

An Appraisal-Based Generalized Regression Estimator of House Price Change

Jan de Haan and Rens Hendriks¹

Revised draft, 14 August 2012

Abstract: The house price index compiled by Statistics Netherlands relies on the Sale Price Appraisal Ratio (SPAR) method. The SPAR method combines selling prices with prior government assessments of properties. This paper outlines an alternative approach where the appraisals serve as auxiliary information in a generalized regression (GREG) framework. An application on Dutch data demonstrates that, although the GREG index is much smoother than the ratio of sample means, it is very similar to the SPAR series. To explain this result we show that the SPAR index is an estimator of our more general GREG index and in practice almost as efficient.

Key words: generalized regression estimation, house price index, property assessments, sampling.

¹ Jan de Haan, OTB Research Institute for the Built Environment, Delft University of Technology and Division of Process Development, IT and Methodology, Statistics Netherlands, P.O. Box 24500, 2490 HA The Hague, The Netherlands. E-mail: j.dehaan@cbs.nl; Rens Hendriks, Division of Economic and Business Statistics and National Accounts, Statistics Netherlands. E-mail: r.hendriks@cbs.nl.

1. Introduction

When attempting to construct constant-quality house price indexes, statistical agencies face a number of problems. First, exact matching of properties over time is problematic as their quality will likely have changed; houses depreciate and they may also have had major repairs, additions or remodelling done to them. In other words, every property in each period can be viewed as a unique good. Second, the turnover of houses is generally low compared to the housing stock and the mix of properties sold changes over time, so a quality mix problem arises. Third, there is often a lack of data on characteristics. Data availability issues have implications for the choice of measurement method.

Three main types of house price indexes can be found in the literature: median or mean indexes, repeat sales indexes and hedonic indexes. A median (mean) index tracks the change in the price of the median (mean) house traded from one period to the next. This method is problematic in that the characteristics of, e.g., the median house changes over time. The problem is often tackled by stratifying the samples according to region, type of dwelling, etc., a procedure which is also known as *mix adjustment*. Stratification obviously requires additional data.

Repeat sales methods address the quality mix problem by restricting the data set to houses that have been sold twice or more during the sample period. This ensures that 'like is compared with like', assuming that the quality of the individual houses remains unchanged. Repeat sales methods are based on regressions where the repeat sales data pertaining to different periods are pooled. A potential drawback is revision; when new data is added to the sample, previously computed index numbers will change. The repeat sales method is originally due to Bailey, Muth and Nourse (1963). Case and Shiller (1987, 1989) argue that changes in house prices include components whose variances increase with the interval of sales and propose a Weighted Least Squares approach to adjust for this type of heteroskedasticity. An alternative weighted method has been suggested by Calhoun (1996). Jansen *et al.* (2008), using Dutch data, compare the unweighted repeat sales method with various weighted methods and conclude that the unweighted method performs satisfactorily.

Unlike repeat sales methods, *hedonic regression methods* can in principle adjust for quality changes of individual properties (in addition to quality mix changes). These methods utilize information on housing characteristics, such as number of bedrooms, lot size and location, to estimate quality adjusted price indexes using regression techniques.

Today, hedonic house price indexes are computed in many countries. For example, the French statistical agency (INSEE), jointly with Conseil Supérieur du Notariat, compiles a hedonic index (Gouriéroux and Laferrère, 2009) as does Statistics Finland (Saarnio, 2006). The UK has three hedonic house price indexes, compiled by different institutes. RPData-Rismark computes hedonic indexes for the capital cities in Australia (Hardman, 2011). Hedonic indexes come in two main varieties. The time dummy method models the log of price as a function of property characteristics and a set of dummy variables indicating the time periods. Since the data of all periods are pooled, this method suffers from revision as well. Hedonic imputation methods, which estimate the ‘missing prices’, do not have this drawback. Hill and Melser (2008) discuss numerous hedonic imputation methods in the housing context. Diewert, Heravi and Silver (2009) and de Haan (2010) provide a comparison between time dummy and hedonic imputation price indexes.

A fourth approach to estimating house price indexes is the *use of assessment or appraisal data*. One option is to augment a repeat sales dataset by using assessment data as estimates for past or current values of properties that have not been resold during the sample period. Some of the data on which the repeat sales index is based would then be pseudo rather than genuine repeat data. For more on the use of assessment information in a repeat sales price index and the removal of appraisal bias, see e.g. Geltner (1996), Edelstein and Quan (2006), and Leventis (2006). Another option, which also controls for quality-mix changes, is to combine current selling prices with appraisals from an earlier period to compute price relatives in a standard matched-model framework. An advantage over the repeat sales approach is that index numbers will not be revised. This so-called Sale Price Appraisal Ratio (SPAR) method has been applied in New Zealand for a long time now and is currently also being used in the Netherlands and a few other European countries. Bourassa, Hoesli and Sun (2007) describe the New Zealand SPAR index which is compiled by Quotable Value, a state-owned property valuation company. Other studies into the SPAR method include Rossini and Kershaw (2006), van der Wal, ter Steege and Kroese (2006), de Vries *et al.* (2009), de Haan, van der Wal and de Vries (2009), Shi, Young and Hargreaves (2009), and Grimes and Young (2010).

In this paper we outline an alternative appraisal-based method to measure house price change. The appraisals serve as auxiliary information in a *generalized regression* (GREG) estimation framework. GREG is a model-assisted technique that can be used to

increase efficiency as compared to simpler estimators such as sample means (Särndal, Swensson and Wretman, 1992), provided that population information is known for one or more variables that exhibit a strong linear correlation with the variable under study. In our case we regress selling prices in each time period on appraisals. Appraised values are available in the Netherlands for all properties in stock in some reference period, and we expect them to be highly collinear with selling prices. Although the method is based on regression, the resulting price index is not a hedonic index as the regression model is descriptive rather than explanatory.

The paper is organized as follows. To set the stage, in Section 2 we describe the SPAR method and its relation to the sample means of sale prices and appraisals. Due to compositional change and the relatively low number of transactions, the Dutch SPAR series exhibits strong volatility, especially for small market segments. In Section 3 we outline a simple GREG estimator of house price change and two alternatives. The first alternative is a stratified version of the original index whereas the second one uses an alternative model specification. Section 4 contains empirical evidence using Dutch data. The GREG index numbers turn out to be very similar to the SPAR index numbers and are equally volatile. In Section 5 we explain this result by showing that the SPAR index is in fact an estimator of the GREG index and almost as efficient. Section 6 concludes and suggests a topic for further research in this field.

2. Horvitz-Thompson Estimators and the SPAR Index

The typical aim of survey sampling is to estimate the total or (arithmetic) mean of some variable for a finite population. In a housing context we may want to estimate the total value of the housing stock in, say, period 0. Let U^0 denote the housing stock of size N^0 and p_n^0 the value of house n ($n = 1, \dots, N^0$). The target to be estimated is

$$V^0 = \sum_{n \in U^0} p_n^0. \quad (1)$$

Suppose we have a sample S^0 consisting of n^0 houses sold in the base period. If the houses were selected by simple random sampling from the housing stock U^0 , where each house had the same inclusion probability, then the Horvitz-Thompson estimator

$$\hat{V}^0 = (N^0 / n^0) \sum_{n=1}^{n^0} p_n^0 \quad (2)$$

is an unbiased estimator of (1); see e.g. Cochran (1977).

A natural target – though not the only possibility – for a house price index would be the value change of a fixed housing stock. Conditioning on the *base period stock* has two implications: additions to the stock (mostly newly-built houses) should be excluded and the price changes of existing properties should be adjusted for quality changes, i.e. for the impact of depreciation, renovations and extensions. For convenience we assume that such quality changes are negligible. In that case the target price index going from the base period 0 to the comparison period t (> 0) is defined as

$$P^{0t} = \frac{\sum_{n \in U^t} p_n^t}{\sum_{n \in U^0} p_n^0}, \quad (3)$$

with obvious notation. Suppose that we also have a sample S^t , consisting of n^t houses sold in period t and assume that it is an independent random draw from the base period stock. The ratio of the Horvitz-Thompson estimators (the sample means) in both periods

$$\hat{P}^{0t} = \frac{(N^0 / n^t) \sum_{n \in S^t} p_n^t}{(N^0 / n^0) \sum_{n \in S^0} p_n^0} = \frac{\sum_{n \in S^t} p_n^t / n^t}{\sum_{n \in S^0} p_n^0 / n^0} \quad (4)$$

might seem a natural estimator of our target index (3). However, if the samples S^0 and S^t are independently drawn, the variance of estimator (4) can be substantial. Moreover, an estimated ratio such as (4) has a bias that depends on the variance of the numerator and the covariance of the numerator and the denominator (Cochran, 1977). From an index number perspective the issue at stake is that the mix of properties traded in period t differs from that in period 0. That is, we are not comparing like with like.

The standard approach to estimating price indexes relies on the matched model methodology where prices p_n^0 and p_n^t are observed for a fixed panel of items. The use of panel data ensures that like is compared with like and will reduce the variance of the ratio estimator because p_n^0 and p_n^t are typically positively correlated. However, unless the samples S^0 and S^t are extraordinary large, there will only be few matched houses, if any. Hence, while prices p_n^t are observed for the houses belonging to S^t , for most of those houses the base period prices p_n^0 are ‘missing’. What may be available instead are government assessments a_n^0 . We could use these as base period values and construct the following (pseudo) matched-model estimator of house price change:

$$\tilde{P}^{0t} = \frac{\sum_{n \in S^t} p_n^t / n^t}{\sum_{n \in S^t} a_n^0 / n^t}. \quad (5)$$

A problem associated with estimator (5) is that the base period index number will differ from 1 because the appraisals a_n^0 differ from the selling prices p_n^0 . Rescaling (5) by dividing it by its base period value is an obvious solution, yielding

$$\hat{P}_{SPAR}^{0t} = \frac{\sum_{n \in S^t} p_n^t / n^t}{\sum_{n \in S^t} a_n^0 / n^t} \left[\frac{\sum_{n \in S^0} p_n^0 / n^0}{\sum_{n \in S^0} a_n^0 / n^0} \right]^{-1} = \frac{\sum_{n \in S^t} p_n^t / n^t}{\sum_{n \in S^0} p_n^0 / n^0} \left[\frac{\sum_{n \in S^0} a_n^0 / n^0}{\sum_{n \in S^t} a_n^0 / n^t} \right]. \quad (6)$$

Note that the rescaling factor is stochastic, as it is a ratio of sample means for the base period, and will increase the variance of (6) as compared to the estimator given by (5), depending on the correlations between the appraisals and the selling prices. Details can be found in de Haan (2007). But we cannot circumvent rescaling since a price index that does not start at the value 1 would be meaningless.

