

Developing a House Price Index for The Netherlands: A Practical Application of Weighted Repeat Sales

S. J. T. Jansen · P. de Vries · H. C. C. H. Coolen ·
C. J. M. Lamain · P. J. Boelhouwer

Published online: 11 July 2007

© Springer Science + Business Media, LLC 2007

Abstract This paper describes the development of a house price index that has been introduced in May 2005 in The Netherlands. This monthly index, called Woningwaarde Index Kadaster (House Price Index Kadaster), is designed to detect changes in the price of the overall stock of owner-occupied homes. Fifty-five indices are calculated: one overall index, four regional indices, 12 provincial indices and 38 indices based on combinations of region/province and dwelling type. We used Case and Shiller's geometric Weighted Repeat Sales Model to calculate monthly house price indices. We used recorded data on the sales of over 500,000 owner-occupied homes in The Netherlands, all representing repeat sales between January 1993 and December 2006. The accuracy of the index was determined using the 95% confidence interval. We observed that accuracy might become a problem in smaller sub samples. Revision volatility was explored by comparing the index values computed from all available data until December 2005 with the index values computed from the data available until December 2006. Our analysis showed that revision volatility does not seem to be a major problem to the index. We also explored heteroskedasticity in the Repeat Sales method but did not find conclusive evidence for the proposed heteroskedasticity. Given our target (a geometric mean index value) and the characteristics of the dataset (very large but without property characteristics) the Repeat Sales Method seems to be adequate for calculating a house price index for The Netherlands.

S. J. T. Jansen (✉) · P. de Vries · H. C. C. H. Coolen · C. J. M. Lamain · P. J. Boelhouwer
OTB Research Institute, Delft University of Technology,
P.O. Box 5030, 2600 GA Delft, The Netherlands
e-mail: s.j.t.jansen@tudelft.nl

P. de Vries
e-mail: p.devries@tudelft.nl

H. C. C. H. Coolen
e-mail: h.c.c.h.coolen@tudelft.nl

C. J. M. Lamain
e-mail: c.j.m.lamain@tudelft.nl

P. J. Boelhouwer
e-mail: p.j.boelhouwer@tudelft.nl

Keywords Weighted repeat sales · House price index · Revision volatility · Accuracy · Heteroskedasticity

Introduction

In The Netherlands, as elsewhere, there is a need for a house price index that would, amongst other things, enable financial organizations to value the collateral behind mortgage portfolios. In fact, the Dutch central bank, De Nederlandsche Bank, requires that financial institutions specify their risks with regard to their mortgage portfolios by estimating the actual liquidation value for every home in their portfolio. Another application of a house price index in The Netherlands is to allow brokers and homeowners to calculate the current value of an individual dwelling as well as the amount of equity gained (or lost) through house price appreciation (or depreciation). These two arguments apply to regional or provincial indices. Next to these indices, a national index would be useful to keep track of the national development of house prices in The Netherlands from year to year. Furthermore, regional or provincial indices could be compared to the national index to examine whether they differ from the national tendency of growth in house prices. Lastly, Eurostat, the Statistical Office of the European Communities recommends associated European countries to develop a national house price index in order to be able to make comparisons between European countries. The goal of our index is to follow the mean price development of an existing home in the entire stock of owner-occupied homes in The Netherlands.

Worldwide, the most frequently used methods for calculating house price indices are: (1) a summary measure of central tendency (e.g., mean, median); (2) hedonic price models; (3) Repeat Sales Models; and (4) variants on and hybrids of the latter two.

Until recently, only the summary methods were applied in The Netherlands. Once a month the Dutch Land Registry Office (Kadaster¹) published the mean selling price and the National Association of Property Brokers (NVM) published the median selling price of existing homes. However, one intrinsic flaw in the summary methods is that they are not adjusted for quality. They are unable to distinguish between price movements and changes in the composition of sold dwellings from one period to the next (Bourassa et al. 2006). For example, if for some reason, a disproportionate number of high-priced homes were sold in a given month, the mean or median price would still rise, even though not a single house had increased in value (Case and Shiller 1987). Furthermore, the quality of new houses is likely to rise. Since these houses ultimately become existing houses, the median or mean price of existing houses will rise even if individual properties are not appreciating (Bailey et al. 1963; Case and Shiller 1987). The shortcomings in the summary methods meant that an alternative method had to be found for calculating a house price index for The Netherlands.

¹ Kadaster, or the Dutch Land Registry Office, collects information about registered properties in The Netherlands, records them in public registers and in cadastral maps and makes this information available to members of the public, companies and other interested parties in society.

The second option, hedonic regression analysis, is based on the principle that the price of a house can be accurately estimated from its characteristics. The selling price is regressed on a set of important qualitative variables, e.g., the number of rooms and lot size, and several variables for measuring time effects (Rosen 1974). The regression coefficients can be interpreted as implicit price attributes; for example, an extra room will push up the value of the property by a specific amount. However, the challenge posed by this method is to compute a functionally correct mathematical model for house prices. A correct set of explanatory variables must be specified and the relationships between these and the response variable must be correctly determined beforehand (Wang and Zorn 1997). Another drawback of this method is that quality characteristics are both numerous and difficult to measure. Hence the hedonic model may not yield useful results (Bailey et al. 1963).

Bailey et al. (1963) state that most of the difficulties of specifying and measuring quality characteristics can be avoided by basing the price index on the selling prices of the same properties at different times. This method—the Repeat Sales Model—checks quality characteristics by comparing the same property over time. It uses data on properties that have actually been sold more than once during the period in question and focuses on price changes rather than prices themselves (Wang and Zorn 1997). The greatest drawback of Repeat Sales is that it wastes data by only using information on repeat sales (Wang and Zorn 1997).

Finally, hybrid models avoid the inefficiency of the Repeat Sales Model because they also use information from houses that are only sold once (Wang and Zorn 1997). They might avoid the problem of misspecification to which the hedonic method is susceptible. However, like the hedonic method, hybrid models require a large database with a detailed set of property attributes.

In 2004, yet another method for calculating house price indices was introduced in The Netherlands. It was developed by Von Dewall et al. (2004) and called the Integrated House Price Index (Geïntegreerde Woningprijs Index/ GWI). Basically, the GWI calculates the mean appreciation rate of groups of properties that are purchased in the same period (e.g., month, quarter, year) and re-sold later. The appreciation rate is obtained for the various time periods by comparing the appreciation rates of groups of properties with the same purchase date and a different selling period, and by repeating this procedure for every purchase period. The method uses properties that are sold at least twice. The calculation method for the GWI seems to have a lot in common with the chain index described in Bailey et al. (1963). One benefit of such a method is that it is computationally simple. However, it is also inefficient, especially in the earlier periods, because it neglects index data for earlier periods contained in price relatives with final sales in later periods. Another drawback of such a method is that it does not provide standard errors for the index values.

