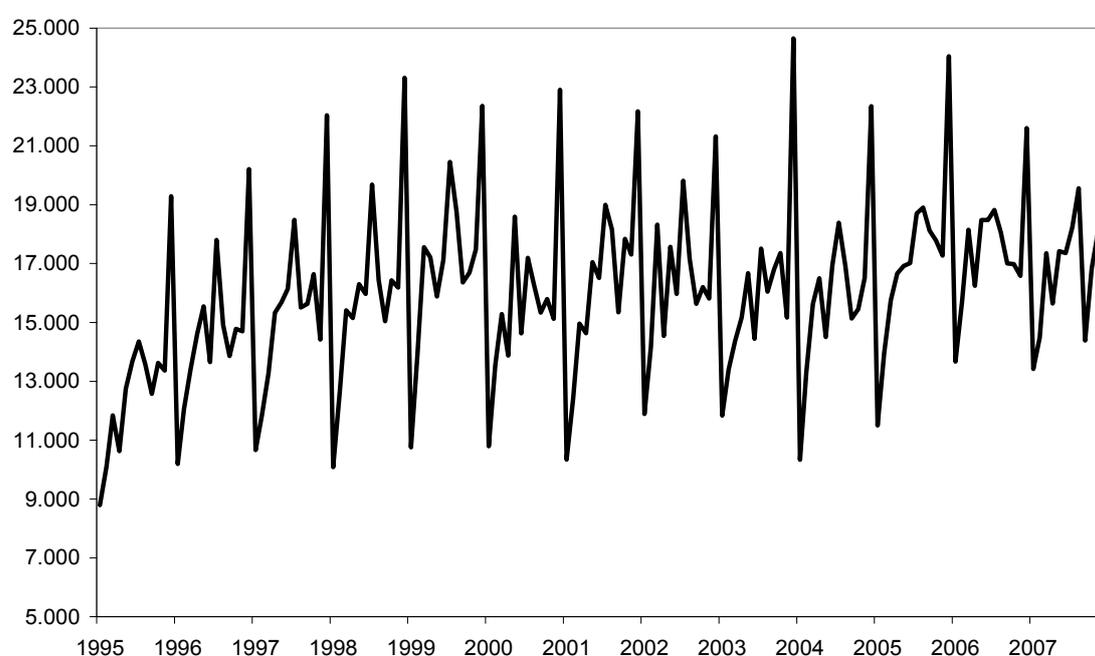
Figure 1. Number of sales for second-hand houses

Official government appraisals are available from January 1995 onwards. Under the Real Estate Law the municipalities are obliged to re-estimate the market value of all dwellings periodically. So far there have been four appraisal reference months: January 1995, January 1999, January 2003 and January 2005. As from January 2007 appraisals should be made every year. The land registry's data set has accordingly been split into four sub-periods: January 1995 – January 1999, January 1999 – January 2003, January 2003 – January 2005, and January 2005 – December 2007. January 1999, January 2003

and January 2005 serve as overlap periods. This is to create a time series by multiplying or 'chaining' the SPAR index numbers for the sub-periods during the overlap months.

The land registry's sale price data and the official appraisals relate to the same concept of price, including land. During the reference months we expect to find a high correlation between actual sale prices and appraisals. Figure 2 plots appraisals against sale prices (in thousand euros) for second-hand houses sold in January 2005, the most recent appraisal reference period. To investigate the empirical relationship between sale prices and appraisals, we ran various regressions, including logarithmic ones. The linear model did not seem to perform worse than the logarithmic models, which is reassuring. Using the notation of Section 3, the estimated linear model shown in Figure 2 reads $\hat{p}_i^0 = 1.835 + 0.975 a_i^0$, with $R^2 = 0.963$. Note that we did some data cleaning. Houses for which the appraisals differed more than 50% from the sale prices were deemed unacceptably large outliers and were deleted from the initial data set. It concerns less than 3% of the observations.

Figure 2 confirms that the distribution of house prices is skewed to the right. It also points to heteroskedasticity: the variance of the sale price seems to increase when the appraised value increases. Thus, if we would split the housing stock up into a lower and a higher priced segment of equal sizes, the standard error of the SPAR index for the higher priced segment is likely to be greater for two reasons: fewer observations (giving rise to larger sampling variance) and a larger variability of the sale price appraisal ratios (resulting in larger model variance).

Although the correlation in Figure 2 is high, one may wonder why 4% of the variation in sale prices is still not 'explained' by appraisals. One reason is that the way in which the appraisals are carried out has not yet been fully harmonised across Dutch municipalities. Moreover, the appraisals are established retrospectively – usually there is a time lag of one or two years involved. For example, many appraisals relating to January 2005 may not have been determined before the end of 2006. The municipalities differ both in the valuation method as such – though most of them look at sale prices of comparable houses in the neighbourhood – and in the way the values are 'deflated' to obtain appraisals for the reference month.[15]

---

[15] If a house has been renovated between the reference month and the date the appraisal has actually been performed, the appraised value includes the improvement. This may perhaps seem somewhat strange but a positive aspect is that the SPAR price indexes will be partially adjusted for such quality improvements.
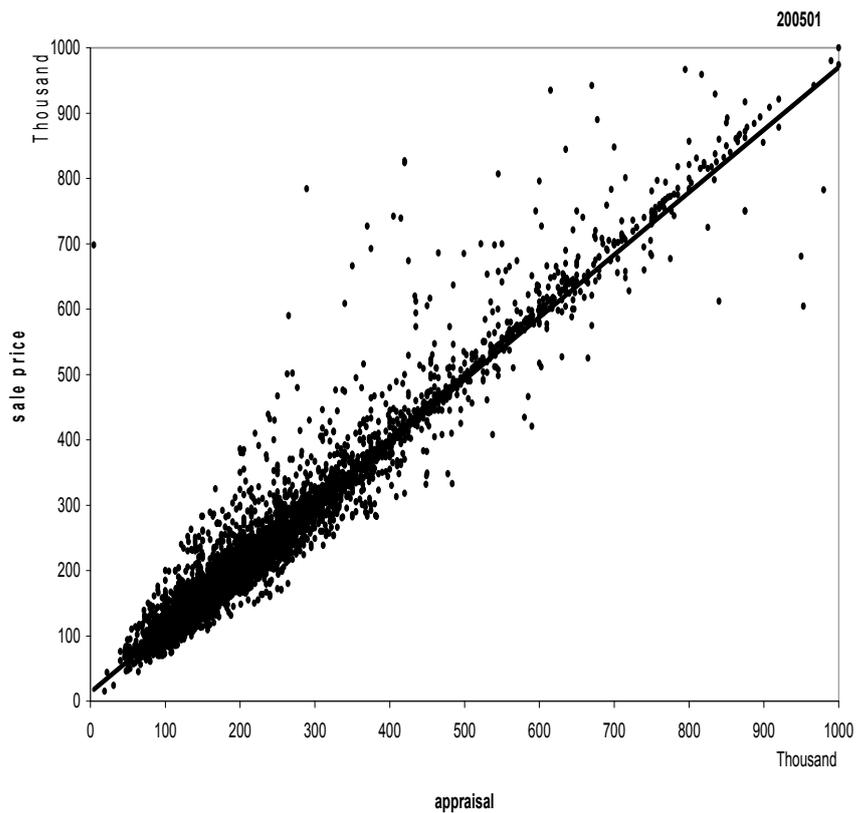
Figure 2. Scatter plot and linear OLS regression line of sale prices and appraisals, January 2005

