A House Price Index Based on the SPAR Method

Draft, May 2, 2008

Paul de Vries
Delft University of Technology
OTB Research Institute for Housing, Urban and Mobility Studies
The Netherlands
p.devries@tudelft.nl

Jan de Haan
Statistics Netherlands
Division of Macro-Economic Statistics and Dissemination
The Netherlands
jhhn.cbs.nl

Gust Mariën
Delft University of Technology
OTB Research Institute for Housing, Urban and Mobility Studies
The Netherlands
a.a.a.marien@tudelft.nl

Erna van der Wal
Statistics Netherlands
Division of Macro-Economic Statistics and Dissemination
The Netherlands
eral@cbs.nl

Abstract
The Sale Price Appraisal Ratio (SPAR) method to construct house price indexes makes use of the ratios of transaction prices and previous appraisal values. It has been applied in New Zealand, Sweden and Denmark. This paper describes the SPAR method, compares it with repeat sales and hedonic methods and assesses the reliability of appraisal values in the Netherlands. Empirical results covering a ten-year period are also presented. Since the SPAR method performs well compared with the more traditional methods, government agencies elsewhere might consider it when developing house price indexes, provided that reliable appraisals are available.

Key words: Appraisal Value, Hedonics, House Price Index, Repeat Sales.

1. Introduction

The measurement of house price changes is a challenging area for researchers, statistical agencies and financial organisations because homes constitute a significant proportion of household wealth and can have important ramifications for the economy. While academic researchers use house price indexes to gain a better understanding of the housing market, European statistical agencies are participating in a Eurostat project on the measurement of house price indexes and the treatment of owner-occupied housing in the Harmonised Index of Consumer Prices (HICP).[1] Furthermore, the Dutch Central Bank requires that financial institutions specify their risks by estimating the actual liquidation value for each dwelling in their mortgage portfolio.

In this paper we focus on the construction of an index that aims at measuring price changes across the entire stock of owner-occupied houses. This, in itself, implies a certain advantage since a price index of the housing stock is presumably what the general public understands by a house price index. It is the market value of the property that should be observed, which is defined as the price a house is actually sold for. The vast majority of existing properties is not sold in a certain month or quarter; their market values must be estimated. The market value of a house, whether observed or estimated, is usually called the ‘price’. In view of the lack of data, any house price index will in practice be based on sales of second-hand dwellings.

In 2004 the land registry office [*Kadaster*] asked the OTB Research Institute for Housing, Urban and Mobility Studies (Delft University of Technology) to develop a price index for owner-occupied dwellings. A set of 55 monthly indexes was introduced in May 2005, consisting of an overall index, four regional indexes and indexes for combinations of region and type of dwelling. These indexes (Jansen *et al*, 2006) were estimated with a weighted repeat sales approach (Case and Shiller, 1987; Abraham and Schauman, 1991; Calhoun, 1996), which is a modification of the repeat sales method originally developed by Bailey *et al.* (1963).

According to Bailey *et al*. (1963) the repeat sales method is more efficient than alternative methods because it utilises information on prices pertaining to earlier periods and includes it in selling prices in later periods. However, there are certain drawbacks which, if left unresolved, make these indexes unsuitable for official statistics or as input for the HICP. The most serious drawback is revision: past index numbers are revised by present-day information (Baroni, 2004). In other words, additional sales reverberate on the index numbers because new pairs provide information on changes in house prices that goes beyond the information obtained from the sample.

These drawbacks are also mentioned by Bourassa *et al.* (2006), who conclude that historical index numbers should not require revision when data are added for subsequent periods. They present the Sale Price Appraisal Ratio (SPAR) method as an alternative to

[1] HICPs are consumer price indexes produced in EU Member States on the basis of harmonised standards. They are used in particular for assessing price convergence, monitoring inflation and conducting monetary policy in the euro zone. For a discussion on alternative methods to incorporate owner-occupied housing into a CPI (or HICP), see Diewert (2003).

hedonic or repeat sales indexes. Like the repeat sales method, the SPAR method is based on matched pairs, but, in contrast, it uses (nearly) all the price data that is available for the period under observation. Since the majority of the houses sold during the observation period were not sold during the index reference or base period, there is a general shortage of transaction prices for the base period. The base period prices are therefore estimated using official government appraisals of the dwellings. In the Netherlands appraisals are collected under the Real Estate Law [*Wet Waardering Onroerende Zaken*]. In contrast with a repeat sales index, the SPAR index is not revised when data for new periods are added. Bourassa *et al.* (2006) "maintain that the advantages and the relatively limited drawbacks of the SPAR model make it an ideal candidate for use by government agencies in developing house price indexes".[2]

Indexes can be equally weighted or value weighted. A value-weighted price index explicitly or implicitly weights the indexes of individual dwellings by their base period prices (values). The literature stresses that the choice between an equally-weighted and a value-weighted index should depend on the aim of the index (Wang and Zorn, 1997). The aim of the Central Bank and financial institutions is to have an index relating to a 'mean dwelling', which is a good match for an unweighted geometric mean index arising from a standard repeat sales approach. For an index that should track the changes in the value of the total housing stock, the weighted arithmetic variant seems a natural choice. *Statistics Netherlands* aims at an acquisitions-based index (that approximates the value change of the housing stock), for which the value-weighted SPAR index seems a suitable candidate.

This paper compares the SPAR and the repeat sales methods. Notwithstanding the differences between these methods, the intention was to develop a single index. A simple hedonic approach is also included in the analysis to ensure that an adequate conclusion could be reached on the issue of choice of method for measuring house price changes.[3] The paper is organised as follows. Section 2 provides a brief overview of the literature on hedonic and repeat sales approaches and presents background information on the SPAR method. Section 3 discusses why in the Netherlands individual property appraisals can be used for constructing the SPAR index and presents empirical evidence. In Section 4 we compare hedonic, repeat sales and SPAR price index numbers covering a ten-year period. Section 5 concludes.