The choice of method for calculating an index depends on the ‘target’ (Wang and Zorn 1997) and the characteristics of the available dataset (Abraham and Schauman 1991). The target is the statistic that users of an index need to know regardless of the method (Wang and Zorn 1997). Our target is the geometric mean index value—which matches well with the Repeat Sales Model. Moreover, whereas the hedonic and hybrid methods can be used only if information is available on the characteristics of individual homes (e.g., number of rooms, lot size), Repeat Sales can be

applied when only the purchase and selling prices and the dates of sale are known. In The Netherlands, data on all houses sold are recorded by the Dutch Land Registry Office (Kadaster) since January 1993. However, as no details are recorded on house characteristics apart from built surface area and type of dwelling (detached house, corner house, terraced house, apartment, semi-detached house), hedonic and hybrid methods cannot be applied. For these reasons, Repeat Sales seems a logical choice for a house price index for The Netherlands. One disadvantage of Repeat Sales is that it requires a large dataset, because only houses that are sold more than once are used to calculate the index values. Fortunately, the dataset of the Dutch Land Registry Office is quite large, containing all the sales of owner-occupied homes since January 1993 in The Netherlands (more than 2.5 million transactions, more than 700,000 of which are repeat sales). This is why we chose the Repeat Sales Model as the method for calculating a house price index for The Netherlands. In the next section, our practical application of the (Weighted) Repeat Sales method will be described.

Materials and Methods

Weighted Repeat Sales Model

As the (weighted) Repeat Sales Model is extensively addressed in the literature (see e.g., Bailey et al. 1963; Case and Shiller 1987, 1989; Goetzmann 1992; Calhoun 1996; Dreiman and Pennington-Cross 2004), we believe that a brief description here will suffice. A more detailed description of our application of the (Weighted) Repeat Sales method can be found in Jansen et al. (2005).

Bailey et al. (1963) were the first to develop a house price index that was based on the Repeat Sales Model. Essentially, Repeat Sales uses a collection of the prices paid for single properties at different points in time to estimate a vector of numbers that ‘best’ explains the observed changes in price over the sample period (Abraham and Schauman 1991). In practice, the Repeat Sales Model uses ordinary least squares regression analysis in which the dependent variable is the logarithm of the price relative from the twice-sold property. The log price relatives are then regressed on a set of dummy variables corresponding with the time periods. A dummy variable is added for each period, except the first (base) period. The dummy variable for the first sale has the value ‘−1’ and the dummy variable for the second sale has the value ‘+1’. All other dummy variables have the value ‘0’. There is no constant term in the analysis, the coefficients are estimated only on the basis of changes in house prices over time. The estimated coefficients represent the log of the cumulative price index for each period. The time dummy for the initial period is set at zero to normalize the index at 1. The regression equation is (Bailey et al. 1963):

$$r_{it'} = \sum_{j=1}^T b_j x_j + u_{it'} \quad (1)$$

where $r_{it'}$ is the log of the ratio of the final sales price in period t' to initial sales price in period t for the i th pair of transactions with initial and final sales in these two

periods, b is a column vector of unknown logarithms on the index numbers to be estimated, and x is an $n \times T$ matrix with values -1 , 0 , and 1 , as explained above. Finally, u_{it} are the residuals in log form with zero means, equal variances, and uncorrelated with each other.

In 1987, Case and Shiller published an adapted version of the Repeat Sales Model of Bailey et al. (1963): the Weighted Repeat Sales method. Case and Shiller argued that the longer the time between transactions the more variance there is in individual house price appreciation; for example, because some houses are very well maintained whereas others are not maintained at all. As a result, the variance of the residuals (i.e. the differences between predicted and observed house prices) will increase with the length of the holding period. This phenomenon—known as heteroscedasticity—undermines efficiency as the variance of the index values becomes too great (Wang and Zorn 1997). This may not be a problem if the application relies solely on the indices themselves and are based on plentiful data (Wang and Zorn 1997). However, heteroscedasticity is certainly a problem if confidence intervals are calculated (Wang and Zorn 1997). To minimize the effect of heteroscedasticity, Case and Shiller (1987) proposed a three-step procedure, which is described below.

The first step is exactly the same as the first step of the Repeat Sales Model described by Bailey et al. (1963). In the second step, a regression analysis is performed on the squared residuals from the first step. Time is incorporated as an independent variable (predictor) in the model and a constant term (intercept) is also included. This intercept is an estimate of the variance of twice the house-specific random error variance, once for the first sale and once for the second sale (Case and Shiller 1987). The time coefficient is an estimate of the increase in variance for each additional period. This is called the ‘Gaussian Random Walk’. The random walk model implies that the variance of house prices (and growth rates) increases linearly with time (Wang and Zorn 1997). Thus, the second step explores the assumption that the error variance increases linearly with the holding interval and that there is a fixed component to the property specific variance that is not related to the holding period (Goetzmann 1992).

In the third step of the procedure, a weighted regression analysis (Generalized Least Squares Regression) is applied where the weights are the reciprocals of the square roots of the fitted values of the second-stage regression. This procedure minimizes the impact of houses with a relatively long holding period on the regression analysis (Abraham and Schauman 1991). The log price of the i th house at time t is given by (Case and Shiller 1987):

$$P_{it} = C_t + H_{it} + N_{it}, \quad (2)$$

where C_t is the log of the citywide level of housing prices at time t ; H_{it} is a Gaussian random walk that represents the drift in individual housing value through time, and N_{it} is a house-specific random error that has zero mean and equal variance and is serially uncorrelated.

Various authors have proposed additions and corrections to (weighted) Repeat Sales. In 1991, Abraham and Schauman (1991) argued that the variance of the error term associated with any Repeat Sales pair would not indefinitely increase linear to

the holding period. Instead, they proposed a quadratic model so that the increase in variance would decrease as the holding period increased:

$$E[d_i^2] = A(t-s) + B(t-s)^2 + 2C \quad (3)$$

where d_i^2 refers to the squared residuals, $t-s$ refers to the number of periods between acquisition and sale, the constant term $2C$ provides an indication of the variance of twice the house-specific random error, A is an estimate of the increase in variance for each additional period, and, finally, B is an estimate of the increase in variance for each additional period squared. We followed this approach in the second step of our calculation of the Woningwaarde Index Kadaster, just like Calhoun (1996) for the OFHEO index.

Furthermore, in 1992, Goetzmann proposed an ex-post correction to the model by Case and Shiller (1987). Goetzmann states that the Repeat Sales method provides an estimate of the geometric mean growth rate and not of the arithmetic mean growth rate. Because the log function is concave, the average of the logs is less than the log of the average, when there is any variance in the data (Goetzmann 1992). The log transformation results in a downward bias of the arithmetic mean at each point in time (Goetzmann 1992). Goetzmann (1992) argues that the geometric return has a natural interpretation for a times series where it represents the growth rate of an investment over time. However, for a cross-sectional interpretation an arithmetic return seems more natural. Goetzmann (1992) suggests a relatively simple scalar adjustment to the estimated geometric means based on adding half the variance in house price growth rates associated with the diffusion of house prices over time. Calhoun (1996) proposes to also include a term in this calculation for time squared, as in the second step of the procedure.

We do not directly apply the Goetzmann correction in our calculation of the house price index for various reasons. Firstly, one goal of the Woningwaarde Index Kadaster is to provide a measure for homeowners and brokers to calculate the growth rate for an individual dwelling. In such a longitudinal context the geometric mean is an adequate measure of center (Wang and Zorn 1997). Secondly, the parameters needed to calculate the Goetzmann correction have to be provided separately if the value of a portfolio of dwellings is to be calculated, because the form of the correction function is non-linear (e.g., the increase in the variance between the first two periods is larger than for the last two periods). Thus, the parameters are dependent upon the beginning and ending dates of the particular portfolio. In such a case, e.g., when banking institutions want to calculate the value of their entire portfolio of mortgages at once, the necessary parameters can be provided separately and the Goetzmann correction can be calculated for the particular portfolio. This is the strategy that is followed by the OFHEO House Price Index (Calhoun 1996).

