

**WHY AND HOW TO CONSTRUCT A GENUINE  
BELGIAN PRICE INDEX OF HOUSE SALES**

by

**André DECOSTER  
Kris DE SWERDT**

**Public Economics**

**Center for Economic Studies  
Discussion Paper Series DPS 05.15  
<http://www.econ.kuleuven.ac.be/ew/admin/default.htm>**

**July 2005**

D/2005/2020/17

# Why and how to construct a genuine Belgian price index of house sales\*

André Decoster<sup>†</sup> and Kris De Swerdt<sup>‡</sup>  
CES, Katholieke Universiteit Leuven

July 2005

## Abstract

Assessing the price evolution of houses on the basis of average sales prices, as is current practice in Belgium, might be misleading due to changing characteristics of the houses sold in the periods observed. A hedonic index which takes into account changes in characteristics is more appropriate. We use the budget surveys of the Belgian Statistical Institute to illustrate how this also applies for Belgium. The estimated hedonic price index for house sales on the secondary market is practically always below the index based on average sales values for the period considered. This demonstrates the need to collect more extensive data on the characteristics of the dwellings sold in Belgium.

**JEL Number:** R31, C51, E31

**Keywords:** hedonic house price index, hedonic regression, house price inflation.

## 1 Introduction

In recent years the topic of housing has undoubtedly pushed itself to the forefront of the public debate in many Western countries. During the last four months, housing and housing markets figured prominently in at least 6 issues of *The Economist*. The soaring house prices, or at least the perception thereof, might not be alien to this increased interest. Belgium is no exception in this respect. With clock-like regularity newspapers and economic magazines report about the price increases. Policy makers also devote increasing attention to the perceived problem of unaffordability of house acquisition.

There is an abundance of reasons to motivate this increased interest in the "price" of houses. Indeed, whereas equity holdings such as stocks and bonds are mainly concentrated in

---

\*The research forms part of a project funded by the Flemish community, in the framework of *point of support for scientific research* 'Governmental organisation in Flanders'. All calculations on the basis of the NIS budget surveys were done to the account of the Belgian National Statistics Institute. We are grateful to Bart Capéau for comments on an earlier version of this paper. All remaining errors and interpretations are ours. Corresponding author: kris.deswerdt@econ.kuleuven.be.

<sup>†</sup>Center for Economic Studies, Katholieke Universiteit Leuven, Naamsestraat, 69, B-3000 Leuven, Belgium.

<sup>‡</sup>Policy Research Centre "Governmental Organization in Flanders" and Center for Economic Studies, University of Leuven, Naamsestraat 69, B-3000 Leuven, Belgium.

the upper regions of the wealth distribution, housing is the major asset in most households' portfolios, certainly in countries with high rates of homeownership. As such, house price changes affect not only the possibility for younger generations to choose this desired portfolio composition, but can also have potentially large effects on the wealth of owners. It is therefore not surprising that many have wondered in recent years whether or not (at least part of) the recent rapid increase in house prices is a speculative bubble comparable to the stockmarket bubble that burst in 2000. Although opinions about the existence of a speculative bubble in the housing market diverge quite widely, it is beyond all doubt that a minimum requirement to (start to) adequately investigate the item is a reliable house price index. Furthermore, McCarthy and Peach (2004) show that the expected magnitude of macroeconomic effects following a flattening or decline in house prices, crucially hinges on the price concept used for evaluating house price evolution. Also, behavioural models of the housing market that try to explain the purchase of a house, or the length of residence in a given dwelling (such as, among many others, Capéau et al., 2003, Goodman, 1995, Haurin and Chung, 1998, Kan 2000 and Kendig, 1984), cannot do without a proper house price index. Finally, the construction of a welfare measure related to housing needs an estimate of real housing consumption. Starting from nominal house values, the first requirement to make comparisons over time is evidently to divide the nominal values by an appropriate price index.