The basic assumption underlying the SPAR method is that of a fixed ratio between sale prices (or market values) and appraisals in the base or reference appraisal period. Less strictly formulated, for the SPAR method to produce approximately unbiased results the regression line should go through the origin; it is not necessary that the slope coefficient is equal to unity. In January 2005 the estimated intercept term does differ significantly from 0 ($p < 0.001$), but the value of 1.84 thousand euros is so small that we do not have to worry about it too much. Table 1 summarizes the results of the regressions of the sale prices p_i^0 on the appraisals a_i^0 for all four appraisal reference months. Measured by the goodness of fit (R^2) the reliability of the appraisals appears to have increased somewhat in the course of time.

Unfortunately we do not know which houses in the data set it concerns. For more details on the Dutch appraisal data, see De Vries et al. (2008).

Table 1. OLS regression results (standard errors in parentheses)

Reference month	$\hat{\gamma}$ (intercept)	$\hat{\mu}$ (slope)	R^2	n
January 1995	6.442 (0.369)	0.959 (0.004)	0.903	7197
January 1999	-1.190 (0.414)	0.983 (0.003)	0.938	8958
January 2003	3.439 (0.506)	0.975 (0.002)	0.952	10186
January 2005	1.835 (0.488)	0.975 (0.002)	0.963	9665

Table 2 provides some additional summary statistics to assess the reliability of the appraisals. In January 1995 the mean appraisal value $\bar{a}^0(S^0)$ is 3.3% lower than the mean sale price $\bar{p}^0(S^0)$. Note that the ratio $\bar{p}^0(S^0)/a^0(S^0)$ in the third column equals the denominator of the value-weighted SPAR price index given by expression (13). The same picture emerges from the individual ratios p_i^0/a_i^0 in January 1995; the appraisals underestimate the sale prices on average by more than 4%.[16] The standard deviation of the ratios is quite large (0.161). In January 1999, January 2003 and January 2005 the appraisals overestimate the sale prices, which is in fact a little bit surprising. However, the mean value of the sale price appraisal ratios increasingly approaches 1, at least until 2005, and the standard deviation becomes smaller. The conclusion seems justified that the appraisals tend to become better approximations of the sale prices. In a similar study for a single province (Overijssel) of the Netherlands, De Vries et al. (2006) reached the same conclusion.

Table 2. A comparison of sale prices and appraisals in appraisal reference months

Reference month	$\bar{p}^0(S^0)$ (thousand €)	$\bar{a}^0(S^0)$ (thousand €)	$\dfrac{\bar{p}^0(S^0)}{\bar{a}^0(S^0)}$	p_i^0/a_i^0 mean	stand. dev.
January 1995	90.6	87.8	1.033	1.043	0.161
January 1999	130.9	134.4	0.974	0.975	0.114
January 2003	201.3	202.9	0.992	0.994	0.105
January 2005	212.2	215.8	0.983	0.985	0.099

---

[16] The arithmetic mean of the individual ratios is equal to the denominator of the unweighted arithmetic SPAR index given by (22). The corresponding geometric mean – which is equal to the denominator of the unweighted geometric SPAR index (23) – would of course be slightly lower.

10. Results

This section provides some results. Figure 3 shows monthly SPAR price index numbers for the Netherlands during January 1995 – December 2007, where January 1995 serves as the index reference period. Prices of second-hand houses have risen very sharply. In December 2007 the value-weighted SPAR index amounts to 282.5, which is equivalent to an average monthly increase of 0.7%. Prior to 1999 the unweighted arithmetic and geometric SPAR indexes nearly coincide with the value-weighted SPAR, but from 1999 onwards the unweighted indexes, especially the geometric indexes, are slightly lower.
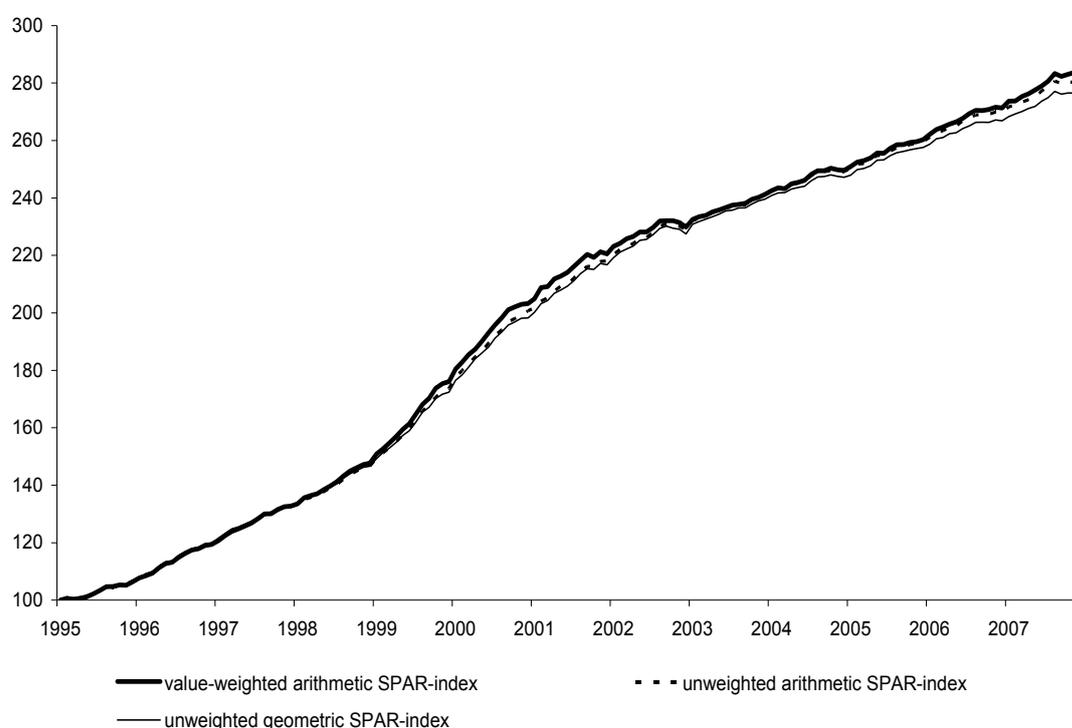
Figure 3. SPAR price indexes (January 1995= 100)