The Dataset

The Dutch Land Registry Office is responsible for the administration of all properties sold in The Netherlands (including all owner-occupied homes). The

dataset contains information on 2,599,449 individual transactions regarding owner-occupied homes between January 1993 and December 2006. A total of 121,666 transactions were deleted because information on either the type of dwelling or the Intramax region (see next section for an explanation of the term Intramax region) was missing, resulting in 2,477,783 transactions.

Table 1 shows the owner-occupied stock in November 2006, the number of dwellings sold at least once between January 1993 and December 2006, the number of dwellings sold twice or more, and the number of pairs of Repeat Sales for the different types of dwellings. It may be deduced from the table that, between January 1993 and December 2006, 47% of all owner-occupied homes were sold at least once. Fifteen percent of dwellings ($n=549,993$) were sold at least twice. Of the dwellings sold since January 1993, 32% were at least sold twice.

Then, the number of transactions related to repeat sales were calculated. First, all transactions ($n=1,057$) related to dwellings that were sold more than ten times ($n=46$) were deleted. This was done for reasons of validity. Dwellings that are frequently resold may not be representative, for example, because they have hidden drawbacks that become overt only after sale (so-called ‘lemons’). This resulted in 2,476,726 transactions. Next, transactions that related to only one sale or that related to the first sale of multiple sales were deleted ($n=1,740,685$) in order to obtain pairs of repeat sales (two successive sales form one pair). This resulted in 736,041 pairs of repeat sales.

Next, we deleted 54,518 pairs of repeat sales (7,4%) that were transactions related to dwellings that were sold within 12 months, because a short interval between the acquisition and divestment of a house may imply an unusual transaction (Englund et al. 1998). On the one hand, these may represent distressed sales arising from divorce or job loss. On the other hand, they may be speculative sales. No conveyance tax needs to be paid in The Netherlands if a house is resold within 6 months. In a period of rapidly rising house prices, as observed between 1998 and 2001 in The Netherlands, a number of sales will have taken place purely for speculative reasons. Clapp and Giacotto (1999) advise that transactions, which they refer to as ‘flips’, be removed or weighed down. Flips are houses that are resold within 1 or 2 years of purchase. Clapp and Giacotto suggest that flips are (cosmetically) improved after purchase and have therefore appreciated at a higher rate when they are sold again soon afterwards. Thus, they introduce an upward bias to the index values. Finally, Steele and Goy (1997) argue that the opportune buyer rationale for the existence of bias in the price change of repeat sales properties implies that the bias should be greater the shorter the holding period. They too suggest eliminating very short holds from the dataset.

To explore the potential impact of very short holds, we calculated the monthly growth rate for every dwelling (including the ‘flips’):

$$\text{Monthly growth rate} = (((P_t/P_{t-1})^{**}(1/t)) - 1) * 100 \quad (4)$$

Where P_t represents the price at the second sale, P_{t-1} represents the price at the first sale, and t indicates the period in months between sales.

Figure 1 confirms that deviating changes occur in the growth rate of homes resold within 12 months. For example, the mean growth rates are 8.3, 5.3, 1.2, and 0.9%

Table 1 Owner-occupied stock (November 2006), number of dwellings sold and not sold, and number of pairs of repeat sales up till December 2006

	Owner-occupied stock	Number of dwellings not sold	Percent (%)	Number of dwellings sold at least once	Percent (%)	Number of dwellings sold twice or more	Percent (%)	Pairs of repeat sales
Overall	3,709,921	1,968,995	53	1,740,926	47	549,993	15	735,796
Types								
Apartments	520,384	161,470	31	358,914	69	157,364	30	235,394
Single-family homes	3,189,573	1,807,561	57	1,382,012	43	392,629	12	482,829
Sub-types								
Terraced houses	1,326,070	661,489	50	664,581	50	211,760	16	265,310
Corner houses	525,916	273,235	52	252,681	48	72,695	14	89,142
Semi-detached	569,560	347,010	61	222,550	39	57,953	10	69,734
Detached	767,991	525,791	68	242,200	32	50,221	7	58,643

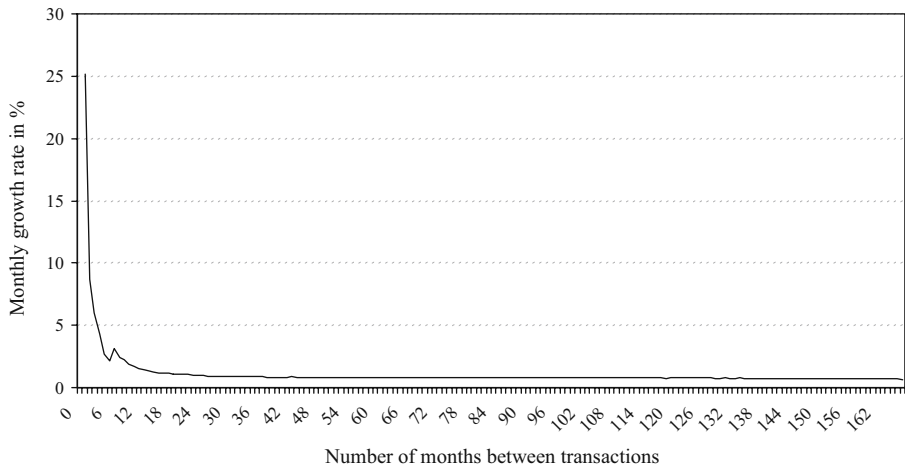


Fig. 1 The mean growth rate value per month (%) across the number of months between two transactions

for houses sold within 6 months, within 12 months, within all periods, and between 12 months and the end of period, respectively. Homes sold within a few months realize, on average, a very high increase in value per month, which may bias the index.

Transaction or Sample Selection Bias

The repeat sales sample consists of a selection of houses that have been sold at least twice between January 1993 and December 2006. This sample may not, however, be representative of the overall stock of owner-occupied homes in The Netherlands. In other words, a problem will arise if the price changes in the sample are different from those in the rest of the housing stock. This phenomenon is known as ‘sample selection bias’ or ‘transaction bias’. For example, Table 1 shows that 30% of the apartments have been sold at least twice since January 1993 whereas only 7% of detached homes were sold at least twice in that same period.

Samples of repeat sales may differ from the overall housing stock for different reasons (Bourassa et al. 2006). First, properties may have been bought explicitly for the purpose of renovation and resale. Second, properties that are repeatedly sold may not meet buyer expectations (so-called lemons), and third, starter homes sell more frequently as the owners tend to move on to larger (and better) dwellings. Costello and Watkins (2002) discuss the ‘starter home hypothesis’ (2002) and point out that houses which are sold more frequently tend to be smaller and cheaper and to appreciate more rapidly than houses which are sold less frequently. One of the explanations for this finding is that younger homeowners may upgrade their home more frequently (Costello and Watkins 2002). Thus, in general, properties in the repeat sales sample may be in a poorer condition and worth less (at least at the time of the purchase; Bourassa et al. 2006).

As stated in “Introduction,” the goal of our index is to follow the mean price development of an existing home in the entire stock of owner-occupied homes in The Netherlands. One can imagine that houses with different values will show

different appreciation rates; however, the value of houses in the overall stock of owner-occupied homes is not known until the actual sale is transacted. Thus a correction according to value is not possible. Another factor worth considering is that the rate at which house prices appreciate may vary from region to region. Houses from different regions may not be represented in the repeat sales sample in the same proportion as they are represented in the overall stock of owner-occupied homes.