A price index for houses is, however, less straightforwardly constructed than in many other cases. Indeed, dwellings are outstanding examples of non-homogeneous commodities. They differ considerably in structural characteristics such as the presence of amenities, surface and location. Although this is not a problem as such – price indices precisely aim at aggregating prices of non-homogeneous commodities – it becomes a problem when the composition of the sample on which the price index is based changes over time as far as these structural characteristics are concerned. Therefore, assessing the price evolution on the average or median sales values of houses sold in a given period, is not the scientifically sound way to measure price changes over time. McCarthy and Peach (2004) show that the difference is relevant. Over the period 1977-2003 they report a percentage increase of the "median price of new homes sold" in the US to be 311, whereas the "constant quality new home price" increases by only 199 percent (p. 3). One third of the price increase is removed if quality is kept constant. They conclude:

*"[...] a perceived increase in home prices may reflect only the demand for more housing services that can be obtained through better quality homes with more amenities, even as the price of a standardized unit of housing services may stay the same". (p.4).*

Also in Belgium, statements made in the public domain (by real estate agencies, press, but also government officials) about the house price evolution seem to be based on the evolution of average sales prices through time. The most often used and cited index in Belgium is the Stadim index, which is essentially an average (or quantile value, like the third quartile) of the sales values in a given year without correction for the changing structural characteristics of the dwellings sold that year.<sup>1</sup> Yet, the methodology to implement more accurately the

---

<sup>1</sup>Stadim is a Belgian-based research agency for real estate. Every year they publish a report on the evolution of house prices in Belgium for some broad categories of dwellings (with base year 1953). More

theoretically sound concept itself of a price change, by keeping quality constant, is available in the form of hedonic price indices. The estimation of a hedonic regression for the sales value of houses sold in different periods, allows the effect of changing characteristics on the sales value to be isolated. This enables one to estimate the evolution of prices over time under the *ceteris paribus* clause that characteristics (and hence quality) are kept unchanged, exactly as is required to construct a genuine price index.

In other countries such as the US, the use of hedonic indices is more widespread (e.g. McCarthy and Peach, 2004; Wallace, 1996; Meese and Wallace, 1991; Mills and Simenauer, 1996). Also in the UK an often cited index, the Halifax index, is based on hedonic regression analysis.<sup>2</sup> And recently, academics and notaries have put efforts together in France to construct hedonic price indices for several French communities and districts (Laferrère, 2003). The notaries record all sales in a given period together with the characteristics of the dwellings sold. The methodology of estimation of the hedonic regression and the regression itself is done by the French statistical agency INSEE.

In this paper we apply the methodology of hedonic regression to illustrate the possible divergence between the price evolution based on average sales values and one based on a hedonic index for Belgium. We use the sales values of houses and their characteristics from the last four available budget surveys of the Belgian Statistical Institute. Unfortunately, the limited nature of the data do not allow us to unambiguously conclude exactly how much of the reported price increases by e.g. Stadim, are due to quality changes. Contrary to many other countries, however, there does not exist a publicly available database in Belgium that records both sales values and detailed characteristics of houses sold. Our empirical work is therefore at best illustrative of the possibilities of the methodology, and a plea in favour of the recognition of the need for more accurate and reliable data for the construction of a genuine house price index in Belgium.

The remainder of the paper is organised as follows. In section 2 we describe the technique of hedonic regression. Section 3 describes the data we have at our disposal, and in section 4 we show and discuss the results. Section 5 concludes.

## 2 Hedonic regression

The fact that price  $p_t$  in time period  $t$  is the factor by which we multiply the quantity  $q_t$  to obtain the value  $v_t$  of this quantity at time  $t$  ( $v_t = p_t \cdot q_t$ ), makes clear that price  $p_t = \frac{v_t}{q_t}$  is to be interpreted as the value of one unit of the commodity under consideration. Therefore, a price index based on sales values  $v_t$ , only traces the evolution of the price  $p_t$  through time if  $q_t$  is kept constant. Certainly for markets in which the traded commodities are very heterogeneous as far as quality is concerned, the evolution of uncorrected sales values is more of a non-starter if the aim is to analyse price evolutions. The housing market is a

---

information can be found on [www.stadim.be](http://www.stadim.be) and in Janssens and De Wael (2005).

<sup>2</sup>More information on house price indices in the UK and the methodology used to construct them can be found on <http://www.nationwide.co.uk/hpi/methodology.htm> and <http://www.hbosplc.com/economy/indexmethodology.asp>.

prominent example. A newly built house with multiple bedrooms, two garages, and a garden obviously contains more units of housing than a small outdated flat which lacks modern and comfortable amenities. It is therefore not surprising that, in most cases the value of this large comfortable house will be higher than the small outdated flat:  $v_t$  is higher, not because  $p_t$  is higher, but because the house contains more units of housing  $q_t$ .