Expression (6) is called a Sale Price Appraisal Ratio (SPAR) index. The SPAR method has been applied in the Netherlands since January 2008 to measure the price change of owner-occupied dwellings. As mentioned earlier, we assume that the SPAR index aims at tracking the price change of the *housing stock*, which is a measure of the change in wealth. In the context of the Harmonized Index of Consumer Prices on the other hand, the house price index should measure the price change of the *houses sold* during the base period (Makaronidis and Hayes, 2006; Eurostat, 2010). Under the latter concept there would be no sampling involved if all transactions are recorded and used in the compilation of the index, as is the case in the Netherlands.

The second expression on the right-hand side of (6) writes the SPAR index as the product of two factors, the ratio of sample means and a factor between brackets. As the SPAR index is essentially based on the matched model methodology (using base period appraisals instead of sale prices), this factor adjusts the ratio of sample means for changes in the quality mix of the samples that occur between period 0 and period t . A potential problem is that the SPAR index is *not a panel-type estimator*. A SPAR time series, say for periods $t = 0, \dots, T$, might therefore suffer from short-term volatility due to mix changes, especially when the number of sales is low.

3. Generalized Regression Estimation

3.1 A Simple GREG Method

In this section we will outline an alternative approach to measuring house price change that makes use of appraisal data. The appraisals now serve as auxiliary information in a generalized regression (GREG) framework. Consider the following simple two-variable linear regression model:

$$p_n^0 = \alpha^0 + \beta^0 a_n^0 + \varepsilon_n^0, \quad (7)$$

where ε_n^0 is the error term. Unlike hedonic regression models, which postulate a causal relation between the selling price p_n^0 and a set of characteristics relating to the structure and the location of the housing units, this model does not say anything about how house prices are generated; equation (7) is merely a descriptive model.

Estimating model (7) by least squares regression on the data of sample S^0 yields predicted prices

$$\hat{p}_n^0 = \hat{\alpha}^0 + \hat{\beta}^0 a_n^0. \quad (8)$$

The regression residuals for $n \in S^0$ are $e_n^0 = p_n^0 - \hat{p}_n^0$. Assuming random sampling, as before, we can write the Horvitz-Thompson estimator $\sum_{n \in S^0} p_n^0 / n^0$ of the mean value $\sum_{n \in U^0} p_n^0 / N^0$ as

$$\sum_{n \in S^0} p_n^0 / n^0 = \sum_{n \in S^0} \hat{p}_n^0 / n^0 + \sum_{n \in S^0} e_n^0 / n^0 = \hat{\alpha}^0 + \hat{\beta}^0 \sum_{n \in S^0} a_n^0 / n^0 + \sum_{n \in S^0} e_n^0 / n^0. \quad (9)$$

Replacing the sample average of appraisals, $\sum_{n \in S^0} a_n^0 / n^0$, by its population counterpart $\sum_{n \in U^0} a_n^0 / N^0$ yields the generalized regression (GREG) estimator:

$$\hat{p}_{GREG}^0 = \hat{\alpha}^0 + \hat{\beta}^0 \sum_{n \in U^0} a_n^0 / N^0 + \sum_{n \in S^0} e_n^0 / n^0 = \sum_{n \in U^0} \hat{p}_n^0 / N^0 + \sum_{n \in S^0} e_n^0 / n^0. \quad (10)$$

Model-assisted sampling theory shows that GREG estimators are *asymptotically design unbiased* (Särndal, Swensson and Wretman, 1992), irrespective of the choice of regressors. Unless the sample would be small, the bias can be neglected. It is obvious that the GREG estimator (10) will be more efficient – in the sense that it has a lower variance – than the Horvitz-Thompson estimator (9). As a result, the GREG estimator will usually outperform the Horvitz-Thompson estimator in terms of the mean square error (the sum of the variance and the squared bias).

The same procedure can be applied to the comparison period t . After estimating the model

$$p_n^t = \alpha^t + \beta^t a_n^0 + \varepsilon_n^t \quad (11)$$

through least squares regression on the data of the current period sample S^t , we obtain predicted prices

$$\hat{p}_n^t = \hat{\alpha}^t + \hat{\beta}^t a_n^0, \quad (12)$$

which lead to the GREG estimator of the mean value of the housing stock in period t :

$$\hat{p}_{GREG}^t = \hat{\alpha}^t + \hat{\beta}^t \sum_{n \in U^t} a_n^0 / N^t + \sum_{n \in S^t} e_n^t / n^t = \sum_{n \in U^t} \hat{p}_n^t / N^t + \sum_{n \in S^t} e_n^t / n^t, \quad (13)$$

where $e_n^t = p_n^t - \hat{p}_n^t$ denote the period t regression residuals. For a fixed housing stock we have $U^t = U^0$, hence $\sum_{n \in U^t} a_n^0 / N^t = \sum_{n \in U^0} a_n^0 / N^0$, and it follows that

$$\hat{p}_{GREG}^t = \hat{\alpha}^t + \hat{\beta}^t \sum_{n \in U^0} a_n^0 / N^0 + \sum_{n \in S^t} e_n^t / n^t = \sum_{n \in U^0} \hat{p}_n^t / N^0 + \sum_{n \in S^t} e_n^t. \quad (14)$$

The GREG estimator of house price change results simply from taking the ratio of equations (14) and (10):

$$\hat{p}_{GREG}^{0t} = \frac{\hat{p}_{GREG}^t}{\hat{p}_{GREG}^0} = \frac{\hat{\alpha}^t + \hat{\beta}^t \bar{a}^0 + \sum_{n \in S^t} e_n^t / n^t}{\hat{\alpha}^0 + \hat{\beta}^0 \bar{a}^0 + \sum_{n \in S^0} e_n^0 / n^0} = \frac{\sum_{n \in U^0} \hat{p}_n^t / N^0 + \sum_{n \in S^t} e_n^t / n^t}{\sum_{n \in U^0} \hat{p}_n^0 / N^0 + \sum_{n \in S^0} e_n^0 / n^0}, \quad (15)$$

where $\bar{a}^0 = \sum_{n \in U^0} a_n^0 / N^0$. Some additional small sample bias will be introduced due to the non-linear (ratio) structure. When using Ordinary Least Squares (OLS) regression to estimate the models (7) and (11), the unweighted sample means of regression residuals in (15), $\sum_{n \in S^0} e_n^0 / n^0$ and $\sum_{n \in S^t} e_n^t / n^t$, will be equal to 0 and the GREG index reduces to

$$\hat{p}_{GREG,OLS}^{0t} = \frac{\sum_{n \in U^0} \hat{p}_n^t / N^0}{\sum_{n \in U^0} \hat{p}_n^0 / N^0} = \frac{\hat{\alpha}^t + \hat{\beta}^t \bar{a}^0}{\hat{\alpha}^0 + \hat{\beta}^0 \bar{a}^0} = \frac{\hat{\alpha}^t / \bar{a}^0 + \hat{\beta}^t}{\hat{\alpha}^0 / \bar{a}^0 + \hat{\beta}^0}. \quad (16)$$

As the first expression on the right-hand side of (16) indicates, the (OLS) GREG approach essentially imputes prices pertaining to the base period and the current period using equations (8) and (12). The difference with the hedonic *double imputation* method is twofold: a descriptive model, not a hedonic one, is used to estimate predicted prices –

so that we cannot speak of unbiased predicted prices – and prices are imputed for all houses of the housing stock instead of the sub-set of sampled houses.

3.2 Properties of the GREG Index

The (OLS) GREG index has several properties worth mentioning. First, the computation of the GREG index is very simple. Once the population mean of appraisals \bar{a}^0 and the base period regression coefficients $\hat{\alpha}^0$ and $\hat{\beta}^0$ have been calculated, all that is needed is running a regression each month of selling prices against appraisals and plugging the coefficients $\hat{\alpha}^t$ and $\hat{\beta}^t$ into (16). Note that the GREG index can be written as a pseudo chain index:

$$\hat{P}_{GREG,OLS}^{0t} = \frac{\hat{\alpha}^t / \bar{a}^0 + \hat{\beta}^t}{\hat{\alpha}^0 / \bar{a}^0 + \hat{\beta}^0} = \prod_{\tau=1}^t \frac{\hat{\alpha}^\tau / \bar{a}^0 + \hat{\beta}^\tau}{\hat{\alpha}^{\tau-1} / \bar{a}^0 + \hat{\beta}^{\tau-1}}. \quad (17)$$

This can be helpful in practice, particularly when new appraisal data becomes available. New appraisal data often becomes available to the statistical agency with a considerable time lag, up to more than a year. There are two reasons for using the latest appraisal information. The quality of the appraisals may improve over time, which seems to have been the case in the Netherlands (de Vries *et al.*, 2009). Also, the assumption of a fixed housing stock can be relaxed so that newly-built properties can be incorporated through chaining; the resulting chained GREG index takes the dynamics of the housing stock into account. The same advantages of chaining apply to the SPAR method. Suppose new appraisals, relating to period T ($0 < T \leq t$), are available in period $t+1$. The time series can then be updated through chain-linking, i.e. by multiplying $\hat{P}_{GREG,OLS}^{0t}$ by the month-to-month change $(\tilde{\alpha}^{t+1} / \bar{a}^T + \tilde{\beta}^{t+1}) / (\tilde{\alpha}^t / \bar{a}^T + \tilde{\beta}^t)$, where the coefficients now pertain to a regression of selling prices on the period T appraisals.

Second, *standard errors* of the GREG index can be estimated rather easily using the variance-covariance matrix of the regression coefficients, which is standard output of most statistical packages. An expression for the approximate standard error is derived in the Appendix. The standard error of the GREG index depends on the goodness of fit (R^2) of the regression model. It is most likely that R^2 for the base period regression is higher than that for the current period regressions. This is because we expect to find a strong linear relation between appraisals and sale prices in the appraisal reference period while in later periods this relation will probably be weaker due to differing price trends

across different types of houses or regions. The derivation of approximate standard errors for the SPAR index is a bit more complex because there is an additional source of sampling error, namely the sampling variability of the mean appraisals; see de Haan (2007).