It is for these reasons that we decided to weigh the repeat sales sample so that it resembles the overall stock of owner-occupied homes as closely as possible. However, as only a few characteristics were available in the dataset of the Dutch Land Registry Office (Kadaster), we were only able to weigh for type of dwelling (corner house, detached house, semi-detached house, terraced house, apartment) and region. Type of dwelling is used as a proxy for value because apartments are more strongly represented in the lower price classes and detached homes in the higher price classes. With regard to weighing by region, we considered regional classification on the basis of four regions (north, east, south, west) and on the basis of our 12 provinces. However, these classifications are based on administrative borders, which may be of little or no importance to house-seekers. For this reason, appreciation rates may differ more within than between provinces. Accordingly, we turned to a classification that is not based on administrative borders but on movements, working and living patterns, and the pressure on regional housing markets (Masser and Scheurwater 1978). This classification, called the Intramax Regions, is used by, among others, Van Kempen et al. (1995) and Goetgeluk (1997). The most recent Intramax classification in 13 Intramax regions was compiled by the University of Utrecht.

In practice, the weighing procedure ensures that the distribution over the 13 Intramax housing market regions and the five dwelling types is reflected in the repeat sales sample as in the overall stock of owner-occupied homes. This procedure reduces the selection bias by down weighting observations from housing types that are sampled “too frequently” in the Repeat Sales sample. For example, in our national analysis apartments have a weighing factor of 0.43, which indicates that they are overrepresented in the repeat sales sample in comparison with the overall stock. Conversely, detached houses are underrepresented (factor of 2.67) in the repeat sales sample. Higher weights indicate more impact in the regression analyses. Table 2 shows the distribution over Intramax regions and types of dwelling in the owner-occupied stock and in the entire Repeat Sales sample. Table 3 shows the resulting weights for the data up to December 2006. Note that with every additional month of data, the weights are determined anew. Note further that in the case when results are calculated for sub samples, such as provinces and regions, the weights, based on type of dwelling and Intramax region, are calculated for every subsample separately.

Furthermore, to eliminate random bias due to, e.g., typing errors, we omitted pairs of cases in which the logarithm of the price relative from the twice-sold property (i.e. the dependent variable in the regression analysis) showed more than five standard deviations from the mean value. In the case of normally distributed data, the odds of that occurring are only about one in a million. However, such cases can distort the analyses since the sum of squares is being minimized in the regression analysis and

Table 2 Distribution of dwellings and pairs of repeat sales over Intramax regions and types of dwellings

Intramax regions														
Dwelling types	Noord	Oost	Arnhem-Nijmegen	Noord-west Veluwe	Utrecht	Amstel-landen	Kop Noord-Holland	Haag-landen	Rotte-landen	Zeeland	West-Brabant	Overig Brabant	Limburg	Total
(%)														
Entire owner-occupied stock														
Apartments	0.8	0.6	0.7	0.3	0.9	2.9	0.3	3.6	2.2	0.2	0.3	0.7	0.5	14.0
Terraced houses	2.3	3.5	1.9	1.5	2.9	5.2	1.6	4.4	3.4	0.9	1.6	3.8	2.7	35.7
Corner houses	1.1	1.5	1.0	0.7	0.9	1.8	0.7	1.5	1.3	0.4	0.8	1.7	0.8	14.2
Semi-detached	2.2	2.9	1.4	0.6	1.0	1.0	0.5	0.6	0.5	0.5	0.7	1.7	1.7	15.4
Detached	4.5	3.5	1.5	0.8	1.1	0.9	1.0	0.7	0.6	0.9	1.1	2.3	1.7	20.7
Total	10.9	12.0	6.5	3.9	6.8	11.8	4.1	10.9	8.1	3.0	4.5	10.1	7.4	100.0
Pairs of repeat sales														
Apartments	1.8	1.6	2.0	0.8	2.6	4.7	0.6	8.6	5.7	0.3	0.8	1.9	1.2	32.5
Terraced houses	3.4	4.2	2.1	1.9	3.4	4.2	1.8	3.6	2.9	1.1	1.9	4.5	2.5	37.5
Corner houses	1.3	1.4	0.8	0.7	0.8	1.3	0.6	1.1	1.0	0.4	0.8	1.6	0.7	12.6
Semi-detached	1.8	1.8	0.8	0.4	0.6	0.7	0.3	0.3	0.3	0.3	0.5	1.0	0.9	9.6
Detached	2.2	1.2	0.4	0.3	0.4	0.4	0.4	0.2	0.2	0.4	0.4	0.7	0.4	7.7
Total	10.4	10.2	6.1	4.1	7.9	11.4	3.6	13.8	10.2	2.5	4.4	9.7	5.7	100.0

such cases may obtain too much weight. In the national sample, about 0.5% of cases ($n=3,329$) were deleted because they were outliers and 678,194 pairs of repeat sales remained for use in the regression analyses. In the case when results are calculated for sub samples, such as provinces and regions, the outliers are determined for every sub sample separately.

The Weighted Repeat Sales Regression Analysis

The results of the three steps of the Weighted Repeat Sales method for the national index and for the 12 provinces of The Netherlands are summarized in Tables 4 and 5. In the first step of the Weighted Repeat Sales method, an Ordinary Least Squares (OLS) regression analysis is performed in which the log price relatives are regressed on a set of dummy variables corresponding with the time periods. The residuals are saved. The results are presented in the first row of Tables 4 and 5.

In a subsequent regression analysis, the squared residuals obtained in the first step are included as dependent variables and the number of months and squared number of months since previous sale are included as predictors in the model (as proposed by Abraham and Schauman 1991). A constant term was also included. Unfortunately, our results show that the estimated coefficient for holding period squared is positive instead of negative for 11 out of 13 indices. This indicates that the error variance increases more than linearly with the holding period and therefore contradicts the assumption by Abraham and Schauman (1991) of diminishing growth. Furthermore, the coefficient for holding period is negative for six indices, indicating that there is a negative effect of holding period on the growth of variance. This is also contradictory to the theory. The results are presented in the second row of Tables 4 and 5 (method Abraham and Schauman).

Calhoun (1996) encountered a similar problem; he observed that the constant turned out to be negative. As the constant represents variance and variance cannot be negative, he formulated an alternative assumption that the normally distributed error term that represents cross-sectional dispersion in housing values arising from purely idiosyncratic differences in the valuation of individual houses at any given point in time is constant for every house (Calhoun 1996). Under this assumption, this term is cancelled from the equation and the squared residuals are estimated only on the basis of 'holding period' and 'holding period squared.' When we follow this procedure, the resulting coefficients are in agreement with the assumption posed by Abraham and Schauman (1991) for all 13 indices. The results are presented in the third row of Tables 4 and 5 (method Calhoun).

The fourth row of Tables 4 and 5 presents the results for the regression analyses based on the method of Case and Shiller. The results are in accordance to the theory, i.e., the amount of variance increases with the holding period.

Note, however, that irrespective of the method that is used to predict the relationship between the squared residuals and the holding period, the amount of explained variance is very small, ranging from 0.03 to 0.5%. So, even in the best situation, only a half percent of the spread in variance is explained by the holding period. Therefore, significant effects may be an effect of the large sample size.