To be freed from the interference of the multiplication of price and quantity, one possibility of course, would be to study the evolution of sales values for homogeneous categories of dwellings. Unfortunately, from a practical point of view, this greatly reduces the scope of possible empirical work simply because each dwelling is a unique commodity, with its own unique characteristics such as location, age, amenities, surface etc.<sup>3</sup>

A more promising route is to try to model the sales value of a commodity sold at time  $t$  as a function of its characteristics at time  $t$ , picking up the change in quantity, and the time period at which it is sold (Malpezzi, 2003). Formally, we disentangle the change in prices and quantities, by writing:

$$\ln v(t) = \ln p(t) + \ln q(t), \quad \text{for } t = 0 \dots T \quad (1)$$

where we will empirically specify the quantity term in function of characteristics of the dwelling, and the price term in function of the period when a dwelling was sold.

Denoting the vector of structural characteristics of a dwelling  $i$  from a sample of dwellings sold in period  $t$  of  $T$  periods as  $\mathbf{S}_i(t)$ , the vector of neighbourhood characteristics as  $\mathbf{N}_i(t)$ , and the vector of geographical characteristics as  $\mathbf{G}_i(t)$ , we choose a linear specification for the second term in (1):

$$\ln q_i(t) = \alpha + \beta' \mathbf{S}_i(t) + \gamma' \mathbf{N}_i(t) + \lambda' \mathbf{G}_i(t) + \varepsilon_i, \quad (2)$$

to estimate the coefficients  $\beta, \gamma$  and  $\lambda$ . The disturbance term is denoted by  $\varepsilon_i$ . For the price term we use the identity:

$$\ln p(t) \equiv \sum_{\tau=0}^{T-1} \delta_{t\tau} \ln p(\tau) \quad (3)$$

in which  $\delta_{t\tau}$  denotes the Kronecker delta, to rewrite the empirical analogue of (1) as:

$$\ln v_i(t) = \alpha + \boldsymbol{\pi}' \mathbf{T}_i + \beta' \mathbf{S}_i(t) + \gamma' \mathbf{N}_i(t) + \lambda' \mathbf{G}_i(t) + \varepsilon_i. \quad (4)$$

The vector  $\mathbf{T}_i$  consists of the Kronecker delta's or dummies for dwelling  $i$  which are all zero, except in the period  $t = \tau$  in which dwelling  $i$  has been sold, where it takes the value of one. For the reference period 0, we normalize  $p(0) = 1$ . This allows us to obtain an estimate of  $\ln p(\tau)$  as  $\hat{\pi}_\tau$ , such that the estimate for the price index for period  $\tau$ , say  $\hat{P}(\tau)$  is obtained

---

<sup>3</sup>The "repeated sales" method tries to exploit the fact that if the same dwelling is sold several times, one can assume that the heterogeneity, though not completely absent of course (e.g. through depreciation and/or renovation) is less of a problem than when observing sales prices of different dwellings. Obviously this makes heavy demands as far as data collection is concerned. Since the sample has to consist of dwellings for which at least two sales have been observed, one needs a large amount of data, and large time periods to infer relevant price evolutions.

as<sup>4</sup>:

$$\widehat{P}(\tau) = \frac{\widehat{p}(\tau)}{p(0)} = e^{\widehat{\pi}\tau}. \quad (5)$$

### 3 The data

We use data from four Belgian household budget surveys (1997-1998, 1999, 2000 and 2001) of the Belgian Statistical Institute to estimate the price index in (5) for Belgium. There is no doubt that better data might be conceived for this purpose. The major limitation lies in the fact that we do not dispose of a description of the characteristics of the dwelling at the moment at which it has been sold, i.e. at the moment for which the sales value is observed. Yet, to the best of our knowledge no better information is publicly available in Belgium at the moment, and at least a considerable amount of potentially useful information on owner-occupied housing is available in these budget surveys.<sup>5</sup>

In table 1 we give an overview of the kind of information readily found in the respective surveys together with some summary statistics. The first two columns refer to the full sample of 8264 owner-occupiers. For the hedonic regression we discard households that had their house newly built and confine the analysis to households that have bought on the secondary market. The information for this subsample of 4833 households can be found in the last two columns of table 1. We first describe the characteristics for the full sample.