A comparison of this index with the value-weighted (arithmetic) SPAR index tells us that more expensive houses exhibited a larger price increase than less expensive ones, especially after 1999. As already mentioned in Section 7, cheaper houses are sold more frequently in the Netherlands than expensive houses. This will introduce sample selection bias into all types of price indexes. However, sample selection bias in SPAR indexes is potentially less severe than in a repeat-sales index. This is because the repeat-sales method by definition excludes houses that were sold only once during the period

January 1995 – December 2007, many of which are relatively expensive. Bourassa et al. (2006) also argue that the SPAR method is less prone to sample selection bias than the repeat-sales method.

A standard repeat-sales index essentially aims at the unweighted geometric mean of price relatives (Wang and Zorn, 1997), so comparing this index with the unweighted geometric SPAR will make a lot of sense. This is done in Figure 4, where the repeat-sales index refers to the house price index published by the Dutch land registry office until January 2008.[17] As a matter of fact the repeat-sales index overstates the geometric SPAR more than expected. In December 2007 the difference has increased to 23.0 index points.
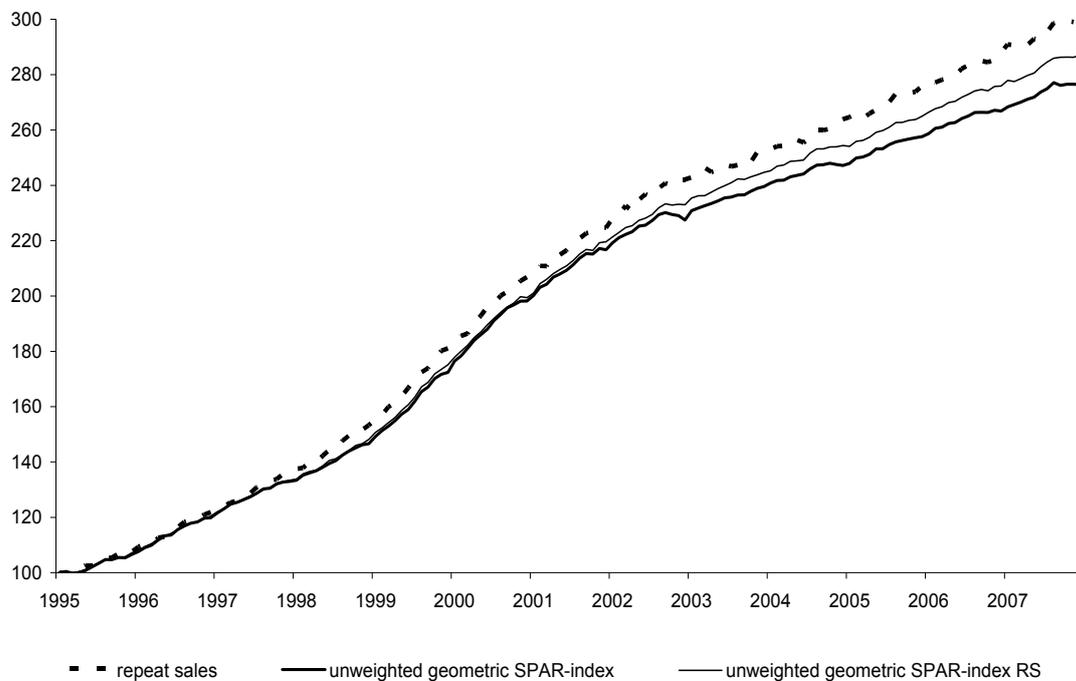
Figure 4. SPAR and repeat-sales price indexes (January 1995= 100)

To gain insight into this large difference we re-computed the geometric SPAR index using the repeat-sales data set. The result is also shown in Figure 4. This accounts

---

[17] This index is actually a weighted variant of the repeat-sales approach, where weighting of the data is applied to reduce heteroskedasticity (Jansen et al., 2005). The motivation behind weighting is that the variance of the log price changes may increase over time, in particular due to differences in maintenance. In practice the weighted repeat-sales index numbers differ not very much from the unweighted index numbers.

for almost half of the difference. Restricting the data set to houses that have been sold more than once clearly gives rise to substantial sample selection bias, which is in an upward direction in this particular case. However, more than half of the difference between the repeat-sales index and the geometric SPAR index is still unexplained. A part of this might be related to the repeat-sales estimation procedure and the revision that goes with it. The regression methodology is such that the estimated growth rate in house prices in a certain period is determined not only by the prices of houses sold in that period, but also by the prices of houses for which consecutive transactions occurred at either side of that period. The most recent periods are subject to particularly large revision as the houses that changed hands in a period are a relatively small proportion of the total number of sale prices that will eventually affect the index in that period.

The revisions of the Dutch land registry's repeat-sales index are indeed typically in a downward direction, something which is generally the case for repeat-sales indexes in other countries. Yet, given the magnitude of the revisions so far, it seems unlikely that future revisions will close the whole gap in December 2007 between the repeat-sales index and the geometric SPAR index based on the repeat-sales data set. While we cannot entirely preclude the possibility of downward bias in the SPAR index, we have difficulties in finding any reasons for this. According to Case et al. (1997) additional upward bias in the repeat-sales index may be due to the fact that estimated house price growth is higher among properties that sell more frequently because purchasers make property investments that are not adequately controlled for.