In the third and final step of the Weighted Repeat Sales method, a weighted regression is performed (Generalized Least Squares) by repeating the regression

Table 4 Results of the three steps of the Weighted Repeat Sales method

Model	National index (<i>n</i> =678,194)	Groningen (<i>n</i> =25,138)	Friesland (<i>n</i> =27,415)	Drenthe (<i>n</i> =21,701)	Overijssel (<i>n</i> =41,673)	Flevoland (<i>n</i> =18,214)	Gelderland (<i>n</i> =71,152)
Step 1: OLS regression (no intercept) ^a							
<i>R</i> ²	82.0	78.9	76.4	76.0	81.5	73.4	86.5
Step 2: Abraham and Schaubman							
<i>R</i> ²	0.2	0.3	0.1	0.2	0.1	0.2	0.1
Intercept	0.0535900, <i>p</i> <0.01	0.0514047, <i>p</i> <0.01	0.0508910, <i>p</i> <0.01	0.0434, <i>p</i> <0.01	0.0606008, <i>p</i> <0.01	0.0670979, <i>p</i> <0.01	0.0501290, <i>p</i> <0.01
Coefficient	0.0000865, <i>p</i> =0.03	0.0004885, <i>p</i> =0.03	0.0011986, <i>p</i> <0.01	0.0013064, <i>p</i> <0.01	0.0000447, <i>p</i> =0.80	-0.0002596, <i>p</i> =0.47	-0.0002099, <i>p</i> =0.02
period							
Coefficient	0.0000017, <i>p</i> <0.01	0.0000003, <i>p</i> =0.83	0.0000006, <i>p</i> =0.75	-0.0000058, <i>p</i> <0.01	0.0000020, <i>p</i> =0.08	0.0000064, <i>p</i> =0.01	0.0000028, <i>p</i> <0.01
period ²							
Step 2: Calhoun ^a							
<i>R</i> ²	5.7	7.4	9.1	6.6	5.9	4.0	5.8
Coefficient	0.0016693, <i>p</i> <0.01	0.0020384, <i>p</i> <0.01	0.0027023, <i>p</i> <0.01	0.0026009, <i>p</i> <0.01	0.0018388, <i>p</i> <0.01	0.0018578, <i>p</i> <0.01	0.00125289, <i>p</i> <0.01
period							
Coefficient	-0.0000080, <i>p</i> <0.01	-0.0000092, <i>p</i> <0.01	-0.0000085, <i>p</i> <0.01	-0.0000137, <i>p</i> <0.01	-0.0000090, <i>p</i> <0.01	-0.0000073, <i>p</i> <0.01	-0.00000609, <i>p</i> <0.01
period ²							
Step 2: Case and Shiller							
<i>R</i> ²	0.1	0.3	0.1	0.2	0.1	0.2	0.1
Intercept	0.0470691, <i>p</i> <0.01	0.0501529, <i>p</i> <0.01	0.0484060, <i>p</i> <0.01	0.0655498, <i>p</i> <0.01	0.0526579, <i>p</i> <0.01	0.0451934, <i>p</i> <0.01	0.0390801, <i>p</i> <0.01
Coefficient	0.0003235, <i>p</i> <0.01	0.0005352, <i>p</i> <0.01	0.0012894, <i>p</i> <0.01	0.0004887, <i>p</i> <0.01	0.0003330, <i>p</i> <0.01	0.0005910, <i>p</i> <0.01	0.0001871, <i>p</i> <0.01
period							
Step 3: GLS regression (no intercept) ^a							
<i>R</i> ²	78.2	74.5	71.2	72.6	77.8	67.9	83.1

^a The amount of explained variance cannot be interpreted in the usual way because no intercept is included

Table 5 Results of the three steps of the Weighted Repeat Sales method (2)

Model	Utrecht ($n=58,384$)	Noord-Holland ($n=92,077$)	Zuid-Holland ($n=168,077$)	Zeeland ($n=16,361$)	Noord-Brabant ($n=99,037$)	Limburg ($n=38,965$)
Step 1: OLS regression no intercept ^a						
R^2	80.1	84.0	83.9	81.3	84.7	82.3
Step 2: Abraham and Schaubman						
R^2	0.2	0.03	0.1	0.5	0.1	0.5
Intercept	0.0402052, $p<0.01$	0.0593466, $p<0.01$	0.0481065, $p<0.01$	0.0478480, $p<0.04$	0.0592625, $p<0.01$	0.0417772, $p<0.01$
Coefficient	0.0008933, $p<0.01$	-0.0000552, $p=0.61$	-0.0000344, $p=0.61$	0.0002297, $p=0.21$	-0.0002552, $p=0.01$	-0.0001519, $p=0.11$
period						
Coefficient period ²	-0.00000025, $p=0.03$	0.00000014, $p=0.06$	0.00000017, $p<0.01$	0.00000013, $p=0.29$	0.00000034, $p<0.01$	0.00000032, $p<0.01$
Step 2: Calhoun ^a						
R^2	5.9	5.0	5.4	10.7	4.8	8.7
Coefficient period	0.0020560, $p<0.01$	0.0017032, $p<0.01$	0.0013836, $p<0.01$	0.0016638, $p<0.01$	0.0014924, $p<0.01$	0.0010767, $p<0.01$
Coefficient period ²	-0.0000096, $p<0.01$	-0.0000094, $p<0.01$	-0.0000070, $p<0.01$	-0.0000075, $p<0.01$	-0.0000072, $p<0.01$	-0.0000042, $p<0.01$
Step 2: Case and Shiller						
R^2	0.2	0.03	0.1	0.5	0.1	0.4
Intercept	0.0505111, $p<0.01$	0.0539732, $p<0.01$	0.0415072, $p<0.01$	0.0427321, $p<0.01$	0.0457489, $p<0.01$	0.0291200, $p<0.01$
Coefficient period	0.0005283, $p<0.01$	0.0001405, $p<0.01$	0.0002040, $p<0.01$	0.0004192, $p<0.01$	0.0002356, $p<0.01$	0.0003083, $p<0.01$
Step 3: GLS regression (no intercept) ^a						
R^2	76.8	81.2	80.6	77.2	81.0	77.6

^a The amount of explained variance cannot be interpreted in the usual way because no intercept is included

analysis from the first step and by dividing each case by the square root of the predicted value that was fitted in the second step (in our case calculated using the ‘Calhoun’ method).

The resulting index (including 95% confidence intervals) for The Netherlands is shown in Fig. 2. The general pattern of the index shows that house prices in The Netherlands increased gradually between January 1993 and December 2006. A relatively large increase in house prices was observed between 1998 and 2001. Figure 3 shows the indices for the 12 provinces of The Netherlands. The figure shows that although in all provinces house prices have gone up since 1993, there are two provinces (Flevoland and Limburg) in which the growth of house prices has been less than in the other provinces, especially after 2004.

The Search for Heteroskedasticity

As described before, Case and Shiller (1987) proposed an adapted version of the Repeat Sales model to correct for heteroskedasticity. They argued that the residuals would increase with the holding period. However, our results showed that, at best, only 0.5% of the spread in variance of the residuals could be explained by the holding period. For this reason, we explored the assumed heteroskedasticity in more detail.

First, we explored whether heteroskedasticity was indeed present, irrespective of the presumed cause. The most simple way to explore heteroskedasticity is to make a scatter plot of the residuals. Note that SPSS was not able to generate scatter plots for the whole sample (sample size too large) so for the national sample we used random samples of 10% of the data. All scatter plots showed that the variance was not spread evenly over the levels of the predictors. Instead, the largest variance was generally observed for the middle category, i.e., the category of dwellings that had not been bought or sold in that particular month. This was also by far the category with the largest number of observations, so this may explain the observed heteroskedasticity.