The latter point brings us to the third property of the GREG index, namely its dependence on the *quality of the appraisal data*. For two reasons at least the appraisals may not exactly represent the transaction prices during the base period so that the model fit is not perfect ($R^2 < 1$). The assessment authorities may not have (real time) access to the actual sale prices and therefore have to make their own judgements based on other information. But even if they knew the selling prices, the authorities may still decide to make adjustments when determining the property values. It can be argued that selling prices do not always properly measure the unknown market values – which can be seen as a latent variable – and tend to be more volatile. In this respect, Francke (2010) and others have used the term transaction noise.

The way in which the appraisals have been determined will affect the standard error of the GREG index. As long as the quality of the appraisal data is the same for all houses in stock, no bias arises since the appraisals only serve as an auxiliary variable in regressions run on the samples S^0 and S^t of properties sold in periods 0 and t ($t = 1, \dots, T$). However, in general we expect the quality of the appraisals to be higher for properties belonging to the appraisal reference (base) period sample S^0 , although this will most likely differ across valuation methods. In the Netherlands the properties are assessed for tax purposes, both for income tax and local taxes. The municipalities are responsible for the valuations. Several municipalities value the houses which are sold during the reference period (January) by the selling price. Houses which were not sold are sometimes valued by comparing them to similar traded houses. Some municipalities apparently use a form of hedonic regression to value the houses, but the methodology is unfortunately not made publicly available. For more information on the Dutch appraisal system, see de Vries *et al.* (2009).

So far we have assumed that the quality of the individual houses stays the same over time. This is a strong assumption. Thus, the fourth property – and most important drawback – of the GREG method is that the resulting price index suffers from *quality change bias* since explicit quality adjustments are not carried out. The same drawback holds true for the SPAR method and for the standard repeat sales method. In principle,

hedonic regression methods can deal with the quality change problem, although it may prove difficult to control for all relevant price determining characteristics, in particular micro location. The SPAR method automatically controls for micro location, provided of course that the appraisals sufficiently account for this, as it is based on the matched-model methodology where the matching is done at the address level.

3.3 Alternative GREG Estimators

Statistics Netherlands not only computes house price indexes for the whole country but also for segments of the housing market, according to type of house (family dwellings and apartments) and region (provinces and large cities), mainly because of user needs. Another motivation behind stratifying the sample can be to mitigate the effect of *sample selection bias*. This type of bias may arise if the set of houses sold in a particular period is not a random selection from the housing stock. The nationwide index should then be indirectly computed as a weighted average of the stratum indexes instead of directly from all observations.

Suppose the total housing stock U^0 is sub-divided into K non-overlapping strata U_k^0 of size N_k^0 ($\sum_{k=1}^K N_k^0 = N^0$). The target price index (3) can now be rewritten as

$$P^{0t} = \frac{\sum_{n \in U^0} P_n^t}{\sum_{n \in U^0} P_n^0} = \frac{\sum_{k=1}^K \sum_{n \in U_k^0} P_n^t}{\sum_{k=1}^K \sum_{n \in U_k^0} P_n^0} = \sum_{k=1}^K s_k^0 P_k^{0t}, \quad (18)$$

where $P_k^{0t} = \sum_{n \in U_k^0} P_n^t / \sum_{n \in U_k^0} P_n^0$ is the target price index for stratum U_k^0 ($k = 1, \dots, K$). The base period stock value shares $s_k^0 = \sum_{n \in U_k^0} P_n^0 / \sum_{n \in U^0} P_n^0$, which serve as weights for the stratum indexes, are unknown and have to be estimated. Assuming the variables that define the strata are known for all $n \in U^0$, a natural choice for the weights would be the appraisal shares $\hat{s}_k^0 = \sum_{n \in U_k^0} a_n^0 / \sum_{n \in U^0} a_n^0 = (N_k^0 / N^0)(\bar{a}_k^0 / \bar{a}^0)$. Obviously, the stratum-defining housing variables should be included in the appraisal data set. In the Netherlands address and type of dwelling are included. This allows a sub-division of the population into cross classifications of location and type of dwelling. Appraisals may not always be accurate estimates of the ‘true’ market values of the individual properties but at the stratum level we expect the accuracy of the average appraisals to be sufficient for the computation of the weights.

Statistical techniques such as GREG estimation are typically applied to estimate totals or means for small domains for which the number of observations is so small that the standard errors using traditional (Horvitz-Thompson) estimators – in our case the ratio of sample means – would become unacceptably high. It should be mentioned that, even with the GREG method, the stratification scheme should not be too detailed since that might unduly raise the variance of the stratum indexes and hence of the aggregate index. More importantly perhaps, small sample bias will increase and may become non-negligible with very small samples.

OLS regressions of selling prices on appraisals should now be run in every time period for each stratum in order to compute the aggregate GREG index. The stratified (OLS) GREG index is

$$\hat{P}_{StrGREG}^{0t} = \sum_{k=1}^K \hat{S}_k^0 \hat{P}_{k,GREG,OLS}^{0t} = \sum_{k=1}^K \hat{S}_k^0 \left(\frac{\hat{\alpha}_k^t / \bar{a}_k^0 + \hat{\beta}_k^t}{\hat{\alpha}_k^0 / \bar{a}_k^0 + \hat{\beta}_k^0} \right); \quad (19)$$

Differences in the slope coefficients $\hat{\beta}_k^s$ ($s = 0, t$) across the strata could be the result of sampling error or reflect a real phenomenon. The latter can be of particular importance for periods t which are very distant from period 0 as different housing market segments tend to show varying price trends. Whether any differences in the slope coefficients reflect a real phenomenon could be tested.

An alternative model, to be estimated on the entire data set, is one with a single intercept term, but where the β 's are allowed to differ across the strata. Let $D_{n,k}$ be a dummy variable that has the value 1 if property n belongs to stratum k and 0 otherwise. In period s ($s = 0, t$) the model

$$p_n^s = \alpha^s + \sum_{k=1}^K \beta_k^s D_{n,k} a_n^0 + \varepsilon_n^s \quad (20)$$

is estimated by OLS regression on the data of the sample S^s , yielding predicted prices $\tilde{p}_n^s = \tilde{\alpha}^s + \tilde{\beta}_k^s a_n^0$ for $n \in U_k^0$. The residuals again sum to zero and the new (unstratified) OLS GREG index becomes

$$\tilde{P}_{GREG,OLS}^{0t} = \frac{\sum_{n \in U^0} \tilde{p}_n^t / N^0}{\sum_{n \in U^0} \tilde{p}_n^0 / N^0} = \frac{\sum_{k=1}^K \sum_{n \in U_k^0} \tilde{p}_n^t / N^0}{\sum_{k=1}^K \sum_{n \in U_k^0} \tilde{p}_n^0 / N^0} = \frac{\tilde{\alpha}^t + \sum_{k=1}^K \left(\frac{N_k^0}{N^0} \right) \tilde{\beta}_k^t \bar{a}_k^0}{\tilde{\alpha}^0 + \sum_{k=1}^K \left(\frac{N_k^0}{N^0} \right) \tilde{\beta}_k^0 \bar{a}_k^0}. \quad (21)$$

Model (20) is more flexible than the original model given by equations (7) and (11), and could be useful if the proportionality between sale prices and appraisals fails. Estimator (21) reduces to the original GREG index (16) if the $\tilde{\beta}_k^s$'s are all equal. In practice this will not happen, and (21) and (16) will give different answers. A common justification for the use of GREG estimators is that, being asymptotically unbiased, they are relatively *robust to model choice*. So we would expect the impact of the alternative model specification (21) to be moderate. On the other hand, it is well recognized in the literature that model dependence can be an issue under specific circumstances, notably when dealing with highly variable and outlier-prone populations. For example, Hedlin *et al.* (2001) stress the importance of a careful model specification search while Beaumont and Alavi (2004) focus on the treatment of outliers. It would therefore be worthwhile examining the effect of this alternative model specification.

4. Empirical Illustration

For the empirical study we used two data sets from different sources. The first data set contains the sale prices of nearly all transactions of existing houses (excluding newly-built houses) in the Netherlands between January 2003 and March 2009 as registered by the Dutch land registry office. The total number of observations amounts to 1,126,242 or approximately 15 thousand per month. The sales were recorded at the time the final agreement was made at the notary's office, on average six weeks after the preliminary sale was agreed on. The second data set contains the government appraisals, relating to January 2003, for all owner-occupied dwellings in the housing stock. Because addresses are available in both data sets, we know the sale price and the appraisal value for each transaction. Because the type of dwelling is also available, we were able to stratify by dwelling type and location.

The first thing we did was run unstratified OLS regressions of selling prices on appraisals, using model (14), for all 75 months. A selection of the results is listed in Table 1; detailed empirical material is available from the authors upon request. Not surprisingly, the coefficients $\hat{\beta}^t$ are different from zero at very low significance levels. In most cases the intercepts $\hat{\alpha}^t$ differ significantly from zero at the 5% level. Roughly 80 to 90% of the variation in selling prices is 'explained' by the variation in appraisals, as shown by the R^2 values. In other words, the correlation coefficient between selling

prices and base period appraisals ranges from 0.89 to 0.95. Figure 1 shows that R^2 diminishes slightly over time. As mentioned earlier, one of the reasons could be that different segments of the market exhibit different price changes. We were a bit surprised to find though that R^2 is not the highest in January 2003, being the appraisal reference period.

[Insert Table 1]

[Insert Figure 1]

Based on the above regression results, we computed GREG price index numbers according to equation (16). From January 2003 until mid 2008 house prices increased by some 25% in the Netherlands but then started to fall, probably due to the financial and economic crisis. Importantly, the GREG index turns out to be a lot smoother than the simple ratio of sample means as Figure 2 makes clear, which is precisely what the index has been designed for.

[Insert Figure 2]

Figure 3 compares the GREG index with the SPAR index. In general the trend of both indexes is very similar, although there appears to be a small difference by the end of the period. Figure 4 shows that the month-to-month changes in the GREG and SPAR indexes do not differ much either, the GREG index being just a little bit less volatile. So we can conclude that, at the nationwide level, both methods generate more or less equal results. Note that the SPAR index in Figures 3 and 4 is not the official SPAR index published by Statistics Netherlands. We computed a fixed base index using appraisals for January 2003 only whereas the official index is a chained index, based on appraisals for various reference periods; see also Section 5.3.