2. Three Approaches

A key starting point for a house price index is a fixed housing stock. However, the quality of the housing stock is likely to rise over time as a result of newly built homes. More importantly in the short run, houses are sold infrequently and the composition, or 'quality

[2] Unfortunately the authors present the SPAR method without referring to index number issues. In earlier work (De Haan *et al.*, 2006) we addressed questions such as: Is a SPAR index aiming at a price index of the housing stock or only a transaction index? Given the target or population price index, what assumptions are needed to derive the SPAR estimator? What are its statistical properties?

[3] It should be stressed that we cannot estimate a functionally correct hedonic model because of the limited number of quality characteristics included in the data set.

mix', of the properties sold will vary substantially from period to period. This introduces bias in simple index measures such as the mean or median price. For example, if in some period a disproportionate number of high-priced houses were sold, the mean or median price would rise, even though not a single house had increased in value (Case and Shiller, 1987). These shortcomings led to the development of alternative methods, in particular to hedonic and repeat sales methods. Bourassa *et al.* (2006) proposed the SPAR (Sale Price Appraisal Ratio) method as another alternative. This section provides a brief review of the hedonic and repeat sales models and presents background information on the SPAR approach.

2.1 Hedonics

Hedonic regression models are widely used today in housing market research (Mason and Quickley, 1996).[4] A hedonic model expresses the price of house *i* in period *t*, denoted P_{it}, as a function of a set of physical (and possibly other) characteristics, Q_i, and time *t*:

$$P_{it} = f(Q_i, t)\,. \tag{1}$$

The hedonic coefficients can be interpreted as shadow prices which reflect the value of a characteristic. For example, an extra room will increase the value of a house by a specific amount. It is essential to specify the correct functional form and to include the correct set of quality characteristics (see Wang and Zorn, 1997). Bailey *et al.* (1963), who laid the foundations for the repeat sales method, reported that quality characteristics are numerous and difficult to measure, and concluded that hedonic models may not yield useful results. Mason and Quigley (1996) argued that the existence of sub-markets might go some way to explaining why the standard hedonic specification may not work. They further argued that the functional form assumption is particularly awkward in a housing context because the hedonic price function summarises not only consumer preferences and production technologies but also various quantities that are historically determined, hard to measure, and inaccessible to economic theory (see also De Vries and Boelhouwer, 2005). Despite these drawbacks, researchers have examined numerous datasets and model specifications to determine the marginal effect of housing characteristics on house prices and construct quality-adjusted house price indexes; see for a recent review Sirmans, Macpherson and Zietz (2005).

In the Netherlands, data on all second-hand houses sold are recorded by *Kadaster*, the land registry office. However, no details are registered on house characteristics other than built surface area and dwelling type (detached house, corner house, terraced house, semi-detached house). This prevents the use of hedonics for constructing quality-adjusted house price indexes.

---

[4] Colwell and Dilmore (1999) have suggested that it was not Court (in 1939), but Wallace (in 1926) or Haas (in 1922) who first estimated a hedonic model. Before it was applied in housing market research, hedonic regression has been used to estimate quality-adjusted price indexes for capital goods and consumer goods like cars (see for example Griliches, 1971).

2.2 Repeat Sales

As the weighted repeat sales model is extensively addressed in the literature (see Bailey, Muth and Nourse, 1963; Case and Shiller,1987, 1989; Goetzmann, 1992; Calhoun, 1996; Drieman and Pennington-Cross, 2004), we make do with a brief description here. Bailey *et al.* (1963) were the first to develop a house price index based on the repeat sales model. Essentially, the repeat sales approach uses the prices paid for single properties at different points in time to estimate a vector of figures that 'best' explains the observed changes in price over the sample period (Abraham and Schauman, 1991). The repeat sales approach models the price changes of houses that are repeatedly sold. In particular, it expresses the logarithm of the ratio R_i of the house price P_{is_2} in the second sale period s_2 and the price P_{is_1} in the initial or first period s_1 ($s_1 < s_2$) as

$$\ln(R_i) = \ln(P_{is_2} / P_{is_1}) = f(D_i, t)\,. \qquad (2)$$

D is a set of time dummy variables. For the first sale of a particular house it has the value -1, for the second sale it has the value +1. All other dummies have the value 0.

Case and Shiller (1987) proposed the weighted repeat sales method, an adapted version of the unweighted method described by Bailey *et al*. (1963). They argued that the longer the holding period becomes, the greater the variance in individual house price appreciation will be. This type of heteroscedasticity may undermine the efficiency of the repeat sales index (Wang and Zorn, 1997). Calhoun (1996) distinguishes three stages in the estimation of the weighted repeat sales model. In the first stage the original model of Bailey *et al.* is calculated. The second and third stages aim to improve the efficiency of the first-stage parameter estimates, accounting for the possibility that the estimation error is positively related to the time interval between subsequent transactions. In the second stage the squared error from the first stage is regressed on the (squared) time interval.

Jansen *et al.* (2007) found that heteroscedasticity was of very little importance in their Dutch dataset; the amount of explained variance is less than one percent. They also encountered a problem with the weights necessary to correct for heteroscedasticity. Thus, Jansen *et al.* (2007) argue that the original repeat sales method of Bailey *et al*. (1963) is more adequate for calculating a house price index in the Netherlands than its weighted counterpart.

2.3 SPAR

The Sale Price Appraisal Ratio (SPAR) method, which has been applied in New Zealand since the early 1960s, was proposed by Bourassa *et al.* (2006) as an alternative approach to measuring house price indexes. Like repeat sales methods, the SPAR method is based on matched pairs but, in contrast, uses (nearly) all the price data that is available for the period of observation. Since the vast majority of houses that are sold during the current period were not sold during the index reference or base period, there is a general shortage of transaction prices for the base period. The base period prices are therefore estimated with official government appraisals of the properties. De Haan *et al.* (2006) indicated that there are various types of SPAR indexes. SPAR indexes can be either value weighted or equally (or unweighted) index. If an equally-weighted index is preferred, the geometric variant would be the best choice. For an index that tracks the changes in the value of the total housing stock, the weighted arithmetic variant seems a natural choice. Equation 3 shows the arithmetic value-weighted SPAR index:

$$I_{SPAR,t} = \frac{\sum_{i=1}^{n_t} P_{it} \Big/ \sum_{i=1}^{n_t} A_{i0}}{\sum_{i=1}^{n_0} P_{i0} \Big/ \sum_{i=1}^{n_0} A_{i0}}, \tag{3}$$

where P_{it} and P_{i0} are the transaction prices in the current period (t) and the period (0) in which the houses were appraised; A_{i0} is the appraisal value. Note that the denominator of equation (3) vanishes if in period 0, which is the base period of the index, the appraisals would be equal to the transaction prices. The index given by (3) is a value-weighted price index because the price indexes of the individual houses are implicitly weighted by their base period prices (appraisal values).[5]