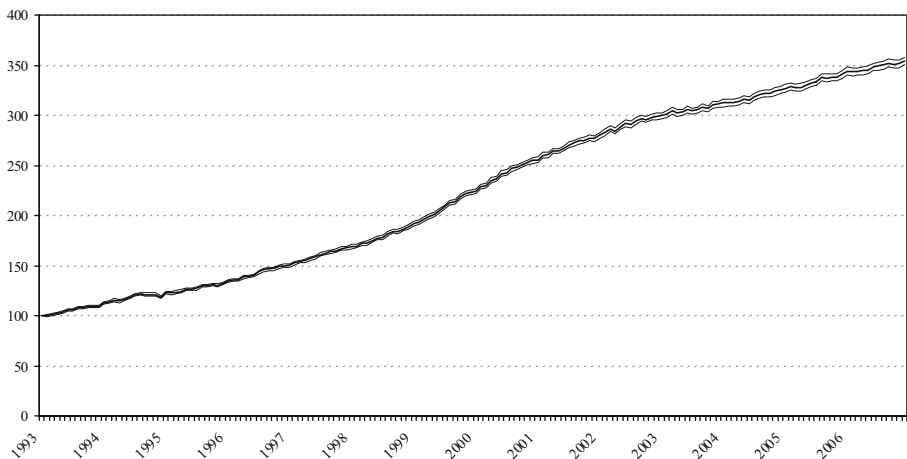


Fig. 2 Index values for owner-occupied homes in The Netherlands and 95% confidence interval

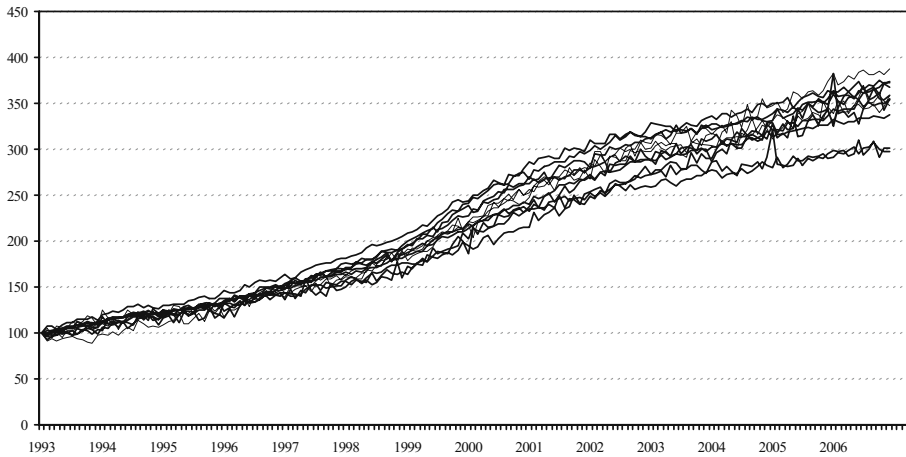


Fig. 3 Index values for the 12 provinces of The Netherlands

Another method to explore heteroskedasticity is the Breusch–Pagan test. For this test, the squared residuals are divided by the sum of the residuals that is divided by the number of observations (in, e.g., Greene 1993, p. 395):

$$z_i^2 = u_i^2 / s^2 \quad (5)$$

$$s^2 = \sum u_i^2 / n \quad (6)$$

where i relates to the observations, u_i^2 to the squared residuals and n relates to the number of cases. Next, a regression analysis is performed on the transformed residuals. In the context of the Breusch–Pagan test, a Lagrange multiplier test can be calculated (in, e.g., Greene 1993, p. 394). The results of this test show that heteroskedasticity is present in the data for all 13 indices. Note, however, that for all indices the amount of explained variance in the regression analysis does not exceed 1%.

Thus far, we explored in general whether heteroskedasticity is present in the data. However, Case and Shiller argue that the heteroskedasticity is related to the holding period. To test this assumption, we made a scatter plot of the residuals against the holding period. We did not find the suggested form in which the variance widens out with time. In fact, the figure suggested the opposite, i.e., that the spread of the residuals would decrease with longer periods between sales. The Breusch–Pagan test indicated heteroskedasticity in the data but, again, the percentage of explained variance was in all cases less than 1%.

We also performed the Goldfeld–Quant Test (see Greene 1993, p. 394). This test is based on the assumption that the sample consists of various groups with different residuals. The holding period ranges from 12 to 168 months. In accordance to the Goldfeld–Quant Test, we made three groups of almost similar group size. Next, we performed the first step of the Repeat Sales regression-analysis in the first and third group separately and compared the amount of squared residuals in both groups. The tests showed that heteroskedasticity was indeed present.

Related to the problem of heteroskedasticity, we encountered a problem with regard to the estimated variance in the second step of the procedure. For example, for the national index, we observed a value of the coefficients for period of 0.0016693 and for period squared of -0.0000080 (see Tables 4 and 5, method Calhoun). Based on these coefficients the squared residuals are estimated:

$$\hat{d}_i^2 = 0.0016693 \cdot t + -0.0000080 \cdot t^2 \quad (7)$$

where t relates to holding period. We calculated a graph of the estimated squared residuals and observed that they increased with a longer holding period and that this increase leveled off as assumed. However, when the holding period is about 107 months, the estimated variance starts to decrease. This means that the weighing procedure in the third step is at stake. Cases are weighted on the basis of the value of the estimated squared residuals, to correct for the heteroskedasticity that is the result of the length of the holding period (according to the theory). The assumption is that cases with longer periods between sales should obtain less weight in the regression analyses. However, cases with a holding period of more than 107 months will now obtain more weight in the analysis instead of less weight. This effect was also observed for the indices of the individual provinces. The point where the estimated variance starts to decrease ranges from 91 to 159 months. A solution to this problem would be to keep the variance constant from the point where the variance starts to decrease. For the national index, we examined whether this finding was dependent upon the number of periods. However, irrespective of whether we calculated a monthly, quarterly, semi-annual or annual index, the decrease in estimated variance took place at about 107 months.

Confidence Intervals and Accuracy

The Repeat Sales Model requires a large number of repeat sales in a market segment to yield reliable estimates. Segmentation according to region, province and type of dwelling will reduce the number of repeat sales upon which the index is based. The accuracy of the measured estimates depends on the sample size, the distribution of the parameter scores in the population (standard error) and the level of confidence considered. A 95% confidence interval was used for the Woningwaarde Index Kadaster, because it is the most commonly used value and because it offers the best compromise between a high level of confidence on the one hand and a high level of accuracy on the other.

We determined the accuracy of an index on the basis of the 95%—confidence interval around the estimated index value. The estimated index value I_t is calculated as follows (Calhoun 1996):

$$I_t = 100 \cdot e^{\hat{\beta}_t} \quad (8)$$

in which $\hat{\beta}_t$ is the estimated coefficient from the ‘generalized least squares’ regression analysis. The standard error of the index figures thus derived is calculated as follows (Calhoun 1996):

$$\sigma_{I_t} = I_t \cdot \sigma_{\hat{\beta}_t} \quad (9)$$

in which σ_{I_t} is the standard error of the index figure for period t ; I_t is the index figure for period t ; and $\sigma_{\hat{\beta}_t}$ relates to the standard error of the estimated coefficient from the third step of the ‘generalized least squares’ regression analysis.