The majority of households live in a house (88%) rather than an apartment (12%), with detached and semi-detached dwellings making up the larger part. The average number of bedrooms is 2.9, and the average floor surface amounts to 118 m<sup>2</sup>. The latter figure is an approximation obtained by adding the surfaces of the different rooms the dwelling consists of. Note that lot size, a variable which is widely used in hedonic price regressions, is not registered in the budget survey. Of all owner-occupiers in the sample, 90% have a garden<sup>6</sup>, and 70% a garage. In the full sample we have more than half of the households living in Flanders, but in the subsample used in the regressions this percentage is reduced to 43%. The exact year of construction is not recorded in the budget survey, but approximated by intervals. Slightly less than one third of the dwellings dates from before 1946. We used these intervals to identify whether the house was acquired on the primary or the secondary market. When the year of purchase is inside the interval of construction we classified the house as newly built. Obviously, in the subsample of the secondary market, older dwellings make up a larger fraction of the sample (those constructed before 1946 now make up more than half of the sample). We also used these intervals of years of construction to approximate the age of the dwelling, by subtracting the mean year of each of these intervals from the year of purchase. The average age of a dwelling at the moment of purchase, was 27.2 years in the

---

<sup>4</sup>To be unbiased, an extra term should be included in expression (5). The correct expression then is:  $\widehat{P}(\tau) = \frac{\widehat{p}(\tau)}{p(0)} = e^{\widehat{\pi}\tau} \cdot e^{-\frac{1}{2}\text{var}(\widehat{\pi}\tau)}$ , where  $\text{var}()$  stands for variance. The results reported in the text take this correction into account, but there was hardly any difference when using expression (5) directly.

<sup>5</sup>The other major Belgian survey, The Panel Study of Belgian Households, does contain many interesting characteristics of the dwelling, but unfortunately lacks information on the sales value at which the house has been bought.

<sup>6</sup>A garden should be broadly interpreted. It also includes for example a terrace.

full sample, and 47 in the subsample used in the regressions. As far as the amenities are concerned, nearly all houses have running water. The heating technology is mainly central heating (more than 75%) using either oil or gas.

With amenities and heating possibilities, we touch upon an important limitation of the information in the budget survey to estimate the price index. All characteristics of the dwelling, except of course the price at which it was acquired, are recorded at the moment of the survey. If some amenities were not present at the moment of purchase (e.g. one has bought a house without central heating but had it installed later), or, more generally, that the characteristics have changed since the moment of purchase, they are not the appropriate variables to take up in a regression which explains the sales value at the moment of purchase. Therefore we have made a selection of characteristics for which the assumption of no change since the moment of purchase, was not too unreasonable. We have indicated them in bold in Table 1. For example, we assume that the number of rooms and their floor surface, recorded at the moment of the survey, is still the same as it was at the time of purchase. This assumption definitely introduces error in the estimation, makes a comparison with the Stadim index aimed too high, and seriously hampers a policy-relevant interpretation of the index we will obtain. However, until detailed quality and price information about the housing market is collected and becomes available, this is the only way to construct at least some sound price index for Belgian dwellings, and it does not prevent us from investigating – on the same set of data – the amount of bias introduced by using average sales values.