[Insert Figure 3]

[Insert Figure 4]

Next we stratified the data by thirteen provinces and five types of dwellings, ran OLS regressions per month for the resulting 65 strata and calculated GREG indexes as well as sample means ratios. Figure 5 displays the results for one stratum, apartments in the province of Friesland. Due to the relatively low number of observations there are some dramatic spikes, for instance in September 2009 when the ratio of sample means increases by 50%. Again, the GREG index is smoother than the ratio of sample means (but still very volatile) and strikingly similar to the SPAR. The same picture emerges for the other strata, so we do not present those results.

[Insert Figure 5]

Finally, using the stratum results, we computed stratified GREG indexes for the whole country according to equation (19), where the base period appraisal shares serve as stock value weights. As can be seen from Figure 6, there are hardly any differences between the stratified and unstratified GREG indexes, suggesting that sample selection bias is not a major issue. Figure 6 also shows a second alternative GREG price index, computed according to equation (21), which is based on OLS regressions of the dummy variable model (20). And again, the differences with the original GREG index appear to be small.

[Insert Figure 6]

It should be noted that even within strata some houses are still more likely to sell than others, in particular during the crisis after 2008, so that some sample selection bias in the GREG and SPAR indexes will remain. The direction and magnitude of this bias can only be predicted if data on property characteristics was available to estimate the likelihood of houses to sell. Also, as was mentioned earlier, a too detailed stratification will increase both the sampling variance and small sample bias in case the number of houses sold is extremely low and may raise rather than reduce the mean square error of the estimators.

5. Discussion

5.1 Comparing GREG to SPAR

The most interesting question arising from Section 4 is: why are the GREG and SPAR index numbers so similar in spite of their very different construction methods? It is not remarkable that the trends are similar: although the GREG index does not rely on the matched-model methodology, this index does aim at the same target as the SPAR index. If the sample sizes n^0 and n^t would approach the population size N^0 – which in reality will of course never happen – then both price indexes approach the value change of the fixed housing stock. Put differently, the two methods are both asymptotically unbiased or ‘consistent’.

What may come as a surprise is that the GREG index exhibits roughly the same amount of volatility over time as the SPAR index. To understand the reason why, recall that, with OLS, the regression residuals sum to zero in every time period. This implies $\sum_{n \in S^0} p_n^0 / n^0 = \sum_{n \in S^0} \hat{p}_n^0 / n^0$ and $\sum_{n \in S^t} p_n^t / n^t = \sum_{n \in S^t} \hat{p}_n^t / n^t$. For the basic regression models (7) and (11), the SPAR index can thus alternatively be written as

$$\hat{P}_{SPAR}^{0t} = \frac{\sum_{n \in S^t} \hat{p}_n^t / n^t}{\sum_{n \in S^0} \hat{p}_n^0 / n^0} \left[\frac{\sum_{n \in S^0} a_n^0 / n^0}{\sum_{n \in S^t} a_n^0 / n^t} \right] = \frac{(\hat{\alpha}^t + \hat{\beta}^t \bar{a}^{0(t)}) / \bar{a}^{0(t)}}{(\hat{\alpha}^0 + \hat{\beta}^0 \bar{a}^{0(0)}) / \bar{a}^{0(0)}} = \frac{\hat{\alpha}^t / \bar{a}^{0(t)} + \hat{\beta}^t}{\hat{\alpha}^0 / \bar{a}^{0(0)} + \hat{\beta}^0}, \quad (22)$$

using (8) and (12) for $n \in S^0$ and $n \in S^t$, respectively, where $\bar{a}^{0(0)} = \sum_{n \in S^0} a_n^0 / n^0$ and $\bar{a}^{0(t)} = \sum_{n \in S^t} a_n^0 / n^t$ for short. There is a striking similarity between the last expression on the right-hand sides of (22) and (16). The only difference is that the SPAR index (22) divides the coefficients $\hat{\alpha}^0$ and $\hat{\alpha}^t$ by the sample means of appraisals, $\bar{a}^{0(0)}$ and $\bar{a}^{0(t)}$, whereas the GREG index (16) divides them both by the fixed, non-stochastic population mean \bar{a}^0 . Essentially, the SPAR index is a fully sample-based estimator of the GREG index.

Compared with the SPAR method, the GREG approach eliminates one source of sampling error, i.e., the sampling variability of the mean appraisals. In accordance with generalized regression theory, we would intuitively expect the GREG method to reduce the sampling error of the price index and produce a less volatile time series (under the reasonable assumption that $\bar{a}^{0(t)}$ and $\hat{\alpha}^t$ are uncorrelated across periods $t = 0, \dots, T$). Put differently, while the GREG method has been designed as an improvement over the

ratio of sample means, we might have expected it to work as a smoothing procedure for the SPAR index also. But, as was shown in Section 4, in practice this is hardly the case. This result can be explained as follows.

The variance reduction of the GREG index relative to the SPAR depends on the value of the intercept terms from the regressions in periods 0 and t . If the regression lines passed exactly through the origin ($\hat{\alpha}^t = \hat{\alpha}^0 = 0$), then the GREG index and SPAR index would both be equal to the ratio of the slope coefficients $\hat{\beta}^t / \hat{\beta}^0$ and no reduction in variance would be achieved. In the less extreme case, when $\hat{\alpha}^t$ and $\hat{\alpha}^0$ are close to 0 and the ratios $\hat{\alpha}^t / \bar{a}^0$, $\hat{\alpha}^t / \bar{a}^{0(t)}$, $\hat{\alpha}^0 / \bar{a}^0$ and $\hat{\alpha}^0 / \bar{a}^{0(0)}$ in (16) and (23) are very small compared to $\hat{\beta}^t$ and $\hat{\beta}^0$, the GREG and SPAR indexes will differ only slightly and the variance reduction will be marginal; see also the Appendix.

The latter is indeed what happens in practice, as can be seen from Figures 7 and 8 where the values of $\hat{\alpha}^t / \bar{a}^0$ and $\hat{\alpha}^t / \bar{a}^{0(t)}$ and those of $\hat{\beta}^t$ are plotted over time. The ratios $\hat{\alpha}^t / \bar{a}^0$ and $\hat{\alpha}^t / \bar{a}^{0(t)}$ are remarkably similar and small as compared to the $\hat{\beta}^t$'s. Although we cannot ignore those ratios, it is the change in $\hat{\beta}^t$ that mainly drives the GREG and SPAR indexes. The SPAR index is not only a fully sample-based estimator of the GREG index, as mentioned above, it appears to be almost as efficient.

[Insert Figure 7]

[Insert Figure 8]

5.2 The Volatility of the Slope Coefficient

Several factors may have contributed to the volatility of the slope coefficients $\hat{\beta}^t$ in our regressions of selling prices on appraisals and hence of the GREG and SPAR indexes. We will briefly discuss three of these factors: sample mix change, heteroskedasticity and outliers.

A sample of houses can be viewed as a sample of locations, or addresses, since houses are attached to the land they are built on. A change in the sample mix is nothing else than a change in the observed mix of locations at the lowest level. A *location mix change* affects the sample composition in terms of the average quality characteristics of the properties, such as the number of rooms, surface area, etc. In our simple framework,

where we observe only one (non-physical) characteristic, namely the appraised value, a location mix change boils down to a change in the sample distribution of the appraisals. This, together with any varying price changes across market segments, induces a change in the sample distribution of the ratios p_n^t / a_n^0 , which in turn leads to a change in $\hat{\beta}^t$ in the two-variable regression model (11).

Other than by stratification there is little we can do about the effect of changes in the sample mix of locations (but stratifying by province and type of dwelling did not help much), so the volatility of $\hat{\beta}^t$ and therefore of the GREG and SPAR indexes, will be difficult to reduce. Controlling for location at the address level is also impossible in hedonic imputation methods. Here, the effect of (location) mix change is mitigated by controlling for region plus a range of physical characteristics. However, this does not necessarily mean that hedonic imputation will produce more stable index series than the GREG or SPAR methods. Most standard hedonic models fit the cross sectional data less well than our model does, and the characteristics' coefficients typically exhibit a great deal of variability over time. So maybe it is not surprising that Bourassa, Hoesli and Sun (2006) find that “the SPAR index [...] reliably tracks house price changes, but exhibits less volatility than index methods that require more parameter estimates.”

We can alternatively look at the variability of the slope coefficient from a purely statistical perspective. It is well known that in a two-variable model the OLS estimator $\hat{\beta}^t$ can be written as

$$\hat{\beta}^t = r(p^t, a^0) \frac{s(p^t)}{s(a^0)}, \quad (23)$$

where $r(p^t, a^0)$ denotes the sample correlation coefficient in period t between selling prices and appraisals, which is equal to the square root of R^2 ; $s(p^t)$ and $s(a^0)$ are the corresponding sample standard deviations. A comparison of Figures 1 and 8 suggests that sudden changes in R^2 are largely responsible for the volatility of $\hat{\beta}^t$. In December 2004 for example, a substantial drop in R^2 coincides with a significant decrease of $\hat{\beta}^t$ (and with a decrease in the GREG and SPAR indexes, as shown by Figure 4).

Least squares regression can either be weighted or unweighted. In the absence of *heteroskedasticity*, i.e., when the variance of the errors is constant, OLS should be used. Weighted Least Squares (WLS) is preferred if there is evidence of heteroskedasticity; using appropriate weights, WLS will lead to more stable coefficients than OLS. In this case the unweighted sample sum of the residuals differs from zero so the estimator (15)

has to be applied. To facilitate the interpretation of the GREG index and the comparison with the SPAR index, in Section 3 we assumed away the problem of heteroskedasticity and restricted ourselves to OLS. Note that the (OLS) GREG estimator (16) remains asymptotically design unbiased if heteroskedasticity is present.

The most interesting form of (classical) heteroskedasticity – and, given our data set, the only form we would have been able to reduce – would arise if the variance of the errors of our regression model (11) depended on the appraisal value, being the only regressor. However, the residuals from our OLS regressions do not point to substantial heteroskedasticity of this type. This is illustrated in Figure 9 for three months, including the base period (January 2003), where the sale prices are plotted against the appraisals; the regression lines are also given. To be sure, we also performed the White (1980) test. This test did not point towards the presence of this form of heteroskedasticity either.