The borders of the confidence interval (CI) can then be calculated by combining the standard error with the common procedure for obtaining the 95% confidence interval (Cohen et al. 2003).

$$\text{UpperCI}_t = I_t + (1.96 * \sigma_{I_t}) \quad (10)$$

$$\text{LowerCI}_t = I_t - (1.96 * \sigma_{I_t}) \quad (11)$$

The distance between the upper and lower border indicates the width of the confidence interval (Wci). To determine the accuracy per period, the width of the confidence interval for the Woningwaarde Index Kadaster was then divided by the value of the index itself and multiplied by 100:

$$\text{Accuracy} = (Wci_t / I_t) * 100 \quad (12)$$

We found no indications in the literature on how narrow a confidence interval had to be in order to be described as ‘accurate.’ Nor was there any consensus on the minimum required accuracy of a sample. Table 6 shows the actual number of repeat sales, the mean standard error (i.e., the mean over all 168 periods) and the accuracy of the national index and the indices for the provinces. The mean actual standard error (SE) was calculated by taking the average of the standard errors of the 168 index values (I_t) for the various months. The results show that the accuracy ranges between 2 and 18%, which we believe is acceptable.

Table 6 Actual and needed number of repeat sales, actual and needed standard error, accuracy, and revision volatility for the national index and the 12 indices for the provinces

Index	Actual			Needed		Revision volatility “December 05–December 06”
	<i>n</i>	Mean SE	Accuracy (%)	Mean SE	<i>n</i>	Mean percent change (%, range)
The Netherlands	678,194	1.2	2.1	5.7	31,006	−0.23 (−0.82–0.17)
Groningen	25,138	7.0	12.7	5.2	40,572	0.03 (−3.36–2.28)
Friesland	27,415	8.0	13.8	5.8	52,438	−1.03 (−4.45–0.93)
Drenthe	21,701	7.7	13.6	5.7	40,029	0.99 (−2.54–3.00)
Overijssel	41,673	5.2	9.1	5.7	34,244	−0.07 (−3.17–1.60)
Flevoland	18,214	9.2	18.0	5.2	58,248	0.14 (−2.31–6.02)
Gelderland	71,152	3.7	6.0	6.2	25,465	−0.66 (−1.60–0.61)
Utrecht	58,384	4.3	7.6	5.7	33,556	0.06 (−1.36–1.16)
Noord-Holland	92,077	3.2	5.3	5.9	25,975	−0.21 (−1.33–0.62)
Zuid-Holland	168,077	2.1	3.8	5.5	24,964	−0.13 (−1.02–0.56)
Zeeland	16,361	7.8	14.6	5.3	34,786	−0.09 (−4.56–1.60)
Noord-Brabant	99,037	3.3	5.5	5.9	30,574	−0.02 (−1.02–0.97)
Limburg	38,965	4.0	7.6	5.3	22,408	0.10 (−1.72–1.55)

Minimum Number of Repeat Sales

Related to the topic of confidence intervals is the number of pairs of cases needed to obtain an accurate estimate. For example, the OFHEO House Price Index is published only if at least 1,000 homes are sold in the region (Calhoun 1996) and at least ten houses are sold per quarter.

However, it is possible to determine the minimum sample size that is needed to obtain acceptable values for the standard error and the confidence interval. We determined a minimum number of repeat sales by applying the following formula (Cohen et al. 2003):

$$n^* = n \left(\frac{SE}{SE^*} \right)^2, \quad (13)$$

in which n^* is the minimum sample size needed; n is the original sample size; SE is the original standard error; and SE^* is the desired standard error.

The desired standard error (SE^*) can be calculated. If we calculate SE^* on the basis of 10% accuracy, the SE^* for The Netherlands as a whole is 5.7. By applying Eq. 13, the minimum needed number of repeat sales (n^*) is:

$$n^* = 678,194 \left(\frac{1.223813996}{5.723631797} \right)^2 = 31,006 \text{ cases.}$$

Table 6 shows the actual and needed numbers for the 15 indices published by the Dutch Land Registry Office, based on 10% accuracy. The table shows that the number of pairs of repeat sales needed to calculate an accurate index is quite different for the various segmentations (range 22,408–58,248). The accuracy of the measurement depends besides on the size of the sample also on the distribution of the parameter scores in the population (standard error). Thus, more homogeneous sub samples will require fewer cases. The picture that emerges does not justify a minimum number of observations, as applied, for example, by the OFHEO. The table also shows that, for a chosen accuracy of 10%, five provinces would have an actual number of cases that is lower than the needed number of cases.

Effect of Revisions: Revision Volatility

According to Bailey et al. (1963), the Repeat Sales Model is more efficient than other methods because it utilizes information about the price index for earlier periods that is contained in sales prices in later periods. Thus, the index values gain precision. Similarly, Shiller (1991) argues that such a revision is the result of increased efficiency in the estimators. However, present-day information changes the past values of the index (Baroni et al. 2004). Thus, additional sales have implications for the index values because new pairs will provide additional information about changes in the price level beyond that obtained from the previous sample. This is termed revision volatility and it may induce problems to the interpretability of the index, as the new index values may not be similar to the old ones. Clapp and Giacotto (1999) showed that revisions may be large, insensitive to sample size, and systematically downwardly directed. Clapp and Giacotto (1999) observed that

properties with only 1 or 2 years between sales (so-called ‘flips’) appreciate at a higher rate than other properties and may therefore be partly responsible for the downward revision of the index. Abraham and Schauman (1991) argue that in periods of weak real estate markets, most of the properties that do trade will be the strongest performers within the market (‘winners’). An index based on these transactions will therefore overstate the rate of property appreciation. However, eventually the preliminary estimates of price appreciation will be revised downwards as the sample expands from the price information for properties held, but not sold, during that period (Abraham and Schauman 1991).

To obtain an impression of the scale of these changes for the Woningwaarde Index Kadaster, we calculated the index values with all the data up to December 2005, and again with all the data up to December 2006 (thus with 12 additional months) for all previously described 13 indices. The mean percentage change and the range for every index is presented in Table 6, last column. The results show that the volatility of the coefficients is usually small when data are added for 12 additional months. The mean percentage change for The Netherlands is $-0,23\%$, thus less than a quarter of a percent. The largest mean revision is 1.3% and is observed for the province of Friesland. The largest individual revision is observed for the province of Flevoland with a value of 6% . Seven hundred and eighty-eight revisions (39%) are directed upwards and 1,227 downwards (61% ; total=13 indices \times 155 months=2,015 potential revisions).

Concluding Remarks

After a thorough literature study and based on the characteristics of our data set (very large but without property characteristics) and the target of our study (a geometric mean index value), we chose the Weighted Repeat Sales method to calculate monthly indices for house prices in The Netherlands.

One major benefit of the (Weighted) Repeat Sales Model is that it theoretically removes quality differences between packages of homes sold in different periods (Bailey et al. 1963). It so distinguishes differences in quality from differences in price (Abraham and Schauman 1991). All the characteristics that could be included in a hedonic regression analysis or in a hybrid method are corrected (theoretically) by the Repeat Sales Model (Abraham and Schauman 1991). By comparing the same dwelling over time, the procedure also corrects for the possibility of a progressive improvement in quality in new-built houses. (Bailey et al. 1963). However, the index is only corrected for quality if properties retain the same physical attributes and if these attributes are accorded the same value by the market over time (Stephens et al. 1995). It is highly plausible that, for some dwellings, the characteristics will be different on the two dates of sale. This would then undermine one of the assumptions that makes for consistency in the repeat-sales approach. On the one hand, houses may depreciate through time, either physically or because of new tastes and fashions. On the other hand, they may have been modernized and upgraded, thereby gaining in value.

However, for estimating the risk of their mortgage portfolio, banking institutes in The Netherlands are interested only in the current value of houses in their portfolio.