Table 1. Characteristics and summary statistics

variable	Full sample (8264 obs)		Sec. market (4833 obs)	
	mean	st.dev.	mean	st.dev.
<b>purchase price (€)</b>	67772.42	66554.88	59912.27	51117.33
<b>year of purchase</b>	1983.56	11.92	1986.63	10.40
<b>detached dwelling (%)</b>	39.6	48.9	25.3	43.5
<b>semi-detached dwelling (%)</b>	22.3	41.6	24.0	42.7
<b>enclosed dwelling (%)</b>	25.9	43.8	36.7	48.2
<b>apartmentbuilding 2 units (%)</b>	0.9	9.7	1.2	10.7
<b>apartmentbuilding 3-4 units (%)</b>	1.7	12.9	2.5	15.5
<b>apartmentbuilding 5-9 units (%)</b>	2.6	15.8	2.9	16.8
<b>apartmentbuilding &gt;=10 units (%)</b>	6.6	24.7	6.8	25.2
<b>other type of dwelling (%)</b>	0.5	7.2	0.7	8.2
<b>number of kitchens (total)</b>	0.98	0.20	0.99	0.21
zero (%)	2.8	n/a	3.0	n/a
one (%)	96.0	n/a	95.5	n/a
more than one (%)	1.2	n/a	1.5	n/a
<b>number of living rooms (total)</b>	0.92	0.33	0.88	0.37
zero (%)	9.6	n/a	12.9	n/a
one (%)	89.3	n/a	85.9	n/a
more than one (%)	1.1	n/a	1.2	n/a
<b>number of dining rooms (total)</b>	0.22	0.43	0.28	0.46
zero (%)	78.2	n/a	72.2	n/a
one (%)	21.5	n/a	27.4	n/a
more than one (%)	0.3	n/a	0.4	n/a
<b>number of sitting rooms (total)</b>	0.19	0.41	0.24	0.44
zero (%)	81.5	n/a	76.4	n/a
one (%)	18.0	n/a	23.1	n/a
more than one (%)	0.5	n/a	0.5	n/a
<b>number of bedrooms (total)</b>	2.90	1.04	2.82	1.10
zero (%)	0.7	n/a	0.7	n/a
one (%)	6.8	n/a	8.4	n/a
two (%)	24.9	n/a	29.5	n/a
three (%)	44.4	n/a	38.8	n/a
more than three (%)	23.2	n/a	22.6	n/a
<b>number of bathrooms (total)</b>	1.10	0.38	1.09	0.39
zero (%)	1.7	n/a	2.5	n/a
one (%)	87.6	n/a	87.2	n/a
more than one (%)	10.7	n/a	10.3	n/a
<b>number of toilets (total)</b>	1.13	0.66	1.03	0.69
zero (%)	13.4	n/a	19.8	n/a
one (%)	62.6	n/a	60.3	n/a
more than one (%)	24.0	n/a	19.9	n/a

Table 1 continued on next page



Table 1. Characteristics and summary statistics (continued)

variable	Full sample (8264 obs)		Sec. market (4833 obs)	
	mean	sd	mean	sd
<b>floor area (square meters)<sup>1</sup></b>	118.04	57.86	116.48	57.46
<b>age dwelling at purchase (years)<sup>2</sup></b>	27.23	28.80	46.55	22.77
<b>garden (%)</b>	89.6	30.5	86.9	33.8
<b>garage (%)</b>	70.0	45.8	55.8	49.7
<b>brussels region (%)</b>	12.3	32.9	16.7	37.3
<b>walloon region (%)</b>	36.0	48.0	40.5	49.1
<b>flanders region (%)</b>	51.7	50.0	42.7	49.5
year of construction before 1946 (%)	31.4	46.4	53.6	49.9
year of construction 1946-1960 (%)	12.1	32.6	16.5	37.1
year of construction 1961-1970 (%)	14.9	35.6	13.1	33.7
year of construction 1971-1980 (%)	15.4	36.1	9.4	29.2
year of construction 1981-1990 (%)	10.9	31.2	3.2	17.6
year of construction after 1991 (%)	11.5	31.9	0.4	6.4
year of construction unknown (%)	3.8	17.4	3.9	16.4
running water (%)	98.8	10.9	98.9	10.3
running hot water (%)	96.6	18.2	95.2	21.4
vegetable garden (%)	22.0	41.4	19.7	39.7
work shop (%)	38.3	48.6	40.7	49.1
elevator in building (%)	8.3	27.7	8.9	28.5
balcony (%)	17.0	37.6	18.4	38.7
central heating oil (%)	39.2	48.8	34.2	47.5
central heating city gas (%)	32.3	46.8	33.3	47.2
central heating gastank (%)	1.0	10.0	1.0	9.9
centr. heat. oil shared with other units (%)	3.4	18.1	4.4	20.4
centr. heat. gas shared with other units (%)	2.2	14.6	2.4	15.4
centr. heat. gastank shared with other units (%)	0.1	2.7	0.1	3.2
electrical heating accumulation (%)	5.5	22.7	3.1	17.4
electrical heating radiators (%)	1.2	10.9	0.7	8.6
heating coal (%)	1.4	11.9	2.1	14.2
heating wood (%)	1.0	10.0	1.3	11.3
heating gas convectors (%)	7.2	25.9	10.7	30.9
heating gas radiators (%)	0.5	7.0	0.6	8.0
heating oil (%)	2.6	15.8	3.6	18.7
heating others (%)	2.5	12.7	2.3	13.2
garbage collection: container with chip (%)	3.1	17.4	2.8	16.5
garbage collection: yearly tax (%)	68.3	46.5	68.5	46.5
garbage collection: bags (mandatory) (%)	24.0	42.7	24.7	43.1
garbage collection: other system (%)	4.6	19.5	4.0	18.0