In the Netherlands the official appraisals are collected under the Real Estate Law [*Wet Waardering Onroerende Zaken*]. In the real estate literature there is some discussion about Appraisal Values (Geltner, 1991; Edelstein and Quan, 2006; Leventis, 2006). Most studies are based on appraisal values of dwellings that are about to be re-financed. That is why, in general, their findings suggest that appraisals tend to be positively biased; they tend to over-predict the selling price of the property (Leventis, 2006). In the Netherlands (and also in Denmark and New Zealand) appraisals are not collected for re-financing but for tax purposes. Dutch households pay local tax according to the value of their dwelling. Households who feel that the appraisal value is too high may lodge an appeal. Though legally the appraisal value should reflect the market value of the house, we expected local authorities to underestimate it in order to avoid court procedures. So initially we assumed that the appraisal value would tend to be negatively biased, i.e. to underpredict the market value of the property. However, subsequent analysis proved us wrong (see Van der Wal, 2006, and De Vries, 2007).

3. Representativity of the Data

To estimate hedonic and repeat sales house price indexes, we have used the dataset of the Dutch Land Registry Office [*Kadaster*]. We call this dataset the 'transaction dataset'. To compute SPAR index numbers, in addition we have used the appraisal values from the

[5] The denominator of (3) is needed to make the index equal to 1 in the base period 0. Put differently, it is a constant scaling factor. It can alternatively be interpreted as a factor that corrects the numerator for possible over-estimation or under-estimation of the appraisals with respect to the transaction prices. The underlying assumption is that a linear relation through the origin exists in the base period between appraisal values and transaction prices. See also Section 3.

municipalities. We will call this dataset the 'appraisal dataset'. This, of course, raises the question whether the quality of the appraisals is satisfactory. Before explaining how the appraisals were determined and presenting evidence on their reliability, we first describe the transaction dataset.

3.1 Transaction Dataset

The transaction dataset contains data on more than 2.5 million individual transactions regarding owner-occupied dwellings between January 1993 and December 2006, which amounts to 67% of the owner-occupied housing stock. A number of transactions have been removed for reasons of validity. We applied price limits between a thousand and five million euros. Dwellings with a very large surface area (over 1,000 square meters) and dwellings that were sold more than twice in the same month were excluded. We also omitted dwellings that were sold to the sitting tenant. The latter accounted for 9% of all transactions between January 1993 and December 2006 and were sold at prices that were ten to fifteen percent below the appraisal value. Eventually, we were left with 2.2 million transactions for the hedonic and SPAR indexes. For the repeat sales index, only dwellings sold twice or more could obviously be used.

The literature on repeat sales suggests that sales with a short time interval imply an 'unusual' transaction (Englund *et al*, 1998; Steele and Goy, 1997; Clapp and Giacotto, 1999). They may be distressed sales arising from, for example, divorce or job loss or they could be speculative transactions. In the Netherlands no conveyance tax needs to be paid on a house that is resold within six months. Jansen *et al.* (2006) showed that a number of speculative sales took place in the boom between 1998 and 2001. Clapp and Giacotto (1999) and Steel and Goy (1997) suggest eliminating very short holds from the dataset. Jansen *et al.* (2006) explored the potential impact of very short holds by calculating the monthly growth rate for each house sold. The mean growth rates were 8.2%, 5.2%, 1.2%, and 0.9% for houses sold within six months, within twelve months, within all periods, and between twelve months and the end of the period, respectively. Thus, homes resold within twelve months realise, on average, a very high increase in value per month, which can bias the index. After removing transactions that took place within twelve months, we had around 737,000 pairs of repeat sales in the dataset.

3.2 Appraisal Dataset

To construct an index based on the SPAR method, appraisals are also needed. The Dutch municipalities are legally obliged to have an up-to-date estimate of the value of each real estate object on the first of January 1995, 1999, 2003, 2005 and 2007. As from January 2007, the appraisal values will be published one year after the valuation date. The records need continuous updating to be correct and complete. The entire process is monitored on the government's behalf by the Council for Real Estate Assessment [*Waarderingskamer*]. There is no prescribed method of appraisal. Nowadays, all municipalities appraise the objects with (hedonic) valuation models in combination with visual inspections and local market information.

For this study data apparaisel data for 2005 and 2007 was not yet available. Thus we distinguish between three appraisal periods when computing SPAR index numbers. During the first period (January 1993 - December 1998) the valuation must be indexed to the reference date of January 1995, during the second period (January 1999 - December 2002) to January 1999 and during the last period (January 2003 – present) to January 2003.

The appraisals for a particular year are worked out in the future. For example, the appraisal values for January 2003 were determined at the beginning of 2005. The values thus include the results of any appeals lodged by home-owners. They also include home improvements carried out between the date of valuation (January 2003) and the date upon which the house was accorded an official value (early 2005). Though the SPAR method, since it is based on matched pairs, controls for changes in the quality mix of the sample, it does not control for qualitative changes of the individual houses. The same goes for the repeat sales method. It has been suggested that we adjust the valuations to take account of home improvements that require planning permission. Unfortunately, in the Netherlands such adjustments are infeasible because planning permission data are available only at the aggregate (project) level and not for individual dwellings.

Reliable appraisal values are essential if the SPAR method is to work effectively. For privacy reasons we are not allowed to publish these kind of research findings without explicit permission from every municipality – it is they who own the appraisals. For this study all municipalities in the Province of Overijssel, except Hengelo and Dinkelland, granted us permission to publish the results. Unpublished research has shown that our results are representative for the Netherlands as a whole.