According to Hwang and Quigley (2004), the quality change issue is not relevant if an index is intended to measure the market value of dwellings transacted in a given time interval. Similarly, Wang and Zorn (1997) argue that researchers looking for an estimate of the change in the value of housing—as we are—may prefer to include the impact of improvements and depreciation in their indices. For this reason, this disadvantage of the Repeat Sales method seems less important for our application of the Woningwaarde Index Kadaster.

We observed that in the second step of the Weighted Repeat Sales Regression Analysis the coefficient for holding period squared is positive instead of negative for 11 out of 13 indices. A similar observation was also made by Clapp and Giaccotto (1999), who found that the coefficient for period squared was positive in all six combinations of region (Fairfax and Los Angeles) and sample size (three different sample sizes for each region) that they analyzed. These findings contradict the assumption of Abraham and Schauman (1991) that the increase in the variance of the residuals will decrease as the holding period increases. Furthermore, the coefficient for holding period was negative for six indices, indicating that there is a negative effect of holding period on the amount of variance. This is also contradictory to the theory. These results call into question the suggested form of the diffusion of the variance of appreciation rates over time. Another argument against the current use of the second step of the Weighted Repeat Sales procedure is our finding that the proposed heteroskedasticity cannot be conclusively demonstrated in the data. Tests show that heteroskedasticity seems to be present, but the amount of explained variance is less than 1%. Significant results may have been the result of the large sample size. Furthermore, we observed a problem with the weights necessary to correct for heteroskedasticity in the third step of the procedure.

With the highest value of 18%, the accuracy of the 13 indices was reasonably acceptable. However, accuracy may become a problem with smaller sub samples. We have no gold standard of which level of accuracy is still acceptable.

We observed that the revision volatility observed for the Woningwaarde Index Kadaster was reasonably small and acceptable. Whereas most of the revisions are downwards directed, even after removing the ‘flips’, it seems that excluding transactions with a holding period of less than 12 months may not be sufficient. In a previous study, Hoesli et al. (1997) examined the effect of revisions on the index. Because they did not observe statistically significant systematic deviations in the revisions, they concluded that each of the original indices is unbiased and that the revised index is a more efficient estimator of the price level. Abraham and Schauman (1991) found similar results. They conclude that while there is a fair bit of volatility in the indices, transactions-bias (responsible for revision volatility) does not appear to be a problem, even down at the city level.

Finally, note that we performed Chow tests to explore whether the data for the 12 separate provinces could be pooled for the calculation of the national index. Our results showed statistically significant differences, indicating that the data could not be pooled. But, the number of cases used for calculating these statistics is so large that significant results will quickly be found, even in the absence of relevant practical differences. In fact, sensitivity analyses showed that in random samples of the data with a sample size of about 12% at largest, the Chow test consistently did not indicate significant differences between provinces. For this reason, we decided

that we could calculate a national index but that we have to keep in mind that the house price development in separate provinces can deviate from the national tendency. This is a limitation to the national index.

In conclusion, given the characteristics of the available dataset and our target, the Repeat Sales Model seems to be an adequate method for calculating a house price index for The Netherlands.

References

- Abraham, J. M., & Schauman, W. S. (1991). New evidence on home prices from Freddie Mac repeat sales. *AREUEA Journal*, 19(3), 333–352.
- Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, 58, 933–942.
- Baroni, M., Barthelemy, F., & Mokrane, M. (2004). Physical real estate: A Paris repeat sales residential index. Working paper, DR04007 (pp. 1–18). Cergy, France: ESSEC, Research Center.
- Bourassa, S. C., Hoesli, M., & Sun, J. (2006). A simple alternative house price index method. *Journal of Housing Economics*, 15(1), 80–97.
- Calhoun, C. A. (1996). OFHEO house price indexes: HPI technical description. Available at: http://www.ofheo.gov/Media/Archive/house/hpi_tech.pdf, 1–14.
- Case, K. E., & Shiller, R. J. (1987). Prices of single-family homes since 1970: New indexes for four cities. *New England Economic Review*, (Sep), 45–56.
- Case, K. E., & Shiller, R. J. (1989). The efficiency of the market for single-family homes. *The American Economic Review*, 79(1), 125–137.
- Clapp, J. M., & Giacotto, C. (1999). Revisions in repeat-sales price indexes: Here today, gone tomorrow? *Real Estate Economics*, 27(1), 79–104.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum.
- Costello, G., & Watkins, C. (2002). Towards a system of local house price indices. *Housing Studies*, 17(6), 857–873.
- Dreiman, M. H., & Pennington-Cross, A. (2004). Alternative methods of increasing the precision of weighted repeat sales house price indices. *Journal of Real Estate Finance and Economics*, 28(4), 299–317.
- Englund, P., Quigley, J. M., & Redfearn, C. L. (1998). Improved price indexes for real estate: Measuring the course of Swedish housing prices. *Journal of Urban Economics*, 44(2), 171–196.
- Goetgeluk, R. (1997). *Bomen over wonen, woningmarktonderzoek met beslissingsbomen*. Dissertation, Universiteit Utrecht, Faculteit Ruimtelijke Wetenschappen/KNAG, Utrecht Geographical Studies, Utrecht, p. 235.
- Goetzmann, W. (1992). The accuracy of real estate indices: Repeat sales estimators. *Journal of Real Estate Finance and Economics*, 5(1), 5–53.
- Greene, W. H. (1993). *Econometric analysis* (2nd ed.). New Jersey: Prentice-Hall.
- Hoesli, M., Giacotto, C., & Favarger, P. (1997). Three new real estate price indices for Geneva, Switzerland. *Journal of Real Estate Finance and Economics*, 15(1), 93–109.
- Hwang, M., & Quigley, J. M. (2004). Selectivity, quality adjustment and mean reversion in the measurement of house values. *Journal of Real Estate Finance and Economics*, 28(2/3), 161–178.
- Jansen, S., de Vries, P., Boelhouwer, P., Coolen, H. C. C. H., Lamain, C. J. M., & Marien, A. A. A. (2005). Methodologie Woningwaarde-Index Kadaster. Research report. Delft: OTB Research Institute for Housing, Urban and Mobility Studies.
- Masser, I., & Scheurwater, J. (1978). The specification of multi-level systems for spatial analysis. In I. Masser & J. Scheurwater (Eds.), *Spatial representation and spatial interaction. Studies in applied science*, 10. Leiden, Boston: Martinus Nijhoff.
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55.
- Shiller, R. J. (1991). Arithmetic repeat sales price estimators. *Journal of Housing Economics*, 1(1), 110–126.
- Steele, M., & Goy, R. (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(1/2), 133–154.

- Stephens, W., Li, Y., Lekkas, V., Abraham, J., Calhoun, C., Kimner, T. (1995). Conventional mortgage home price index. *Journal of Housing Research*, 6, 389–418.
- Van Kempen, R., Goetgeluk, R., & Floor, H. (1995). *De randstad uit? Achtergronden bij het verhuizen en willen verhuizen van randstedelingen*. Universiteit Utrecht: Faculteit Ruimtelijke Wetenschappen.
- Von Dewall, F. A., Fleming, D. J. C., & Pallada, F. W. M. (2004). Een geïntegreerde prijsindex voor de markt van koopwoningen. *Kwartaalschrift Economie*, 1(4), 386–404.
- Wang, F. T., & Zorn, P. M. (1997). Estimating house price growth with repeat sales data: What's the aim of the game? *Journal of Housing Economics*, 6(2), 93–118.