<sup>1</sup> floor area is calculated as the sum of the different room surfaces<sup>2</sup> age of the building is an approximation based on the mean year of each construction interval

## 4 Results

### 4.1 Average sales prices in the budget survey

Figure 1 displays the crude information of sales prices of houses in the budget survey bought on the secondary market (the black line), and compares it with the Consumer Price Index and the often cited Belgian price index for houses by Stadim.<sup>7</sup>

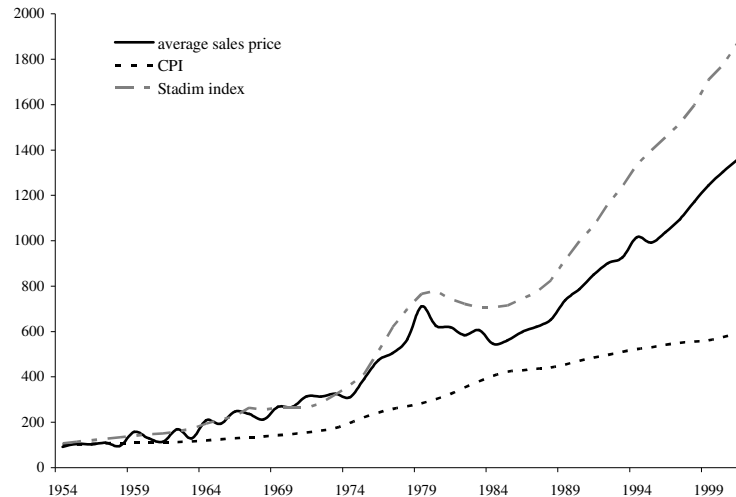


Figure 1: Average sales prices in the Budget surveys versus Stadim index and CPI

House values on the secondary market have risen much faster than the general level of prices as measured by the CPI. The average sales values recorded in the budget survey do rise less sharply than the Stadim index. One of the reasons for this might be that the data in the budget survey suffer from ‘recollection bias’, i.e. people that purchased their house a long time ago might not remember the exact (or even the approximate) price at the time of the survey. But in general the pattern is the same: a sudden and sharp increase from 1974 to 1979, followed by a pronounced slump in the first half of the eighties. From then on, the house prices clearly depart from the CPI, for which the rate of increase slows down. In figure 2, we focus on these rates of change. We confine the picture to the data since 1975, because of the low numbers of sales recorded in the budget survey for the earlier years (see Table A1 in the Appendix).

Figure 2 confirms that the rate of change of average purchase values of houses, as recorded in the budget survey, mimics quite well the pattern displayed by the Stadim index. There was a sharp fall in the rate of increase at the end of the seventies, and even a drop in the prices during the first half of the eighties. In 1985, house values started to rise again, and since 1986 have risen continuously faster than the general level of prices. The question is

<sup>7</sup>The CPI was retrieved from the website of the federal Ministry of Economic Affairs, <http://www.mineco.fgov.be>.