[Insert Figure 9]

Our initial data set of sale prices and appraisals included some obvious *outliers*. To estimate the GREG index we therefore made use of a cleaned data set that has been prepared to compute the official Dutch house price index. Statistics Netherlands applies several data cleaning procedures. Houses that were sold more than once in a month are excluded from the data set. To delete entry errors and outliers that may unduly affect the results, properties with sale prices or appraisals below €10,000 or above €5,000,000 and properties with ‘unrealistic’ sale price-appraisal ratios are also removed. The removal of ‘unrealistic’ observations is done by looking at the distribution of the logarithm of the sale price-appraisal ratios; all observations are deleted for which the log ratio differs more than 5 standard deviations from the mean. For more information, see Statistics Netherlands (2008).

These procedures are rather arbitrary. For regression-based estimators such as the GREG it is more appropriate to delete observations with high leverage, i.e. to delete those sample units that have a big impact on the regression coefficients when they are excluded from the sample. A well-known measure in this context is the DFBETA of a sample unit (Cook and Weisberg, 1982). Since the SPAR can be written as a regression-based index, this measure could be used here as well to detect and delete outliers. The scatter plots in Figure 9 show that the cleaned data set still contains some big outliers.

Whether these have high leverage, and whether removing them will reduce the volatility of the $\hat{\beta}'$'s and the GREG and SPAR indexes, remains to be seen.

5.3 Some Further Points

The GREG method is based on the premise of a fixed housing stock. That is, we have assumed that there are no entries (e.g., newly-built houses) or exits (discarded houses) and that housing quality remains fixed over time. Our approach is non-symmetric in that we condition on the *base period* stock. From an index number point of view we estimate a Laspeyres price index for the housing stock where the quantities are all equal to 1 because every house is treated as a unique property. An equally justifiable approach would be to measure the price change of the current period stock, which includes additions to the stock in each period, using a Paasche index. Taking the geometric mean of both indexes would lead to the Fisher index. The Fisher index is a preferred measure of price change due to its symmetric form. The construction of a Fisher-type GREG index is, however, infeasible since the Paasche component requires real time assessed values for houses that are new to the stock, which are obviously not available.

The assumption of a fixed (base period) housing stock can be relaxed through annual chaining, provided that the housing stock is re-assessed annually. This is the current state of affairs in the Netherlands; in the past, assessments were undertaken once every three or four years. Annual updating of the appraisals might also adjust for quality changes of the properties, to some extent at least, because the updated appraisals likely account for major repairs, remodelling and depreciation.

One final remark is in order. For some purposes it is desirable to decompose the overall house price index into two components: a component that measures the change in the price of the structure and a component that measures the change in the price of the land. Neither our GREG method nor SPAR and repeat sales methods are fit for that purpose. Hedonic imputation methods might work, notwithstanding practical problems like multicollinearity; see Diewert, de Haan and Hendriks (2012) for a first attempt. If data on structure size, plot size and other price-determining attributes became available for all properties in the housing stock, then we would be able to estimate a “hedonic imputation GREG index”, including the land-structure split. The chances of getting such data in the Netherlands are unfortunately negligible.

6. Conclusion

The simple GREG method outlined in this paper, which is based on OLS regressions of selling prices on appraisals, substantially reduces the volatility of a house price index as compared to the ratio of sample means. The SPAR index can be viewed as an estimator of the OLS GREG index (which itself is an estimator, of course) where the base period population mean of appraisals is replaced by the sample means in the base period and the comparison period. Our empirical results for the Netherlands indicate that the SPAR index is almost as efficient as the GREG index, even for small sub-populations. We have checked this by drawing a random sample of 50 observations each month from the total number of monthly sales (15,000 on average). The month-to-month changes of the SPAR index were only slightly bigger than those of the GREG.

Due to compositional change of the properties sold, the GREG (and SPAR) time series exhibit strong short-term volatility. An increase in a particular month is typically followed by a decrease in the next month. Put differently, the month-to-month changes do not tell us much about the true price change of the housing stock which, except under unusual circumstances, should behave smoothly. An improved outlier detection method might help reduce the index volatility, but the effect will probably be limited. Applying a smoothing procedure would seem to be an option. However, that will typically lead to revisions of previously published price index numbers, and the lack of revisions is one of the strengths of the GREG and SPAR approaches. Another option would be to reduce the frequency of observation, for example to quarters, but that may be undesirable as well.

From a purely statistical point of view, in our two-variable model the variability of R^2 seems to be responsible for a large part of the volatility of the slope coefficient and therefore of the volatility of the price index series. Future research could focus on the relation between compositional changes in terms of the property characteristics and changes in R^2 . As many housing characteristics are unavailable, we cannot investigate this issue with our data. Fortunately, Statistics Netherlands has access to a data set from the largest Dutch association of real estate agents that might be useful for this purpose. This data set covers around 70% of all housing sales in the Netherlands during 1999-2008, includes many property characteristics and has been enriched with appraisal data. In the past we already used the data set to compare the SPAR index with various types of hedonic indexes.

Acknowledgements

The authors would like to thank the participants at the Economic Measurement Group Workshop, 1-3 December 2010, University of New South Wales, Sydney, Australia and the participants at an Applied Economics Seminar, 22 November 2011, University of Queensland, Brisbane, Australia for their helpful comments on preliminary versions of the paper. Comments and suggestions made by the editor and two anonymous referees also helped to improve the paper. The assistance of Erna van der Wal, who provided us with the data, is gratefully acknowledged. The views expressed in this paper are those of the authors and do not necessarily reflect the views of Statistics Netherlands.

Appendix: Approximate Standard Errors of the GREG Index

The GREG index defined by equation (16) in the main text is a ratio of two estimators, \hat{p}_{GREG}^t and \hat{p}_{GREG}^0 ; for brevity we delete ‘‘OLS’’. Using a first-order Taylor expansion, the variance of the index can be approximated by (see e.g. Kendall and Stuart, 1976)

$$\text{var}(\hat{P}_{GREG}^{0t}) \cong \left[\frac{E(\hat{p}_{GREG}^t)}{E(\hat{p}_{GREG}^0)} \right]^2 \left[\frac{\text{var}(\hat{p}_{GREG}^t)}{\{E(\hat{p}_{GREG}^t)\}^2} + \frac{\text{var}(\hat{p}_{GREG}^0)}{\{E(\hat{p}_{GREG}^0)\}^2} + \frac{\text{cov}(\hat{p}_{GREG}^t, \hat{p}_{GREG}^0)}{E(\hat{p}_{GREG}^t)E(\hat{p}_{GREG}^0)} \right], \quad (\text{A.1})$$

where $E(\hat{p}_{GREG}^t)$ and $E(\hat{p}_{GREG}^0)$ denote expected values.

The covariance term in (A.1) is equal to 0 since, by assumption, the samples in periods 0 and t are independently drawn. Replacing the expected values in (A.1) by the estimators and subsequently taking the square root leads to the following expression for the standard error of \hat{P}_{GREG}^{0t} :

$$se(\hat{P}_{GREG}^{0t}) \cong \hat{P}_{GREG}^{0t} \left[\frac{\text{var}(\hat{p}_{GREG}^t)}{(\hat{p}_{GREG}^t)^2} + \frac{\text{var}(\hat{p}_{GREG}^0)}{(\hat{p}_{GREG}^0)^2} \right]^{1/2}. \quad (\text{A.2})$$

Equation (A.2) can be estimated in practice using $\hat{p}_{GREG}^s = \hat{\alpha}^s + \hat{\beta}^s \bar{a}^0$ ($s = 0, t$), hence $\text{var}(\hat{p}_{GREG}^s) = \text{var}(\hat{\alpha}^s) + (\bar{a}^0)^2 \text{var}(\hat{\beta}^s) + 2\bar{a}^0 \text{cov}(\hat{\alpha}^s, \hat{\beta}^s)$. Estimates of the (co)variances are readily available in most statistical packages from the variance-covariance matrix.

Dividing (A.2) by \hat{P}_{GREG}^{0t} yields an expression for the relative standard error or coefficient of variation, $CV(\hat{P}_{GREG}^{0t}) = se(\hat{P}_{GREG}^{0t}) / \hat{P}_{GREG}^{0t}$, of the GREG index:

$$CV(\hat{P}_{GREG}^{0t}) \cong \left[\frac{\text{var}(\hat{p}_{GREG}^t)}{(\hat{p}_{GREG}^t)^2} + \frac{\text{var}(\hat{p}_{GREG}^0)}{(\hat{p}_{GREG}^0)^2} \right]^{1/2} = \left[\{CV(\hat{p}_{GREG}^t)\}^2 + \{CV(\hat{p}_{GREG}^0)\}^2 \right]^{1/2}. \quad (\text{A.3})$$

Of more importance is the relative standard error of the *percentage change* of the index, i.e. $CV(\hat{P}_{GREG}^{0t} - 1) = se(\hat{P}_{GREG}^{0t} - 1) / (\hat{P}_{GREG}^{0t} - 1)$. This is generally greater than $CV(\hat{P}_{GREG}^{0t})$, given that $se(\hat{P}_{GREG}^{0t} - 1) = se(\hat{P}_{GREG}^{0t})$ and $\hat{P}_{GREG}^{0t} - 1 < \hat{P}_{GREG}^{0t}$.

If both regression lines almost pass through the origin, hence $\hat{\alpha}^s \cong 0$ ($s = 0, t$), we have $\hat{P}_{GREG}^{0t} \cong \hat{\beta}^t / \hat{\beta}^0$ and (A.2) simplifies to

$$se(\hat{P}_{GREG}^{0t}) = se(\hat{P}_{GREG}^{0t} - 1) \cong \hat{P}_{GREG}^{0t} \left[\frac{\text{var}(\hat{\beta}^t)}{(\hat{\beta}^t)^2} + \frac{\text{var}(\hat{\beta}^0)}{(\hat{\beta}^0)^2} \right]^{1/2}. \quad (\text{A.4})$$

In this particular case the GREG and SPAR indexes nearly coincide, so (A.4) also holds for the SPAR index (using \hat{P}_{SPAR}^{0t} rather than \hat{P}_{GREG}^{0t}).