The problem when comparing the current house price P_i with the appraisal value A_{i0} is the difference in observation period. We therefore computed a 'real' house price, rP_{i0}, using the (repeat sales) House Price Index (*HPI*) which was published by *Kadaster* at that time:

$$rP_{i0} = (HPI_0 / HPI_t)P_{it} . \tag{4}$$

If the appraisals accurately represent the market value in period 0, the relationship will be linear with a zero intercept term. The scatterplot in Figure 1, which is based on data of January 2003 for the Province of Overijssel, shows the coherence between sale prices and appraisals. The linear regression line nearly crosses the origin; the value of R^2 is 0.95. For 1995 and 1999 the R^2 values are 0.90 and 0.94.

< *Figure 1 about here* >

We conducted another simple but efficient comparison of the real house price and the appraisal values, again for January 1995, 1999, and 2001. The percentual differences between the mean appraisal and the mean real house price declined over time, indicating that the reliability of the appraisals has improved. The decrease in the standard deviation endorses these findings. In the first period, the appraisal value underestimated the price by more than four percent on average. In the next period, starting in 1999, the appraisal values overestimated the sale prices, but there was a significant decrease in the absolute difference between the transaction prices and the appraisals and in the standard deviation as well. The same trend is observable in the last period.

< Table 1 about here >

We also compared the real house price and the appraisal value using Cohen's *d*, which provides insight into the relative size of the difference; this is in contrast with the R^2, which provides insight into the coherence. Cohen's *d* is the appropriate effect size measure in the context of a t-test on means. It is defined as the difference between two means divided by their pooled standard deviation. Hence,

$$d = \frac{mean_1 - mean_2}{\sqrt{(SD_1^2 + SD_2^2)/2}}, \qquad (5)$$

where $mean_k$ and SD_k are the mean and the standard deviation for group *k* (for *k* = 1, 2). In general, *d*= 0.2 indicates a small effect size, *d*= 0.5 a medium effect size and *d*= 0.8 a large effect size. The greater Cohen's *d*, the greater the difference between the appraisal A_{i0} and the 'true' market value rP_{i0}. If *d* was lower than 0.2, we assumed the existence of a significant relationship.

As the municipalities are responsible for calculating the appraisals, we analysed Cohen's *d* for each type of dwelling, municipality and appraisal decile. Table 2 shows that Cohen's *d* is below 0.2 for all three appraisal dates for the Province of Overijssel, which indicates that there is no significant difference between the appraisals and (real) house prices. At the detailed levels, we observe that in 2003 the appraisal value deviates from the house price ($d > 0.20$) only for the lowest two appraisal value deciles. These market segments are more difficult to estimate, so differences materialise sooner. On the whole, the municipalities seem to be getting better at working out the appraisal values. In 1995, 71% of Cohen's *d* was above 0.2, whereas in 2003 there was no longer a significant difference in any municipality between the mean appraisal and the mean house price.

< Table 2 about here >

Finally, we analysed the ratio F_{i0} of the real house price and the appraisal value:

$$F_{i0} = (rP_{i0} / A_{i0}). \qquad (6)$$

In accordance with the principles of the Dutch Real Estate Law, the ratio F_{i0} should be (approximately) equal to 1. Figure 2 depicts the distribution for each appraisal date using twenty classes of equal size on the x axis; the y axis shows the corresponding percentual shares. The two middle staffs (0.95-1.00 and 1.00-1.05) are shaded black to indicate the anticipated mid-point. The graphs show that the distribution becomes increasingly steeper over time, indicating that more and more dwellings acquire a 'normal' fraction. In 1995 F_{i0} was between 0.90 and 1.10 for only 56% of all houses. It rose to 79% in 2003. This shows once again that house prices and appraisal values have drawn closer together. We therefore conclude that the official Dutch appraisals – certainly as of January 2003 – can be used for calculating a house price index based on the SPAR method.

< Figure 2 about here >

4. A Comparison of SPAR, Hedonic and Repeat Sales Index Numbers

4.1 Trends and Fluctuations

We computed SPAR, hedonic and repeat sales indexes for the province of Overijssel. Unpublished research confirmed that our findings are representative for the Netherlands as a whole. For the hedonic and SPAR indexes data on 111,858 individual sales that took place between January 1993 and December 2006 were used, whereas for the repeat sales index we used 32,496 pairs of repeat sales. Note that the hedonic index is a conventional time dummy variable index, estimated by Ordinary Least Squares (OLS) regression. The parameter estimates are listed in the upper part of Table 3. The hedonic results should be interpreted with great caution since there are only very few characteristics included in the model. The reason why we have estimated the hedonic index is merely to show that such a crude method may still perform reasonably well.

Figure 3 displays the three indexes as well as the development of the mean price. Although the trend is very similar, in December 2006 the SPAR index is slightly below the hedonic and repeat sales indexes. The mean price index is obviously a bad measure of house price appreciation as it is affected by compositional changes. Indeed, it fluctuates most heavily. Interestingly, however, its trend hardly differs from that of the three other indexes, which all should, at least to some extent, control for changes in the quality mix. There is a simple relation between the mean price index and the (value-weighted) SPAR index:

$$I_{SPAR,t} = \left[\sum_{i=1}^{n_0} \frac{A_{i0}}{n_0} \Big/ \sum_{i=1}^{n_t} \frac{A_{i0}}{n_t} \right] I_{MEAN,t} \,. \qquad (7)$$

Equation (7) shows that $I_{SPAR,t} < I_{MEAN,t}$ ($I_{SPAR,t} < I_{MEAN,t}$) if the 'sample' of the current period t contains relatively many expensive (cheap) houses – in terms of the appraisal value – as compared to the base period 'sample'. In view of this it is not so surprising that the mean price index fluctuates around the trend of the SPAR index. Although the SPAR index controls for compositional change with respect to the base period (i.e., the appraisal reference month), it does not specifically control for compositional changes on a month-to-month basis.[6] Yet, the hedonic index is more erratic than the SPAR index, which must be caused by the over-simplified specification of the hedonic model. Table 3 shows that considerable monthly fluctuations occur for all indexes. For example, in December 2005 the hedonic index rose by 0.7% whereas the repeat sales index and the SPAR index fell by -0.6% and -0.8%, respectively.