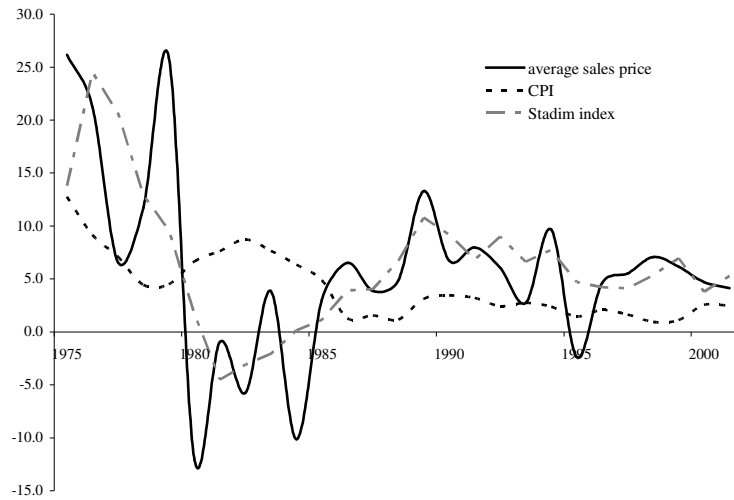


Figure 2: Yearly percentage change in the average sales prices in the Budget surveys versus Stadim index and CPI

which part of this change reflects a change in the characteristics of the sample of houses sold on the secondary market each year and which part constitutes a pure price change. To deduce this, we turn to the hedonic regression.

#### 4.2 A hedonic regression for the secondary market in Belgium

In table 2 the regression results for specification (4) are shown. We also included a series of district dummies, but for obvious space limitations, we do not show the results here.<sup>8</sup> The same applies to the year dummies in (4).<sup>9</sup>

<sup>8</sup>In figure A1 of the appendix we display the differences in the hedonic price index for the several districts.

<sup>9</sup>A complete table with the regression results is available from the authors upon request.

Table 2. Hedonic regression results

variable	Coef.	Std. Err.	p-value
number of kitchens	-0.0233	0.0436	0.5920
number of living rooms	0.0351	0.0259	0.1750
number of dining rooms	0.0134	0.0196	0.4950
number of sitting rooms	0.0187	0.0236	0.4280
number of bedrooms	0.0631	0.0090	0.0000
number of bathrooms	0.1479	0.0245	0.0000
number of toilets	0.0603	0.0121	0.0000
number of other rooms	0.0326	0.0100	0.0010
floor area	0.0022	0.0003	0.0000
floor area squared	-9.04E-07	1.20E-07	0.0000
semi detached dwelling <sup>1</sup>	-0.0897	0.0228	0.0000
enclosed dwelling <sup>1</sup>	-0.1513	0.0236	0.0000
apartmentbuilding 2 units <sup>1</sup>	-0.0177	0.0770	0.8190
apartmentbuilding 3-4 units <sup>1</sup>	-0.1279	0.0534	0.0170
apartmentbuilding 5-9 units <sup>1</sup>	-0.2323	0.0494	0.0000
apartmentbuilding 10 or more units <sup>1</sup>	-0.2312	0.0410	0.0000
other type of dwelling <sup>1</sup>	-0.4256	0.1082	0.0000
garden	0.1090	0.0277	0.0000
garage	0.1659	0.0173	0.0000
age of dwelling	-0.0134	0.0016	0.0000
age of dwelling squared	7.49E-05	1.77E-05	0.0000
constant	8.8869	0.1381	0.0000
R-squared	0.6391		
F-value	79.0400		

<sup>1</sup> detached dwelling is the reference category

The most common amenities, such as a dining room, a sitting room and a living room, have a positive but statistically insignificant impact on the purchase price of a dwelling. The exception is the number of kitchens, which seems to have a negative influence on the sales value, but again, this effect is not statistically significant. Probably the lack of variability in the data for these variables (see table 1) shows up here as a lack of statistical significance. The number of bedrooms, bathrooms, toilets and other rooms, on the other hand, do have a (statistically) significant effect on the prices paid. Adding an extra bedroom for example leads to an increase in sales value by approximately 6%, whereas an extra bathroom increases the value by about 15%, *ceteris paribus*. A larger floor area results in a higher purchase value, but this increase in value decelerates as the floor area becomes larger and larger, albeit at a slow rate.<sup>10</sup> Similarly, as the building ages the purchase value decreases.<sup>11</sup> If we reduce

<sup>10</sup>The price increases up to floor areas of approximately 1217m<sup>2</sup>. Since all but one observation have floor areas that are less than 1217m<sup>2</sup>, we could state that the effect of surface on the sales value is strictly linear over the relevant range.