References

- Academetrics (2009), House Price Indices – Fact or Fiction, www.academetrics.co.uk.
- Bailey, M.J., Muth, R.F. and Nourse, H.O. (1963). A Regression Method for Real Estate Price Construction. *Journal of the American Statistical Association*, 58, 933-942.
- Beaumont, J.F. and Alavi, A. (2004). Robust Generalized Regression Estimation. *Survey Methodology*, 30, 195-208.
- Bourassa, S.C., Hoesli, M. and Sun, J. (2006). A Simple Alternative House Price Index Method. *Journal of Housing Economics*, 15, 80-97.
- Calhoun, C.A. (1996). OFHEO House Price Indexes: HPI Technical Description. Office of Federal Housing Enterprise Oversight, Washington, DC.
- Case, K.E. and Shiller, R.J. (1987). Prices of Single-Family Homes Since 1970: New Indexes for Four Cities. *New England Economic Review*, September-October, 45–56.
- Case, K.E. and Shiller, R.J. (1989). The Efficiency of the Market for Single Family Homes. *The American Economic Review*, 79, 125-137.
- Cochran, W.G. (1977). *Sampling Techniques*, 3rd edition, New York: Wiley.
- Cook, R.D. and Weisberg, S. (1982). *Residuals and Influence in Regression*, New York: Chapman and Hall.
- Diewert, W.E. (2009). The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Future Directions, pp. 87-116 in W.E. Diewert, B.M. Balk, D.

- Fixler, K.J. Fox and A.O. Nakamura (eds.), *Price and Productivity Measurement: Volume 1 – Housing*, Trafford Press.
- Diewert, W.E., Heravi, S. and Silver, M. (2009). Hedonic Imputation versus Time Dummy Hedonic Indexes, pp. 161-196 in W.E. Diewert, J. Greenlees and C. Hulten (eds.), *Price Index Concepts and Measurement*, NBER Studies in Income and Wealth, vol. 70. Chicago: Chicago University Press.
- Diewert, W.E., de Haan, J. and Hendriks, R. (2012). The Decomposition of a House Price Index into Land and Structures Components: A Hedonic Regression Approach. *Econometric Reviews* (forthcoming).
- Edelstein, R.H. and Quan, D.C. (2006). How Does Appraisal Smoothing Bias Real Estate Returns Measurement? *Journal of Real Estate Finance and Economics*, 32, 41-60.
- Eurostat (2010). *Technical Manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices*, Version 1.9. Available at www.epp.eurostat.ec.europa.eu/portal/page/portal/hicp/documents/Tab/Tab/03_METH-OOH-TECHMANUAL_V1-9.pdf.
- Francke, M.K. (2010). Repeat Sales Index for Thin Markets: A Structural Time Series Approach. *Journal of Real Estate Finance and Economics*, 41, 24-52.
- Geltner, D. (1996). The Repeated-Measures Regression-Based Index: A Better Way to Construct Appraisal-Based Indexes of Commercial Property Value. *Real Estate Finance*, 12, 29-35.
- Gouriéroux, C. and Laferrère, A. (2009). Managing Hedonic House Price Indexes: The French Experience. *Journal of Housing Economics*, 18, 206-213.
- Grimes, A. and Young, C. (2010). A Simple Repeat Sales House Price Index: Comparative Properties Under Alternative Data Generation Processes. Motu Working Paper 10-10, Motu Economic and Public Policy Research, New Zealand.
- de Haan, J. (2007). Formulae for the Variance of (Changes in) the SPAR Index. Unpublished manuscript, Statistics Netherlands, Voorburg (Dutch only; available from the author upon request).
- de Haan, J. (2010). Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and 'Re-Pricing' Methods, *Journal of Economics and Statistics* (Jahrbucher für Nationalökonomie und Statistik), 230, 772-791.

- de Haan, J., van der Wal, E. and de Vries, P. (2009). The Measurement of House Prices: A Review of the Sale Price Appraisal Method. *Journal of Economic and Social Measurement*, 34, 51-86.
- Hedlin, D., Falvey, H., Chambers, R. and Kokic, P. (2001). Does the Model Matter for GREG Estimation? A Business Survey Example. *Journal of Official Statistics*, 17, 527-544.
- Hill, R.J. and Melsner, D. (2008). Hedonic Imputation and the Price Index Problem: An Application to Housing. *Economic Inquiry*, 46, 593-609.
- Jansen, S.J.T., de Vries, P., Coolen, H.C.C.H., Lamain, C.J.M. and Boelhouwer, P. (2008). Developing a House Price Index for the Netherlands: A Practical Application of Weighted Repeat Sales. *Journal of Real Estate Finance and Economics*, 37, 163-186.
- Kendall, M. and Stuart, A. (1976). *The Advanced Theory of Statistics – Volume 1: Distribution Theory*, 4th edition, London: Charles Griffin & Company.
- Leventis, A. (2006). Removing Appraisal Bias from a Repeat Transactions House Price Index: A Basic Approach. Paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, 6-7 November 2006.
- Makaronidis, A. and Hayes, K. (2006). Owner Occupied Housing for the HICP. Paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, 6-7 November 2006.
- Hardman, M. (2011). Calculating High Frequency Australian Residential Property Price Indices. Rismark Technical Paper, available at www.rpdata.com/images/stories/content/PDFs/technical_method_paper.pdf.
- Rossini, P. and Kershaw, P. (2006). Developing a Weekly Residential Price Index Using the Sales Price Appraisal Ratio. Paper presented at the twelfth Annual Pacific Rim Real Estate Society Conference, Auckland, 22-25 January 2006.
- Saarnio M. (2006). Housing Price Statistics at Statistics Finland. Paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, 6-7 November 2006.
- Särndal, C.E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*, New York: Springer-Verlag.
- Shi, S., Young, M. and Hargreaves, B. (2009). Issues in Measuring a Monthly House Price Index in New Zealand. *Journal of Housing Economics*, 18, 336-350.

- Statistics Netherlands (2008). Price Index Owner-occupied Existing Dwellings; Method Description. Statistics Netherlands, The Hague, available at www.cbs.nl/NR/rdonlyres/A49D8542-26EC-40FD-9093-82A519247F4B/0/MethodebeschrijvingPrijsindexBestaandeKoopwoningene.pdf.
- de Vries, P., de Haan, J., van der Wal, E. and Mariën, G. (2009). A House Price Index Based on the SPAR Method. *Journal of Housing Economics*, 18, 214-223.
- van der Wal, E., ter Steege, D. and Kroese B. (2006). Two Ways to Construct a House Price Index for the Netherlands: The Repeat Sale and Sale Price Appraisal Ratio. Paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, 6-7 November 2006.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48, 817–838.

Tables and Figures

Table 1. Regression results

Month	Alpha	t	Beta	t	R squared
January 2003	1900.49	2.26	0.98	275.19	0.87
January 2004	5039.16	5.96	1.01	269.26	0.88
January 2005	-2555.12	2.43	1.08	237.54	0.84
January 2006	1282.14	1.41	1.11	286.39	0.87
January 2007	-7567.99	6.36	1.19	243.72	0.83
January 2008	11007.39	8.48	1.26	231.93	0.83
January 2009	16677.31	9.83	1.30	184.24	0.81