Figure 5 compares the value-weighted SPAR index with a 'naïve' index, i.e. the ratio of the arithmetic mean sale prices in the current period and the base period. There appears to be a systematic difference between both indexes, the naive index being lower than the SPAR since 1997. The extremely volatile behaviour of the naive index accords with a priori expectations. Apart from sampling error, this volatility must be caused by compositional changes affecting the index.
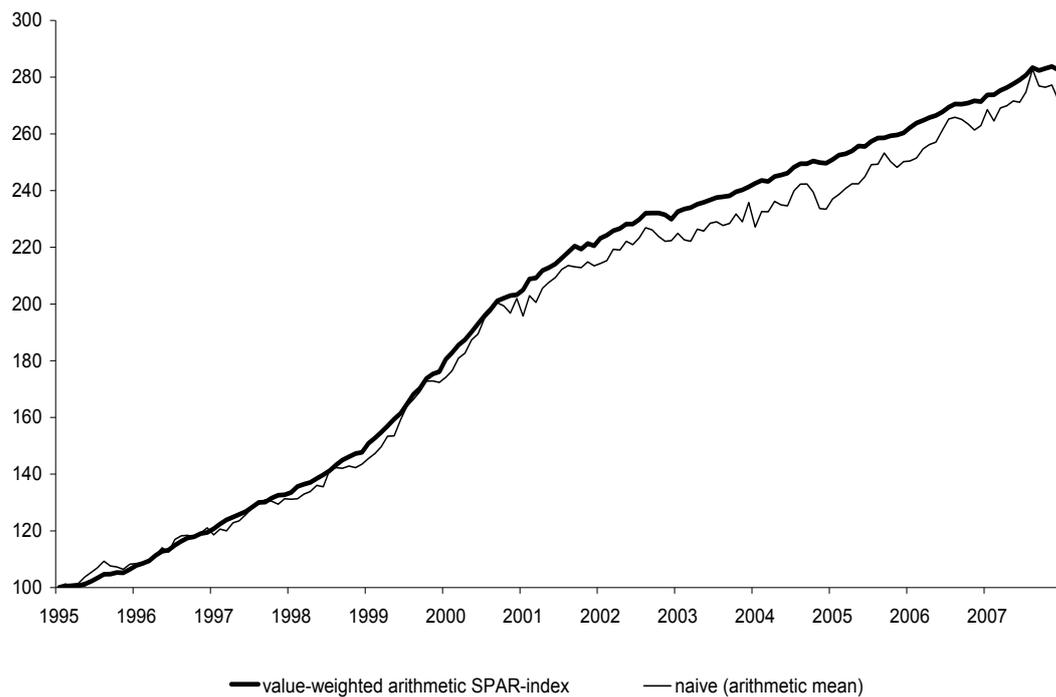
Figure 5. Value-weighted SPAR price index and 'naive' index (January 1995= 100)

In Section 3, equation (14), it was shown that the value-weighted SPAR index can be written as the product of the naive index and a so-called mix-adjustment factor. This factor, which equals the ratio of the average appraisals in the base period sample and the current period sample, essentially adjusts the naive index for any changes in the quality mix, and hence controls for compositional change. Figure 6 illustrates how the mix-adjustment factor has changed over time. Except for 1995-1996 its value is smaller than 1, suggesting that the average quality of the houses in the current period sample is systematically higher than in the base period sample.
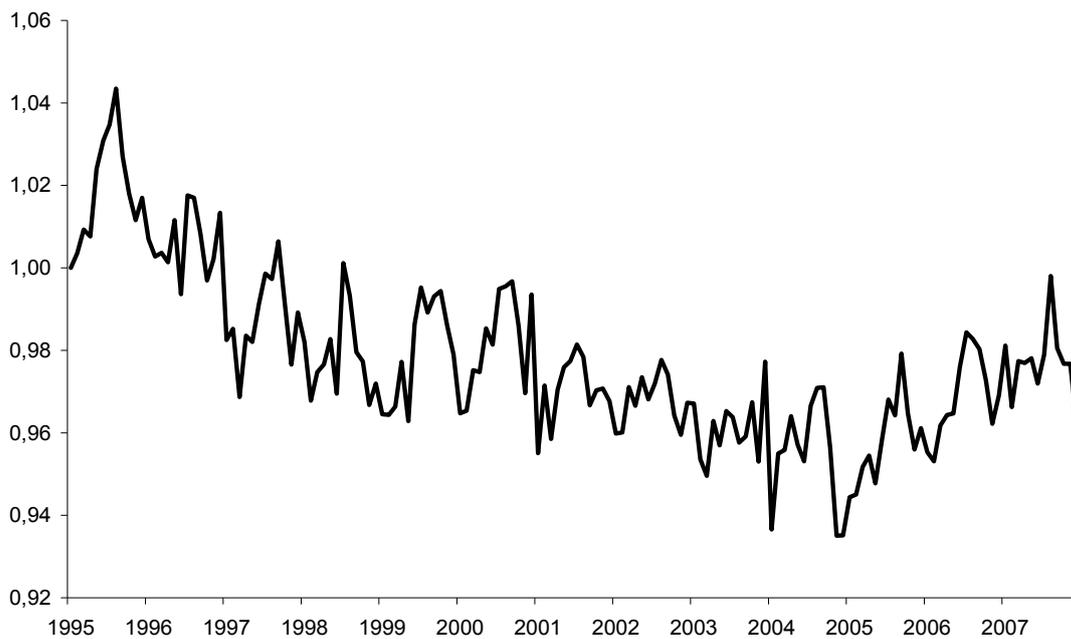
Figure 6. Mix-adjustment factor (ratio of value-weighted SPAR and naive index)

Figure 7 depicts the monthly percentage changes in the value-weighted SPAR index. The time series is definitely smoothed as compared to the naive index, but the monthly fluctuations are still considerable. While these short-term fluctuations will to some extent reflect true price changes, including seasonality, it is most likely that the volatility of the SPAR time series is largely due to compositional changes – which, as mentioned in Section 3, do affect the month-to-month changes – and to sampling error.