< Figure 3 about here >

< Table 3 about here >

4.2 Precision

The mean square error of an estimator – the square root of the sum of its variance and the squared bias – is an inverse measure of its accuracy: it measures how far the estimator is expected to be from the population target it is aiming at. Here we focus on the variance component, or rather on the square root thereof, the standard error. The standard error is an inverse measure of precision: the greater the standard error of an estimator, the lower its precision is. Using the standard errors (*s.e.*) we can calculate 95%-confidence intervals around the estimated index values with bounds given by $I_{(.)t} \pm 1.96 \times s.e.$ The width of the confidence interval gives an idea of the uncertainty surrounding the estimates. Since the standard error depends on the sample size, a very wide interval may indicate that more data should be collected before drawing any definite conclusions about (changes in) the index numbers. The advantage of hedonic and SPAR indexes is that all transactions can be used; the repeat sales index is, by construction, only based on data of houses that were sold more than twice (after January 1993 in our case). This 'waste of data' is often cited as a drawback of the repeat sales approach.

The standard errors and the corresponding 95%-confidence interval for the SPAR index were approximated with Taylor linearization methods.[7] The confidence interval for the hedonic (OLS time dummy) index is easily calculated: statistical packages like SPSS provide standard errors of the estimated coefficients and generally also the bounds of the confidence interval. The computation of the confidence interval for the geometric repeat sales (RS) index is somewhat more difficult. The index number $I_{RS,t}$ is estimated by

$$I_{RS,t} = \exp(\tilde{\beta}_t) \times 100, \tag{8}$$

where $\tilde{\beta}_t$ is the estimated parameter from a generalised least squares regression analysis. The corresponding standard error is

---

[6] Note again that there are three appraisal reference months: January 1995, January 1999 and January 2003. We constructed a time series for the SPAR index by chaining the various index series during these months.

[7] In short, the numerator and the denominator of equation (3) are stochastically based upon the distribution of houses sold in the base period and in the current period. The variance formulas are obtained by assuming that the sets of houses sold are independently drawn random samples. Furthermore, we assume that the relative distribution of the house price in the base period and current periods equals the distribution of the appraisals (in the base period).

$$\sigma_{I_{RS,t}} = I_{RS,t}\sigma_{\hat{\beta}_t}, \tag{9}$$

where $\sigma_{\hat{\beta}_t}$ pertains to the standard error of the estimated coefficient from the third step of the generalised least squares regression analysis.

Because the magnitude of the standard error depends on the level of the index, a relative measure of precision would be more appropriate. One such measure, $PREC_t$, is obtained by dividing the width of the confidence interval, Wc_t, by the index number and multiplying by 100:

$$PREC_t = (W_{C_t} / I_{(.),t}) \times 100. \tag{10}$$

Figure 4 shows the precision of the three estimated price indexes. The repeat sales index was more precise than the hedonic index across the entire period. However, the SPAR index turned out to be even more precise. Table 3 contains the mean precision, which is calculated as the average of $PREC_t$ for the 168 index numbers.

< Figure 4 about here >

4.3 Cause and Effect

There are various possible explanations for the differences in the indexes. We concentrate here on the repeat sales and SPAR indexes. As mentioned earlier, a hedonic house price index of sufficient quality is infeasible in the Netherlands because no details are recorded on house characteristics apart from built surface area and dwelling type. One key factor is that a repeat sales index is based on dwellings sold twice or more, while the SPAR index is based on all transactions. Clapp and Giacotto (1999) and Jansen *et al.* (2006) observed that properties with short time intervals appreciate at a higher rate than properties with longer time intervals. Our own analysis revealed that the mean house price in the repeat sales dataset was approximately 8% lower than the mean price in the SPAR dataset. This need not affect the development of the index if cheaper houses underwent the same price development as more expensive ones. Our calculations did indicate the existence of such an effect, however. Re-calculation of the SPAR index on the repeat sales dataset showed that the SPAR index comes closer to the repeat sales index.

In addition, a repeat sales index might understate price change compared with the SPAR index because it is a geometric mean of individual appreciation rates, which is less than the arithmetic mean unless all appreciation rates coincide (Wang and Zorn, 1997). This difference was negligible in our dataset: the arithmetic repeat sales index (Shiller, 1991) scarcely deviated at all from the usual geometric repeat sales index and conversely, a geometric SPAR index differed only marginally from the arithmetic SPAR.

Last but not least, we come to the revision of a repeat sales index, which is due to the fact that additional sales reverberate on the index numbers because new pairs provide additional information about price changes beyond that found with the previous data.[8] In other words, present-day information changes the past values of the index (Baroni, 2004). The repeat sales method utilises information about the price index for earlier periods and includes it in selling prices in later periods, so it can be argued that this type of revision is the result of increased efficiency of the estimators (Shiller, 1991). Revision volatility may upset the interpretability of the index, however, as the new index values may not even be similar to the old ones. Clapp and Giacotto (1999) and Jansen *et al.* (2006) observed that properties with only one or two years between sales appreciate at a higher rate than other properties and may therefore be partly responsible for the revision of the index. To obtain an impression of the magnitude of the revision, we calculated the repeat sales index index numbers with all data up to December 2005, and again with all data up to December 2006 (thus with 12 additional months). The results showed that the volatility of the coefficients was relatively small: the average percentage change for the Netherlands was –0.23% and only -0.07% for the Province of Overijssel.

5. Conclusion

This paper looked at ways of measuring house price changes. We briefly reviewed the hedonic and repeat sales methods and examined the SPAR method in more detail. Like the repeat sales method, the SPAR method is an alternative to hedonics when insufficient data is available on the characteristics of dwellings. In their standard form, both methods have at least two things in common: they are based solely on price changes of matched pairs, and thus adjust for compositional change, but they make no adjustment for changes in the quality of individual dwellings. From a practitioner's point of view the simplicity and transparency of the SPAR method can be seen as an advantage. It has been used in New Zealand since the early 1960s and is also applied in Sweden and Denmark. Recent experiences with SPAR in Australia are promising as well (Rossini and Kershaw, 2006). SPAR index numbers are not subject to revision. Moreover, sample selection bias is most likely to be smaller for the SPAR index than for a repeat sales index as the latter excludes houses that have been sold only once.[9] Which index is best also depends on the precision of the estimator and on the statistical target aimed at. *Kadaster*, the Dutch Land Registry Office, initially asked for an unweighted (geometric) index, which is a good match for a repeat sales index. *Statistics Netherlands* aims at an (arithmetic) acquisitions-based index, or an index for the value of the housing stock, for which the value-weighted SPAR is a suitable estimator (De Haan *et al.*).