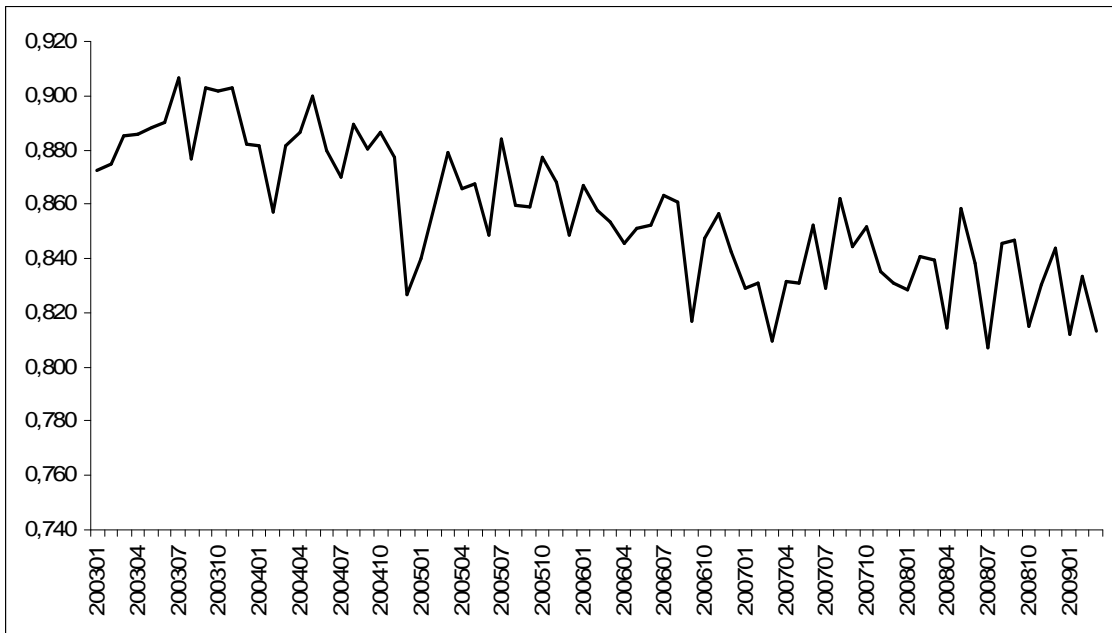


Figure 1. R squared values

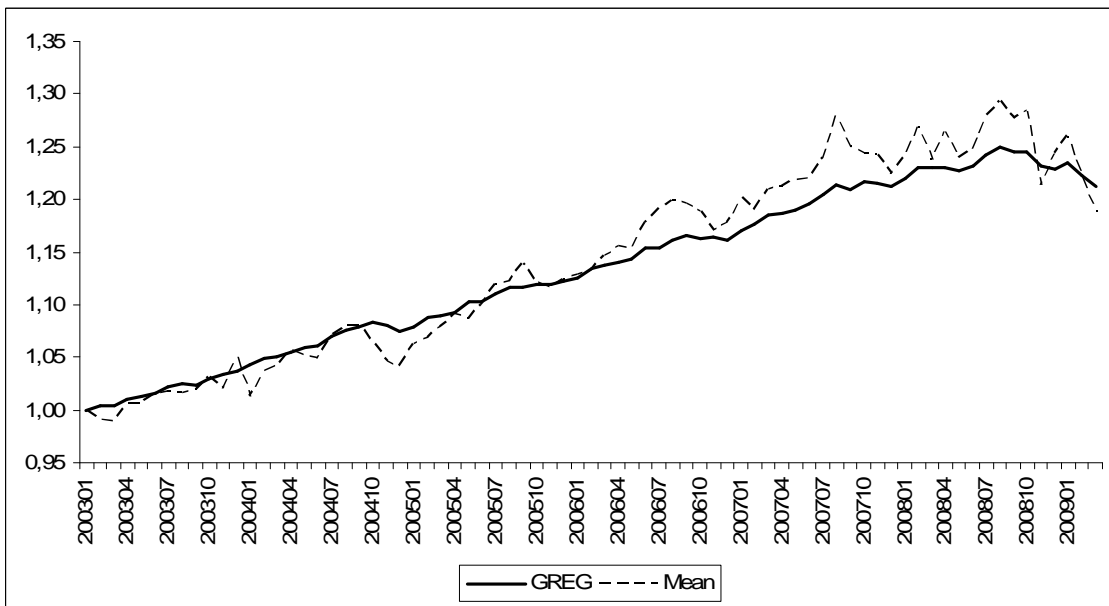


Figure 2. GREG index and ratio of sample means

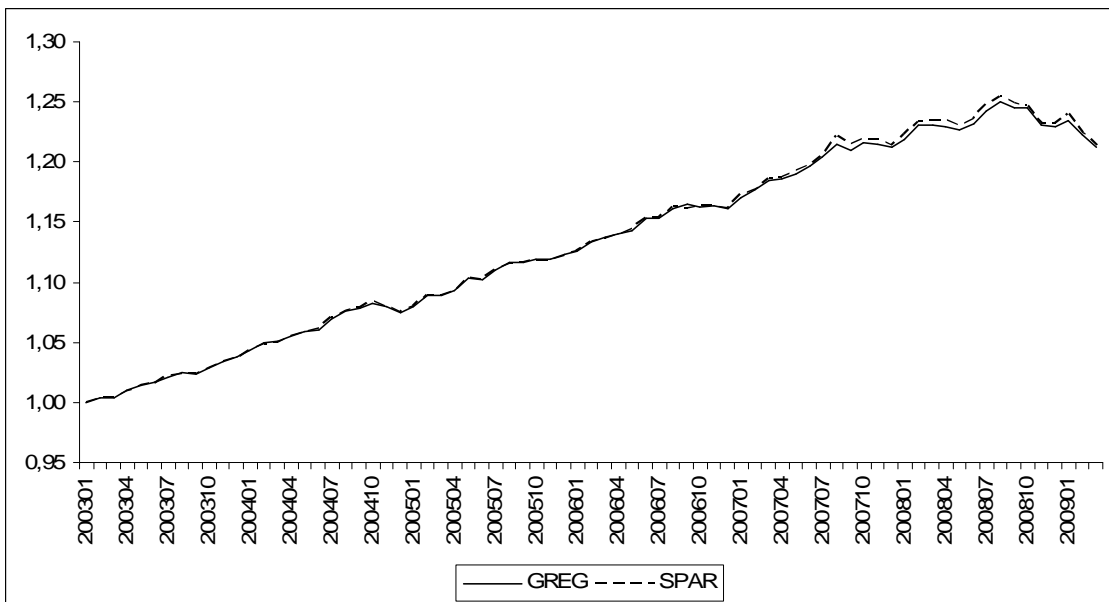


Figure 3. GREG and SPAR indexes

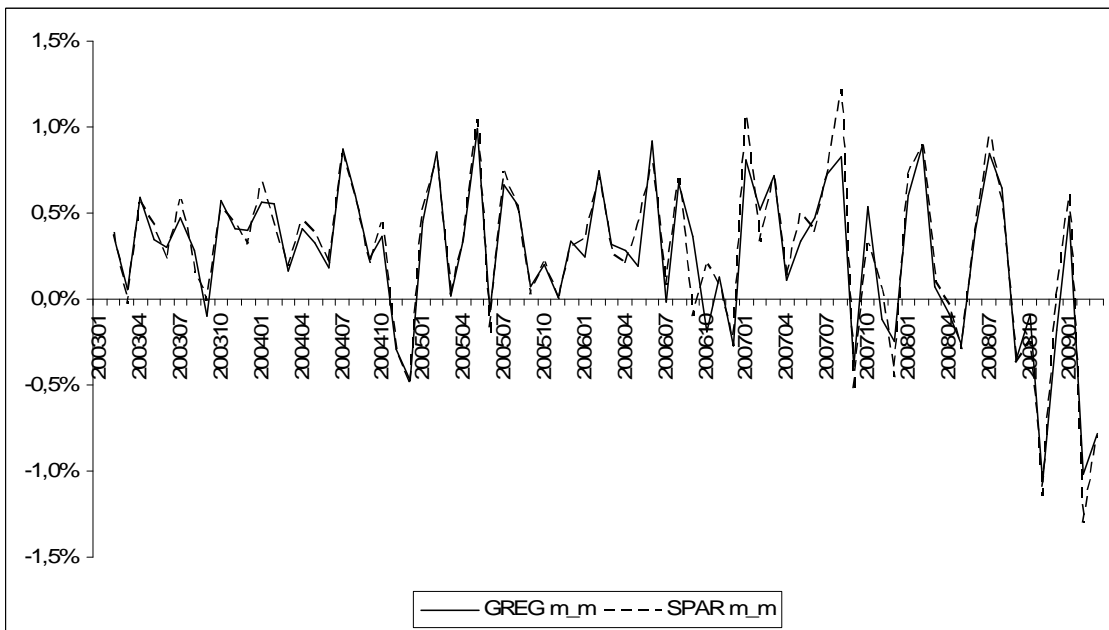


Figure 4. GREG and SPAR: month-to-month percentage changes

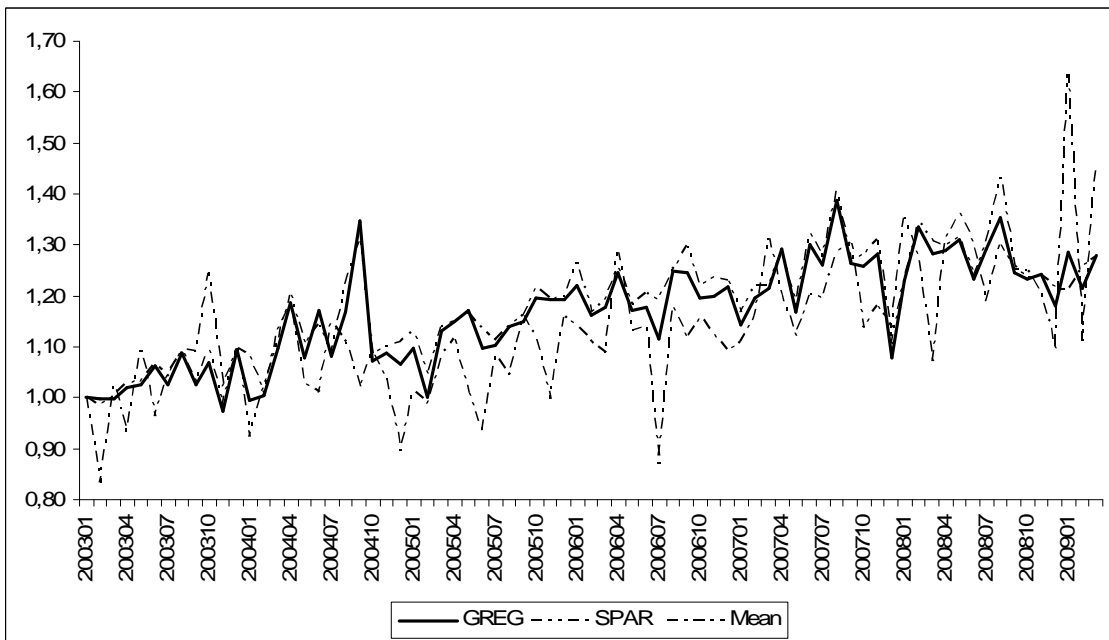


Figure 5. GREG and SPAR indexes and ratio of sample means; apartments in the province of Friesland