Figure 7 also shows the percentage changes with respect to the previous month of the repeat sales index. The volatility of the repeat sales index is greater than that of the value-weighted SPAR index. The reason is that the repeat sales method is based on a much smaller number of observations. Notice that the month-to-month changes of the value-weighted SPAR and the unweighted geometric SPAR almost coincide. Thus, the repeat-sales method performs worse than the SPAR method, both in terms of bias and short-term variability.
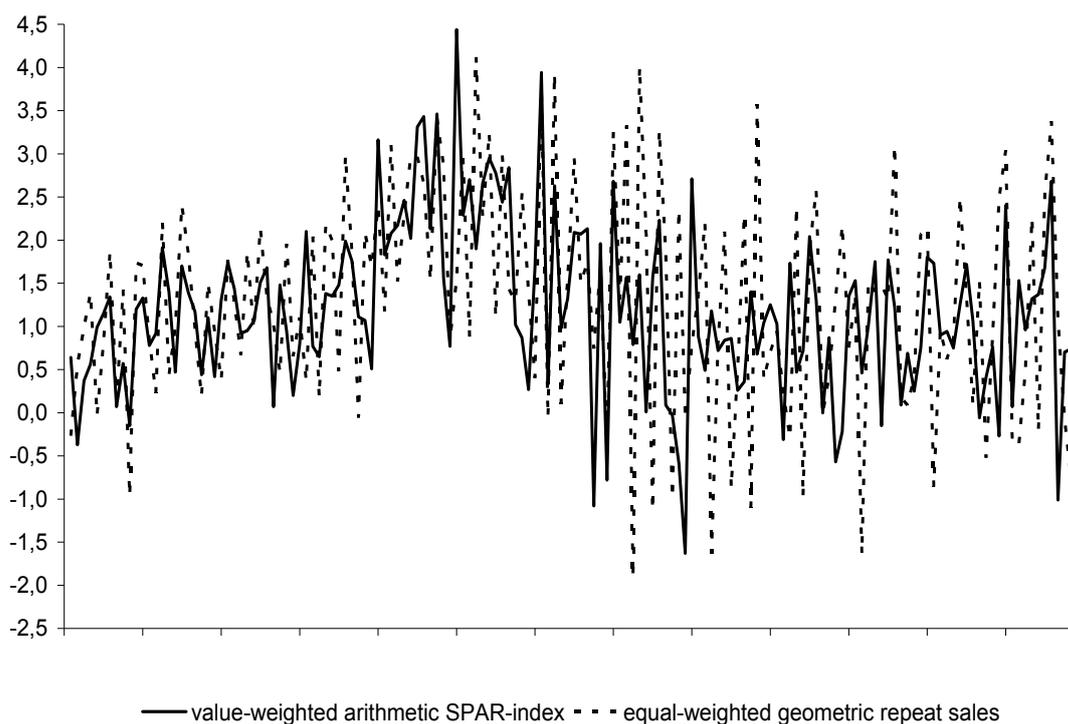
——— value-weighted arithmetic SPAR-index ▪ ▪ ▪ equal-weighted geometric repeat sales

Figure 7. Monthly percentage index changes

We also computed stratified SPAR indexes according to equations (36), (37) and (38).[18] The strata have been obtained by cross-classifying the twelve Dutch provinces and five types of dwellings (apartments, terraced houses, corner houses, semi-detached houses and detached houses), resulting into 60 cells. The average number of monthly sales per stratum is about 230, but in many cases the number of observations was much smaller, sometimes even below 10. Surprisingly, the differences between the ordinary un-stratified SPAR indexes and their stratified counterparts are very small, the stratified index being almost two index points higher during the ten-year period. So while there are a priori reasons to prefer stratification, for example to reduce sample selection bias, in practice it hardly affects the aggregate index. Something similar goes for the stratified Laspeyres-type price index described in Section 8: this index nearly coincides with the ordinary value-weighted SPAR index.

However, because stratum indexes and sub-aggregates based on stratum indexes may be desirable in their own right, it is useful taking a glance at some of the results.

---

[18] To compute stratified unweighted SPAR indexes we need the total number of houses per stratum in the housing stock during the four appraisal reference months. We only had average data at our disposal for the entire reference years (1995, 1999, 2003, and 2005), so we used these instead.

Figure 8 shows the value-weighted SPAR indexes for those strata with the lowest and highest price index numbers in December 2007: apartments in the province of Zeeland (stratum A) and detached houses in Utrecht (stratum B), respectively. Given the small number of observations per stratum, in particular for stratum A, the volatile pattern of the stratum indexes should not come as a big surprise. Figure 8 also confirms that house price growth can differ greatly across strata. This finding, together with the fact that the impact of stratification is almost negligible, suggests that sample selection bias plays a minor role in the SPAR indexes. Note that for small strata, for example stratum A in Figure 8, the value-weighted SPAR index has an upward bias, as explained by equation (17). Thus, a part of the difference between the stratified and the ordinary indexes might be the result of such 'technical' bias instead of sample selection bias.
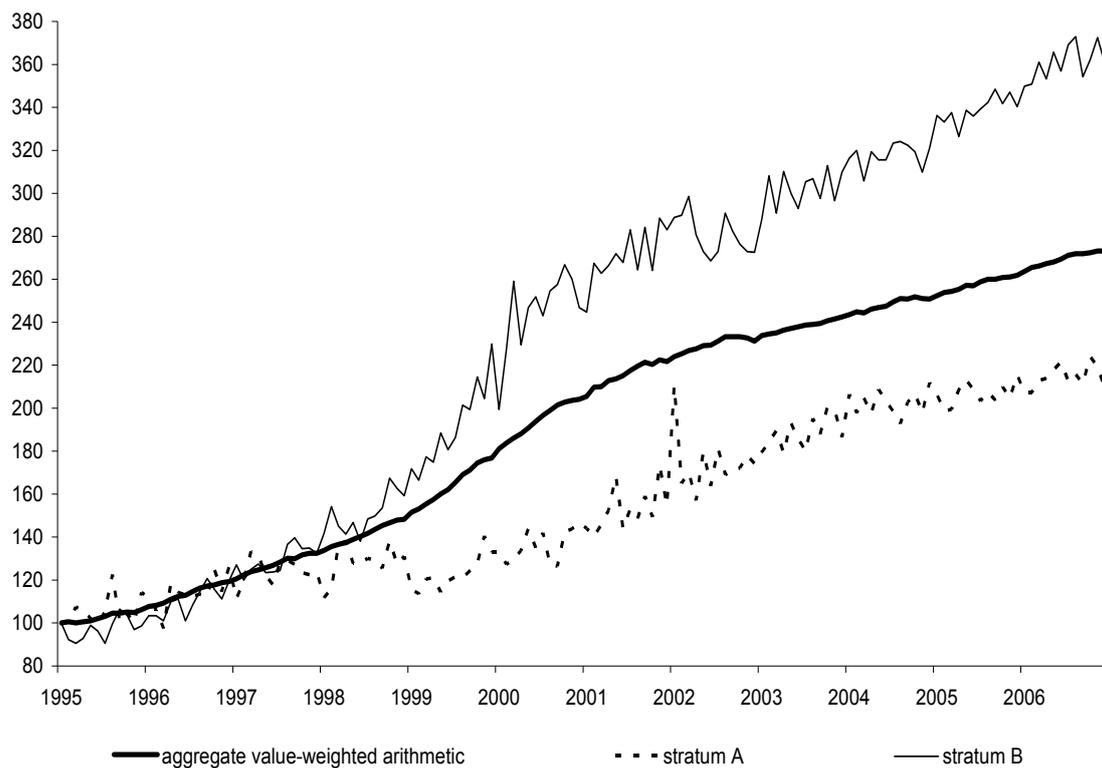