---

[8] Put differently, these indexes violate temporal fixity. The multi-period time dummy variable hedonic price index is subject to revision for the same reason. This type of hedonic index seems to have dominated the literature. Hill and Melser (2007) argue that the double hedonic imputation method may be a better choice: not only are the characteristics parameters allowed to change over time, this method is less prone to omitted variables bias. It is less efficient,however, since data are not pooled across different periods.

[9] Diewert (2007) concludes that methods that rely on assessment information in the base period and sales information in the current period, like the SPAR method, are superior to repeat sales approaches, provided that the appraisals are estimated for tax purposes.

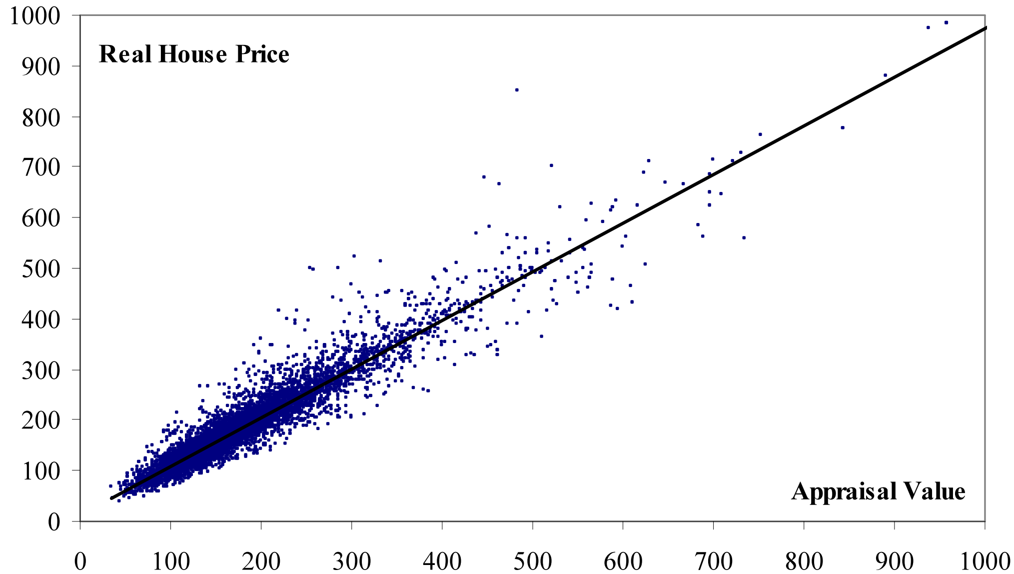
Two main results emerge from our study:

- In the Netherlands, the quality of the official appraisal values, while subject to certain limitations, is sufficient to use them for computing a SPAR index.
- For the Netherlands the difference in the trend of the (unweighted geometric) repeat sales index and the (value-weighted arithmetic) SPAR index, while not negligible in the very long run, differ only slightly for shorter time intervals. Also, the SPAR index is more precise than the repeat sales index.

Given these results, the fact that SPAR price index numbers are not revised, which is an important aspect for many users of official statistics, and the desire to have a single house price index for the Netherlands, *Kadaster* and *Statistics Netherlands* decided to go for the SPAR method. As from January 2008 they jointly publish SPAR price index numbers for the whole country and also for different types of dwellings, for each province, etc.

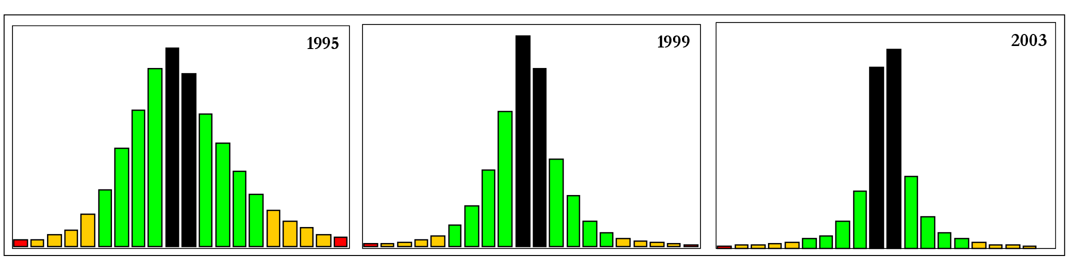
Figures and Tables

Figure 1: Real house prices and appraisal values in January 2003



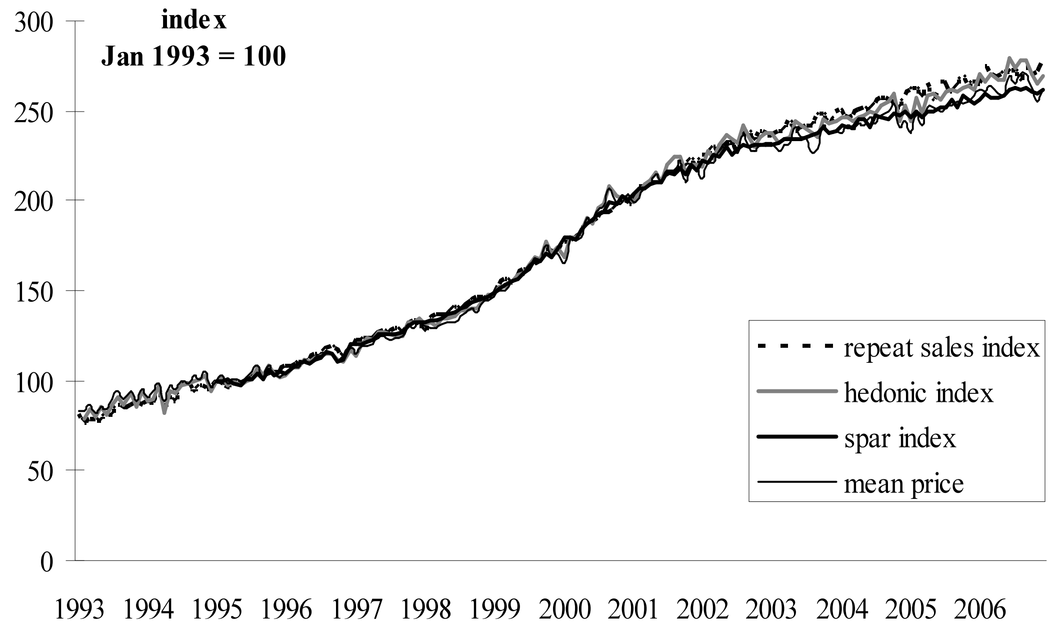
Source: Kadaster Netherlands, computation OTB (10% sample)