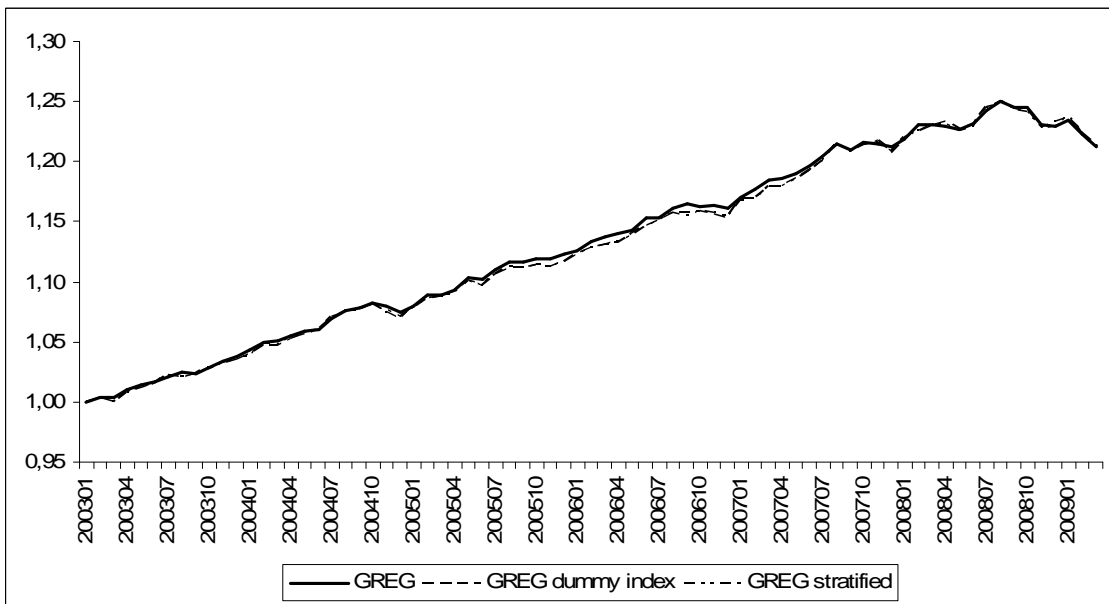


Figure 6. GREG, stratified GREG and dummy variable GREG indexes

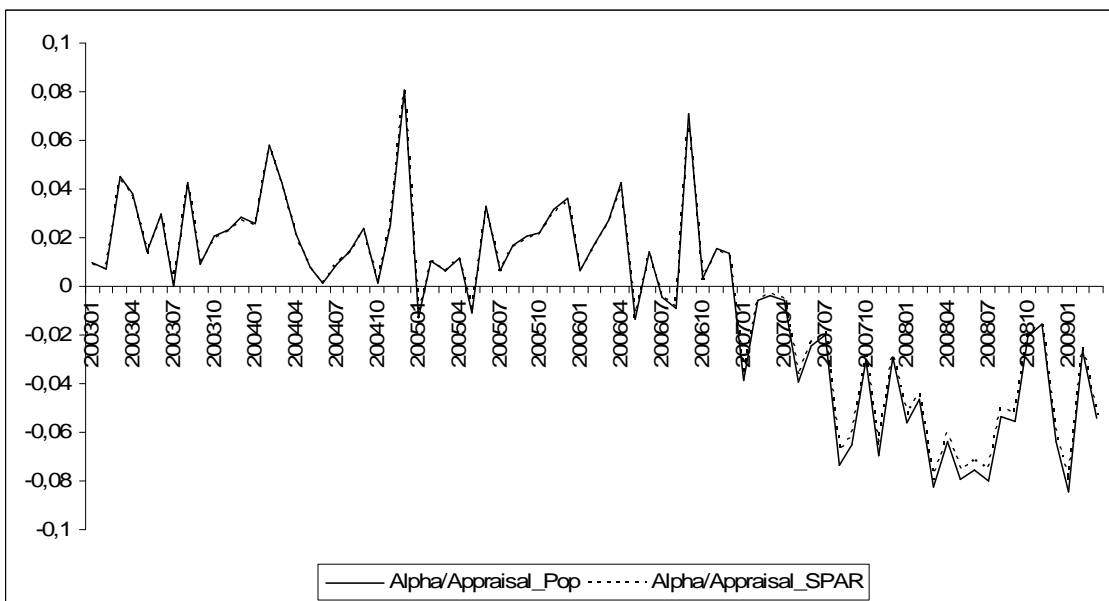


Figure 7. Intercepts divided by appraisal means

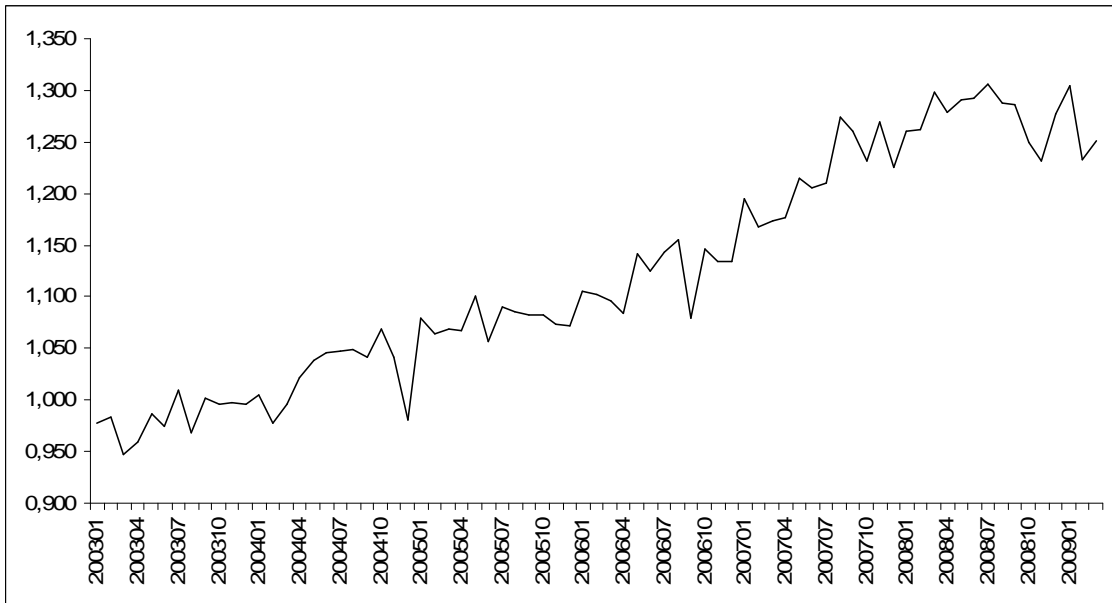
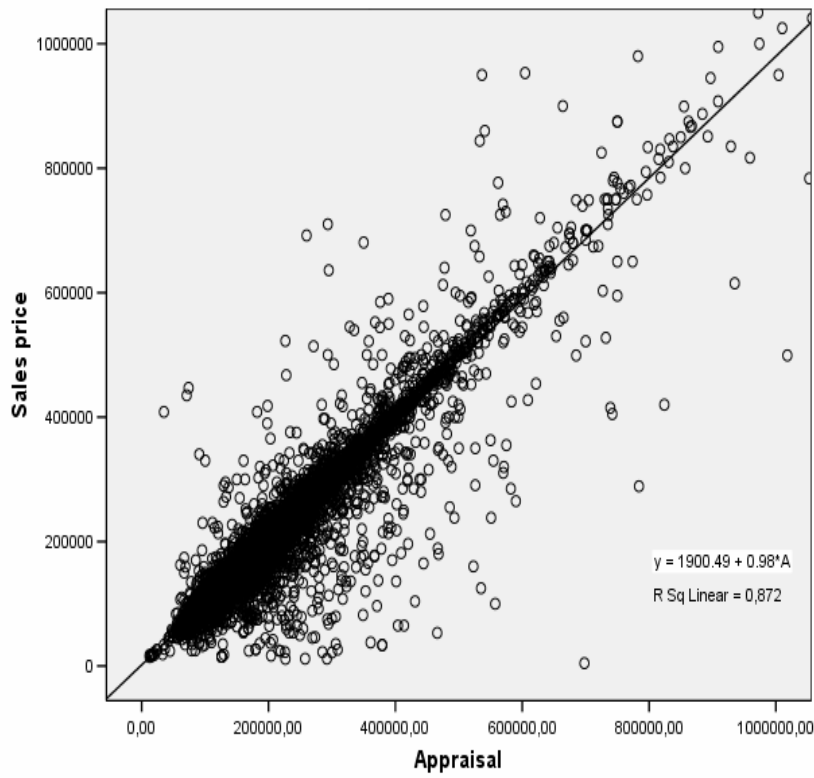
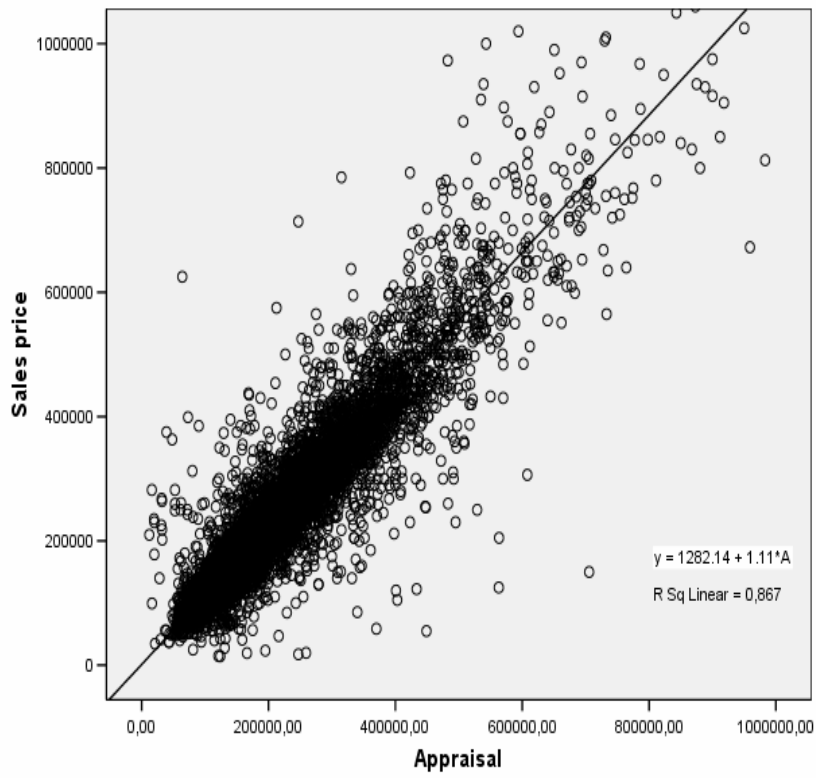


Figure 8. Slope coefficients

January 2003



January 2006



January 2009

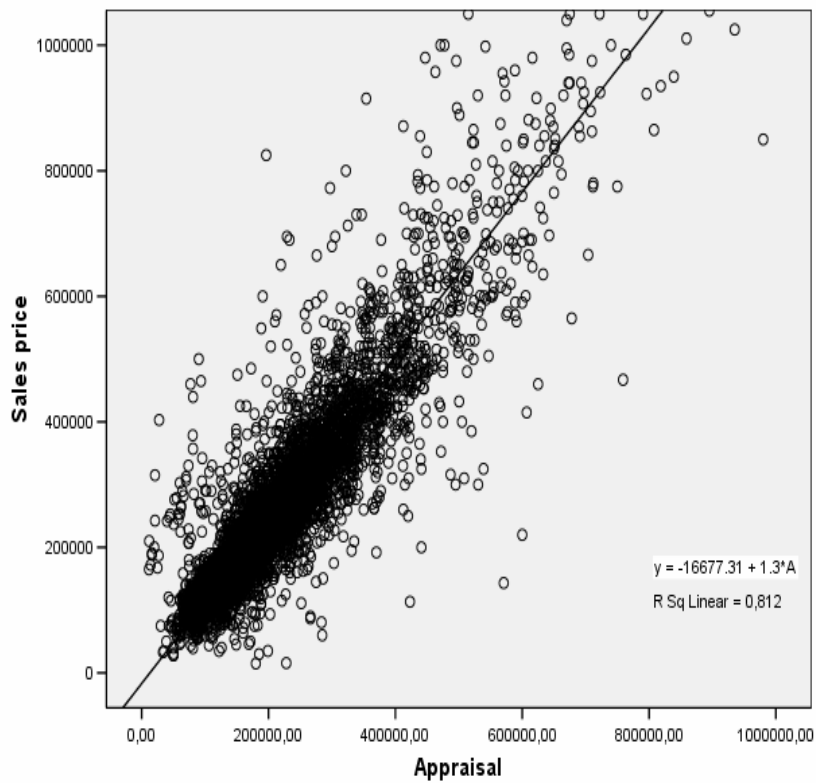


Figure 9. Scatter plots and regression lines