Figure 8. Value-weighted SPAR price indexes (January 1995= 100)

11. Conclusion

In this paper we have reviewed the SPAR approach to measuring house price growth from a statistical and an index number perspective. Just like the repeat-sales method, the SPAR method is an alternative to hedonic regression when data on the characteristics of the houses are unavailable. In their standard form, both methods have at least two things in common: they are solely based on price changes of matched pairs and thus adjust for compositional change, and they do not adjust for quality changes of the houses. From a practitioner's point of view the simplicity and transparency of the SPAR method will be a big advantage. This method has been used in New Zealand since the early 1960s and is also applied in Sweden and Denmark. Recent experiences with the SPAR method in Australia are promising as well (Rossini and Kershaw, 2006). Unlike with repeat-sales, SPAR index numbers will not be revised over time. Also, sample selection bias is likely to be smaller for SPAR indexes than for repeat-sales indexes since the latter exclude houses that have been sold only once. Our empirical work indeed suggests that, while house price growth differs greatly across different regions and types of houses, sample selection bias in the SPAR indexes is of minor importance.

SPAR indexes can either be value-weighted or equally weighted. The literature emphasizes that the choice should depend on the aim of the index. In case an equally-weighted or unweighted index is preferred, the geometric variant should be used since the target price index satisfies the time reversal test whereas the arithmetic target index violates this test. For an index that aims at tracking the change in the value of the total housing stock, the weighted arithmetic variant seems a natural choice. Empirical results for the Netherlands from January 1995 to December 2007 show that the value-weighted and unweighted SPAR indexes differ marginally, so that in practice the choice between them might not be very important. The results also confirm that the repeat sales index suffers from upward bias and that the volatility of the repeat sales index is greater than that of the SPAR index. If a statistical agency has a choice between both methods, we would propose the SPAR method, provided that the quality of the appraisal is deemed sufficient.[19]

---

[19] Preliminary results of our research on the SPAR method have been presented at a recent OECD-IMF Workshop on Real Estate Price Indexes (Van der Wal et al., 2006). Diewert (2007), who discusses the main topics that were dealt with at the Workshop, concludes that " …. these assessment based methods are quite satisfactory (and superior to repeat sales methods) if: the assessed values are used for taxation

As from January 2008 Statistics Netherlands and the Dutch land registry office jointly publish monthly value-weighted SPAR price index numbers. Recently, we have estimated standard errors and confidence intervals for the index numbers as well as for monthly and annual index changes.[20] This would make it possible to provide the public with some measure of accuracy. The failure to adjust for quality changes of the houses remains the most important unresolved issue. One topic for further research might be to investigate to what extent an annually-chained SPAR index more or less automatically adjusts for improvements of the houses, given the fact that most properties are appraised retrospectively and should include the effect of improvements carried out since the last appraisal reference month.

12. Appendix: Derivation of Formulas (3), (4), (6), (7) and (17)

In this Appendix equations (3), (4), (6), (7) and (17) in the main text will be derived. We start by introducing some simplifying notation. Our aim is to estimate the ratio $R = y/x$ of the population variables y and x using the ratio $\hat{R} = \hat{y}/\hat{x}$, where $\hat{y}$ and $\hat{x}$ are (not necessarily unbiased) estimators of y and x.

Given that $\text{cov}(\hat{y}, 1/\hat{x}) = E(\hat{y}/\hat{x}) - E(\hat{y})E(1/\hat{x})$, the expected value $E(\hat{R})$ can be written as

$$E(\hat{R}) = E\left(\frac{\hat{y}}{\hat{x}}\right) = E(\hat{y})E\left(\frac{1}{\hat{x}}\right) + \text{cov}\left(\hat{y}, \frac{1}{\hat{x}}\right). \tag{A.1}$$

Using a second-order Taylor series expansion around the point $E(\hat{x})$, $1/\hat{x}$ can be approximated by

$$\frac{1}{\hat{x}} \cong \frac{1}{E(\hat{x})}\left[1 - \frac{\hat{x} - E(\hat{x})}{E(\hat{x})} + \left(\frac{\hat{x} - E(\hat{x})}{E(\hat{x})}\right)^2\right] = \frac{1}{E(\hat{x})}\left[1 + \frac{\hat{x}^2 - E(\hat{x})\hat{x}}{[E(\hat{x})]^2}\right]. \tag{A.2}$$

Thus, the expected value of (A.2) is

$$E\left(\frac{1}{\hat{x}}\right) \cong \frac{1}{E(\hat{x})}\left[1 + \frac{E(\hat{x}^2) - [E(\hat{x})]^2}{[E(\hat{x})]^2}\right] = \frac{1}{E(\hat{x})}\left[1 + \frac{\text{var}(\hat{x})}{[E(\hat{x})]^2}\right].$$ (A.3)