Figure 2: Distribution of the ratio of real house price and appraisal value in the Province of Overijssel (the Netherlands)



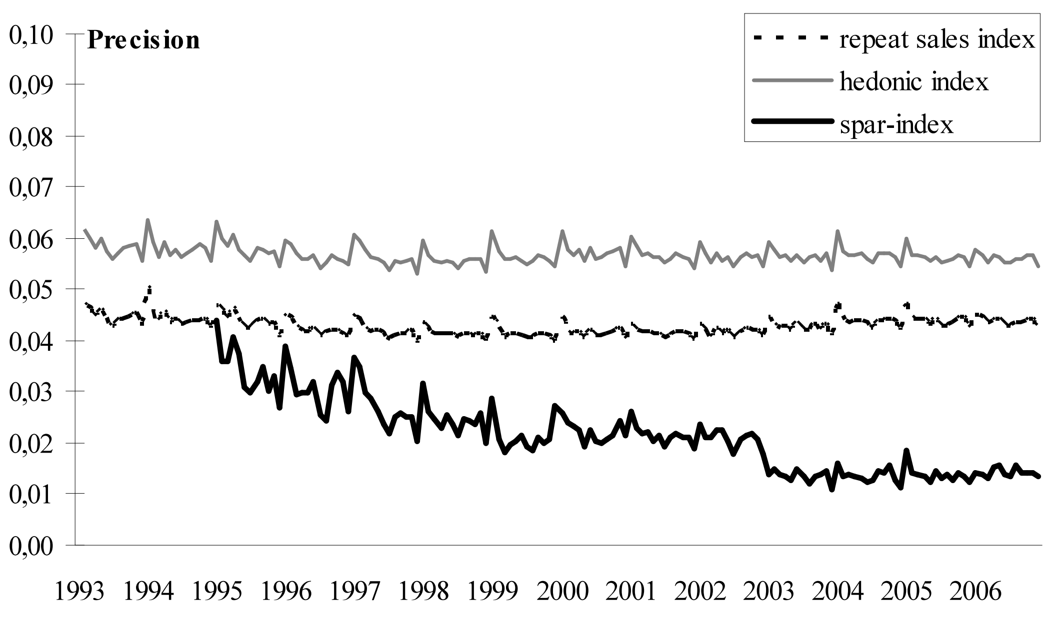
Source: Kadaster; computation OTB

Figure 3: House price indexes for the Province of Overijssel (the Netherlands), January 1993 – December 2006



Source: Kadaster; computation OTB

Figure 4: Precision of the house price indexes for the Province of Overijssel (the Netherlands), January 1993 – December 2006



Source: Kadaster; computation OTB

Table 1: Difference between real house prices and appraisal values

Appraisal date	Mean *A* €	*Mean rP* €	Fraction *rP/A*	Change in Std %	R^2
January 1995	87,607	90,538	0.968	16.2	0.903
January 1999	133,901	130,532	1.026	11.4	0.940
January 2003	202,695	200,167	1.012	10.7	0.951

Source: Kadaster Netherlands, computation OTB and Statistics Netherlands

Table 2: Cohen's *d* for municipalities in the Province of Overijssel (the Netherlands)

	1995		1999		2003	
Dwelling types						
Apartments	0.01		0.01		0.02	
Terraced house	0.01		0.06		0.04	
Corner house	0.06		0.12		0.00	
Semi-detached house	0.07		0.09		0.03	
Detached house	0.09		0.13		0.10	
Municipalities						
A	0.33	*	0.06		0.04	
B	0.11		0.05		0.16	
C	0.27	*	0.08		0.07	
D	0.36	*	0.01		0.05	
E	0.04		0.02		0.03	
F	0.09		0.08		0.01	
G	0.26	*	0.05		0.09	
H	0.30	*	0.05		0.08	
I	0.17		0.08		0.02	
J	0.23	*	0.13		0.02	
K	0.57	**	0.26	*	0.10	
L	0.22	*	0.07		0.01	
M	0.41	*	0.17		0.02	
N	0.53	**	0.17		0.05	
O	0.66	**	0.30	*	0.02	
P	0.43	*	0.18		0.10	
Q	0.46	*	0.03		0.13	
R	0.05		0.03		0.03	
S	0.06		0.01		0.08	
T	0.30	*	0.04		0.01	
U	0.21	*	0.12		0.05	
V	0.25	*	0.08		0.14	
W	0.22	*	0.16		0.06	
X	0.06		0.14		0.01	
Appraisal Value in deciles						
1	0.37	*	0.09		0.28	*
2	0.51	**	0.17		0.24	*
3	0.70	**	0.09		0.07	
4	0.33	*	0.29	*	0.17	
5	0.08		0.43	*	0.04	
6	0.06		0.33	*	0.01	
7	0.23	*	0.15		0.07	
8	0.12		0.40	*	0.02	
9	0.06		0.36	*	0.07	
10	0.01		0.19		0.13	
Total	0.03		0.06		0.01	

Note: * and ** indicate Cohen's d at 0.2 (small effect) and 0.5 (medium effect)
Source: Statistics Netherlands; computation OTB