Using a first-order Taylor series expansion the covariance term in expression (A.1) can be approximated by (see e.g. Kendall and Stuart, 1976)

$$\text{cov}\left(\hat{y}, \frac{1}{\hat{x}}\right) \cong -\frac{\text{cov}(\hat{y}, \hat{x})}{[E(\hat{x})]^2}.$$ (A.4)

Substituting (A.3) and (A.4) into (A.1) yields (compare Cochran, 1977, p. 161)

$$E(\hat{R}) \cong \frac{E(\hat{y})}{E(\hat{x})}\left[1 + \frac{\text{var}(\hat{x})}{[E(\hat{x})]^2}\right] - \frac{\text{cov}(\hat{y}, \hat{x})}{[E(\hat{x})]^2}.$$ (A.5)

The variance of $\hat{R}$ can be approximated using a first-order Taylor series expansion by

$$\text{var}(\hat{R}) = \text{var}\left(\frac{\hat{y}}{\hat{x}}\right) \cong \left[\frac{E(\hat{y})}{E(\hat{x})}\right]^2 \left[\frac{\text{var}(\hat{y})}{[E(\hat{y})]^2} + \frac{\text{var}(\hat{x})}{[E(\hat{x})]^2} - 2\frac{\text{cov}(\hat{y}, \hat{x})}{E(\hat{y})E(\hat{x})}\right].$$ (A.6)

If $\hat{y}$ and $\hat{x}$ are based on independently drawn samples then $\text{cov}(\hat{y}, \hat{x}) = 0$, so that (A.5) and (A.6) reduce to

$$E(\hat{R}) \cong \frac{E(\hat{y})}{E(\hat{x})}\left[1 + \frac{\text{var}(\hat{x})}{[E(\hat{x})]^2}\right];$$ (A.5')

$$\text{var}(\hat{R}) \cong \left[\frac{E(\hat{y})}{E(\hat{x})}\right]^2 \left[\frac{\text{var}(\hat{y})}{[E(\hat{y})]^2} + \frac{\text{var}(\hat{x})}{[E(\hat{x})]^2}\right].$$ (A.6')

Formulas (3) and (4) are obtained by assuming that S^0 and S^t are independently drawn random samples and choosing $\hat{y} = \bar{p}^t(S^t)$ and $\hat{x} = \bar{p}^0(S^0)$ in (A.5') and (A.6'). In this case we know that $E(\hat{y}) = \bar{p}^t(U^0)$ and $E(\hat{x}) = \bar{p}^0(U^0)$. Formulas (6) and (7), in which a single sample appears, are found by choosing $\hat{y} = \bar{p}^t(S^t)$ and $\hat{x} = \bar{p}^0(S^t)$ with the same expected values. The covariance term in equations (A.5) and (A.6) now differs from 0; usually this term will be positive.

The derivation of formula (17) is rather complex since the value-weighted SPAR index is a ratio of two ratios and two Taylor series approximations must be applied. We start from (A.5') where $\hat{y} = \bar{p}^t(S^t)/\bar{a}^0(S^t)$ and $\hat{x} = \bar{p}^0(S^0)/\bar{a}^0(S^0)$. There is small

sample bias, hence $E(\hat{y}) \cong E(\bar{p}^t(S^t))/E(\bar{a}^0(S^t))$ and $E(\hat{x}) \cong E(\bar{p}^0(S^0))/E(\bar{a}^0(S^0))$. To work out the term $\text{var}(\hat{x})/[E(\hat{x})]^2$ we apply (A.6):

$$\frac{\text{var}(\hat{x})}{[E(\hat{x})]^2} \cong \frac{\text{var}(\bar{p}^0(S^0))}{[E(\bar{p}^0(S^0))]^2} + \frac{\text{var}(\bar{a}^0(S^0))}{[E(\bar{a}^0(S^0))]^2} - 2\frac{\text{cov}(\bar{p}^0(S^0),\bar{a}^0(S^0))}{E(\bar{p}^0(S^0))E(\bar{a}^0(S^0))}$$

$$= \left[cv(\bar{p}^0(S^0))\right]^2 \left[1 + \left[\frac{cv(\bar{a}^0(S^0))}{cv(\bar{p}^0(S^0))}\right]^2 - 2\frac{cv(\bar{a}^0(S^0))}{cv(\bar{p}^0(S^0))}\rho\right], \qquad \text{(A.7)}$$

where $\rho = \text{cov}(\bar{p}^0(S^0),\bar{a}^0(S^0))/[\text{var}(\bar{p}^0(S^0))\,\text{var}(\bar{a}^0(S^0))]^{1/2}$ denotes the correlation coefficient of the base period average prices and appraisals $\bar{p}^0(S^0)$ and $\bar{a}^0(S^0)$; the coefficients of variation are defined as $cv(\bar{p}^0(S^0)) = [\text{var}(\bar{p}^0(S^0))]^{1/2}/E(\bar{p}^0(S^0))$ and $cv(\bar{a}^0(S^0)) = [\text{var}(\bar{a}^0(S^0))]^{1/2}/E(\bar{a}^0(S^0))$. The relative variation in the sale prices in S^0 is likely to be of the same order of magnitude as the variation in the appraisals. Thus we expect the coefficients of variation to be approximately equal to each other. Using $cv(\bar{a}^0(S^0))/cv(\bar{p}^0(S^0)) \cong 1$ expression (A.7) reduces to

$$\frac{\text{var}(\hat{x})}{[E(\hat{x})]^2} \cong 2[cv(\bar{p}^0(S^0))]^2(1-\rho). \qquad \text{(A.7')}$$

Substituting (A.7') and the expressions for $E(\hat{y})$ and $E(\hat{x})$ into (A.5') yields

$$E(\hat{P}^t_{W(S)}) \cong \frac{E(\bar{p}^t(S^t))/E(\bar{a}^0(S^t))}{E(\bar{p}^0(S^0))/E(\bar{a}^0(S^0))}\left[1 + 2[cv(\bar{p}^0(S^0))]^2(1-\rho)\right]. \qquad \text{(A.8)}$$

Formula (17) follows from equation (A.8) assuming that S^0 are S^t random samples from the (fixed) population U^0. In that case $E(\bar{a}^0(S^t)) = E(\bar{a}^0(S^0)) = \bar{a}^0(U^0)$, so that $E(\bar{a}^0(S^0))/E(\bar{a}^0(S^t)) = 1$, $E(\bar{p}^0(S^0)) = \bar{p}^0(U^0)$, and $E(\bar{p}^t(S^t)) = \bar{p}^t(U^0)$.

13. References

Australian Bureau of Statistics (ABS) (2005), Renovating the Established House Price Index, Information Paper no. 6417.0.

Bailey, H.J., R.F. Muth and H.O. Nourse (1963), A Regression Method for Real Estate Price Index Construction, *Journal of the American Statistical Association*, 58, 933-942.

Balk, B.M. (1995), Axiomatic Price Index Theory: A Survey, *International Statistical Review*, 63, 69-93.

Balk, B.M. (2005), Price Indexes for Elementary Aggregates: The Sampling Approach, *Journal of Official Statistics*, 21, 675-699.

Bourassa, S.C., M. Hoesli and J. Sun (2006), A Simple Alternative House Price Index Method, *Journal of Housing Economics*, 15, 80-97.

Butler, J.S., Y. Chang and A. Crews Cutts (2005), Revision Bias in Repeat-Sales Home Price Indices, Freddie Mac Working Paper no. 05-03.

Case, K.E. and R.J. Shiller (1987), Prices of Single-Family Homes Since 1970: New Indexes for Four Cities, *New England Economic Review*, 45-56.

Case, B., H.O. Pollakowski and S.M. Wachter (1997), Frequency of Transaction and House Price Modeling, *Journal of Real Estate Finance and Economics*, 14, 173-187.

Clapham, E., P. Englund, J.M. Quigley and C.L. Redfearn (2004), Revisiting the Past: Revision in Repeat Sales and Hedonic Indexes of House Prices, Revised version of a paper presented at the May Meetings of the Homer Hoyt Institute, West Palm Beach, May 2003.

Clapp, J.M. and C. Giacotto (1999), Revisions in Repeat-Sales Price Indexes: Here Today, Gone Tomorrow?, *Real Estate Economics*, 27, 79-104.

Clements, K.W., I.H.Y. Izan and E.A. Selvanathan (2006), Stochastic Index Numbers: A Review, *International Statistical Review*, 74, 235-270.

Cochran, W. (1977), *Sampling Techniques* (3d edition). New York: Wiley & Sons.

Diewert, W.E. (1995), On the Stochastic Approach to Index Numbers, Discussion Paper no. 95-31, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (2003), The Treatment of Owner Occupied Housing and Other Durables in a Consumer Price Index, Discussion Paper no. 03-08, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (2007), The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Future Directions, Discussion Paper no. 07-01, Department of Economics, University of British Columbia, Vancouver, Canada.

Goldberger, A.S. (1968), The Interpretation and Estimation of Cobb-Douglas Functions, *Econometrica*, 36, 464-472.

Haan, J. de (2004a), Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement, *Journal of Economic and Social Measurement*, 29, 427-443.

Haan, J. de (2004b), Hedonic Regression: The Time Dummy Index As a Special Case of the Imputation Törnqvist Index, Paper presented at the eighth meeting of the Ottawa Group, 23-25 August 2004, Helsinki.

Haan, J. de (2007a), Formules voor de variantie van (mutaties in) de SPAR index, Unpublished Paper, Statistics Netherlands, Voorburg (Dutch only).

Haan, J. de (2007b), Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Other Approaches, Paper presented at the Seventh EMG Workshop, 12-14 December, Sydney.

Hansen, J. (2006), Australian House Prices: A Comparison of Hedonic and Repeat-Sales Measures, Research Discussion Paper, Reserve Bank of Australia.

Hill, R. (2004), Constructing Price Indexes Across Countries and Time: The Case of the European Union, *American Economic Review*, 94, 1379-1410.

Hill, R.J. and D. Melser (2007), The Hedonic Imputation Method and the Price Index Problem: The Case of House Price Indexes, Discussion Paper no. 2007/04, School of Economics, University of New South Wales, Sydney.

Hwang, M. and J. Quigley (2004), Selectivity, Quality Adjustment and Mean Reversion in the Measurement of House Values, *Journal of Real Estate Finance and Economics*, 28, 161-178.

International Labour Office (ILO) (2004), *Consumer Price Index Manual: Theory and Practice*. Geneva: ILO Publications.

Jansen, S., P. de Vries, H. Coolen, C. Lamain and P. Boelhouwer (2006), Developing a House Price Index for the Netherlands: A Practical Application of Weighted Repeat Sales, Paper presented at the ENHR Housing Economics Workshop, 16-17 February 2006, Copenhagen.

Kendall, M. and A. Stuart (1976), *The Advanced Theory of Statistics* (Fourth edition). London: Charles Griffin & Company.

Makaronidis, A. and K. Hayes (2006), Owner-Occupied Housing for the HICP, Paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, 6-7 November 2006, Paris.

Rossini, P. and P. Kershaw (2006), Developing a Weekly Residential Price Index Using the Sales Price Appraisal Ratio, Paper presented at the twelfth Annual Pacific-Rim Real Estate Society Conference, 22-25 January 2006, Auckland.

Särndal, C.E., B. Swensson and J. Wretman (1992), *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Thibodeau, T.G. (ed.) (1997), *House Price Indices*, Special Issue of the *Journal of Real Estate Finance and Economics*, 14. Boston/Dordrecht/Londen: Kluwer.

Triplett, J.E. (2004), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*. STI Working Paper 2004/9, OECD, Paris.

Vries, P. de, S. Jansen, P. Boelhouwer, H. Coolen, C. Lamain, E. van der Wal and D. ter Steege (2006), Samenhang tussen marktprijs en WOZ-waarde, *Tijdschrift voor de Volkshuisvesting* , 44-49 (Dutch only).

Vries, P. de, J. de Haan, G. Mariën and E. van der Wal (2008), A House Price Index Based on the SPAR Method, Paper presented at the World Congress on National Accounts and Economic Performance Measures for Nations, 12-17 May 2008, Arlington.

Wal, E. van der, D. ter Steege and B. Kroese (2006), Two Ways to Construct a House Price Index for the Netherlands: The Repeat Sales and the Sale Price Appraisal Ratio, Paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, 6-7 November 2006, Paris.

Wang, F.T. and P.M. Zorn (1997), Estimating House Price Growth with Repeat Sales Data: What's the Aim of the Game?, *Journal of Housing Economics*, 6, 93-118.