Table 3: Results for hedonic, repeat sales and SPAR indexes

	Hedonic		Repeat Sales		SPAR	
Statistics						
R^2		0.58		0.82		Not applicable
Sample size		111,858		32,496		111,858
Precision		5.7%		4.3%		2.1%
Hedonic variables (all significant p=<0.01)						
Type: Apartments		Reference		.		.
. Type: House in a row		-0.11		.		.
. Type: Semi detached		-0.05		.		.
. Type: Detached		0.02		.		.
Neighbourhood: Apartments		Reference		.		.
. Neighbourhood: mixed dwellings types		0.43		.		.
. Neighbourhood: 50%> house in a row		0.61		.		.
. Neighbourhood: 50%> (semi)detached		0.68		.		.
Neighbourhood: Apartments, no gardens		Reference		.		.
. Neighbourhood: mixed ground space		-0.32		.		.
. Neighbourhood: mainly ground 1st quintile		-0.41		.		.
. Neighbourhood: mainly ground 2nd quintile		-0.24		.		.
. Neighbourhood: mainly ground 3rd quintile		-0.19		.		.
. Neighbourhood: mainly ground 4th quintile		-0.32		.		.
. Neighbourhood: mainly ground 5th quintile		-0.41		.		.
House price index						
July 2005	263	(-1.1%)	261	(+1.9%)	264	(-0.1%)
August 2005	260	(-1.2%)	262	(+0.3%)	269	(+1.8%)
September 2005	263	(+1.3%)	260	(-0.7%)	264	(-1.7%)
October 2005	269	(+2.3%)	263	(+1.0%)	271	(+2.5%)
November 2005	264	(-2.0%)	264	(+0.5%)	269	(-0.5%)
December 2005	266	(+0.7%)	262	(-0.8%)	267	(-0.8%)
…….						
July 2006	273	(+0.4%)	274	(-1.8%)	274	(+0.2%)
August 2006	266	(-2.7%)	278	(+1.6%)	274	(+0.0%)
September 2006	275	(+3.4%)	278	(-0.2%)	275	(+0.5%)
October 2006	270	(-1.7%)	270	(-2.7%)	275	(+0.0%)
November 2006	272	(+0.7%)	264	(-2.2%)	271	(-1.4%)
December 2006	277	(+1.9%)	270	(+2.0%)	274	(+0.8%)

Note: monthly percentage changes in parentheses

Source: Kadaster Netherlands; computation OTB

References

Abraham J.M., and W.S. Schauman (1991), "New Evidence on Home Prices from Freddie Mac Repeat Sales", *AREUEA Journal* 19, 333-352.

Bailey, M.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,

Baroni, M., F. Barthélémy and M. Mokrane (2004), "Physical Real Estate: a Paris Repeat Sales Residential Index", Working paper DR04007, ESSEC, Research Center.

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

Calhoun, C. A (1996), "OFHEO House Price Indexes: HPI Technical Description", Internet: 1-14.

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, K.E., and R.J. Shiller (1989), "The Efficiency of the Market for Single Family Homes", *American Economic Review* 79, 125–37.

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

Colwell, P.F., and G. Dilmore (1999), "An Early Hedonic Study", *Land Economics* 75, 620-626.

De Vries, P., and P. Boelhouwer (2005), "Local House Price Developments and Housing Supply", *Property Management* 23, 80-96.

De Vries, P., S. Jansen, C. Lamain, P. Boelhouwer and H. Coolen (2006), "Marktprijs en WOZ-waarde in de provincie Overijssel" [*Market Price and Appraisal Value in the Province of Overijssel*], Kadaster, Apeldoorn, The Netherlands.

De Vries, P., S. Jansen, P. Boelhouwer, H. Coolen, C. Lamain, E. van der Wal and D. ter Steege (2007), "Samenhang tussen marktprijs en WOZ-waarde" [*Coherence between Market Price and Appraisal Value*], *Tijdschrift voor de Volkshuisvesting*, 44-49.

Diewert, W.E. (2003), "The Treatment of Owner Occupied Housing and Other Durables in a Consumer Price Index", Discussion Paper no. 95-01, 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.

Dreiman M.H., and A. Pennington-Cross (2004). “Alternative Methods of Increasing the Precision of Weighted Repeat Sales House Price Indices”, *Journal of Real Estate Finance and Economics* 28, 299-317.

Edelstein, R.H., and D.C. Quan (2006). “How Does Appraisal Smoothing Bias Real Estate Returns Measurement?”, *The Journal of Real Estate Finance and Economics* 32, 41-60.

Englund P., J.M. Quigley and C. Redfearn (1998), “Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices”, *Journal of Urban Economics* 44, 171-196.

Geltner, D.M. (1991), “Smoothing in Appraisal-Based Returns”, *Journal of Real Estate Finance and Economics* 4, 327-345.

Griliches, Z (ed.) (1971), “Price Indices and Quality Change”, Cambridge, MA, Harvard University Press.

Haan, J. de, E. van der Wal, D. ter Steege and P. de Vries (2006), “The Measurement of House Prices: A Review of the SPAR Method”, paper presented at the sixth Economic Measurement Group Workshop, 13-15 December 2006, Sydney.

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.

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 2006 ENHR Congress.
http://enhr2006-ljubljana.uirs.si/publish/W02_Jansen.pdf

Jansen S., P. de Vries, H. Coolen and C. Lamain (2007), “Heteroskedasticiteit in de Woningwaarde-index Kadaster” [Heteroscedasticy in the House Price Index Kadaster], Delft University of Technology.

Leventis, A. (2006), “Removing Appraisal Bias from a Repeat-Transactions House Price Index: A Basic Approach”, OFHEO Working Paper 06-1.

Mason, C., and J.M. Quigley (1996), “Non-parametric Hedonic Housing Prices”, *Housing Studies* 11, 373-385.

Shiller, R. (1991), “Arithmetic Repeat Sales Price Estimators”, *Journal of Housing Economics* 1, 110-126.

Sirmans, G.S., D.A. Macpherson and J. Zietz (2005), “The Composition of Hedonic Pricing Models”, *Journal of Real Estate Literature* 13, 3-43.

Steele, M., and R. Goy (1997), “Short Holds, the Distributions of First and Second Sales, and Bias in the Repeat-Sales Price Index”, *Journal of Real Estate Finance and Economics* 14, 133-154.

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

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.