<sup>11</sup>Sales values would again increase as the dwelling gets older for ages beyond 89 years.

the interpretation of the differentials in sales value to differences at the demand side of the market, the fully detached dwelling, which is the reference category for the dummies taking up the effect of the type of dwelling, seems to be the type of dwelling preferred by Belgian homeowners. The sales values of all other types of dwelling are lower than that of the reference category although not always statistically significantly so. Compared to detached dwellings fully enclosed dwellings sell at values that are about 15% lower and apartments in buildings of 5-9 and more units sell at values that are on average 23% lower. A garden or a garage on the other hand positively influences the sales value.

### 4.3 A constant quality house price index for Belgium

Figure 3 shows the evolution of the hedonic price index calculated on the basis of expression (5) and the average sales values for the same data in the budget survey. If we remove the variation in the "basket" of houses on which the index for average sales prices is based, we find that the hedonic price index has lagged behind the increase in the sales prices. The difference in inflation rates is explicitly displayed in Figure 4. Year by year the difference between the black bars (inflation in the sales values) and the shaded bars (inflation in the hedonic index) shows the importance of the correction for varying characteristics in the stock of dwellings sold over time. Figure 4 shows that this correction is substantial, and that it goes in both directions. Sometimes the hedonic price inflation is higher than the one of the sales values in the budget surveys, sometimes it is lower. Over the whole period, the hedonic index goes up by 825 percent, whereas the sales values increased by 1018 percent, a finding in line with McCarthy and Peach (2004). In terms of yearly inflation rates, the average house price inflation comes down from 8.2 percent when measured by means of the changes in the average sales value, to only 6.0 percent on the basis of the hedonic index. Moreover, if we prevent the cancelling out of the positive and negative values, the average absolute deviation in the yearly inflation of house prices is 12.9%, compared to an average yearly change of the sales value of 17.8% .

We repeat that the hedonic index obtained here cannot serve as a replacement for existing indices for the moment. Our data are too limited for this, which is also illustrated by the rather low  $R^2$ -statistic of 0.64 in table 2, indicating that our hedonic regression is clearly unable to explain the observed heterogeneity of sales values in each period. The negative conclusion out of this might be that the bias in the currently used sales value indices instead of true price indices is only of minor importance. But we prefer a more optimistic conclusion. Since a reliable index requires the observation of characteristics and values to occur simultaneously, the poor fit may not come as a surprise here. Remember that the characteristics are recorded at the time of the survey and considerable renovations and/or improvements might have taken place since the time of purchase. The relation between sales value and characteristics will thus be distorted and therefore, in this note, we limited the set of explanatory variables to those which might be assumed to have remained unchanged between the time of purchase and the time of the survey, inevitably resulting in a bias due to omitted variables. This only reinforces our plea in favour of serious collection of data on

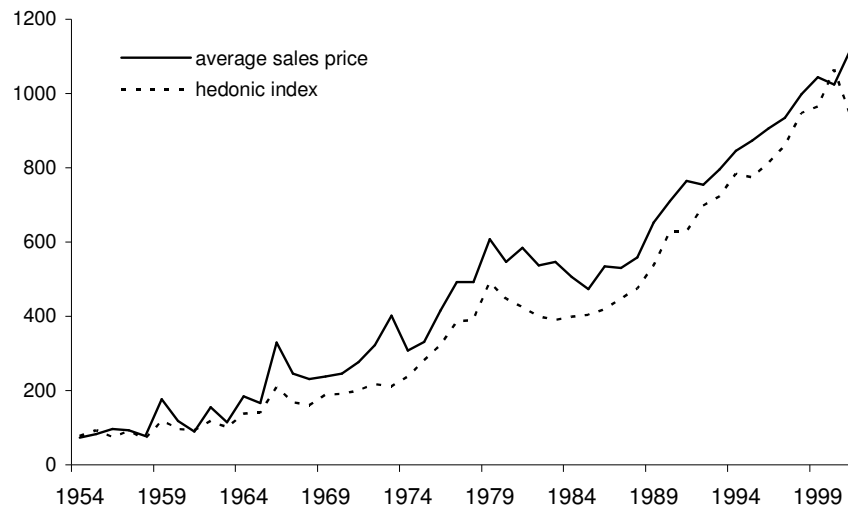


Figure 3: The hedonic index compared to the average sales prices in the budget surveys

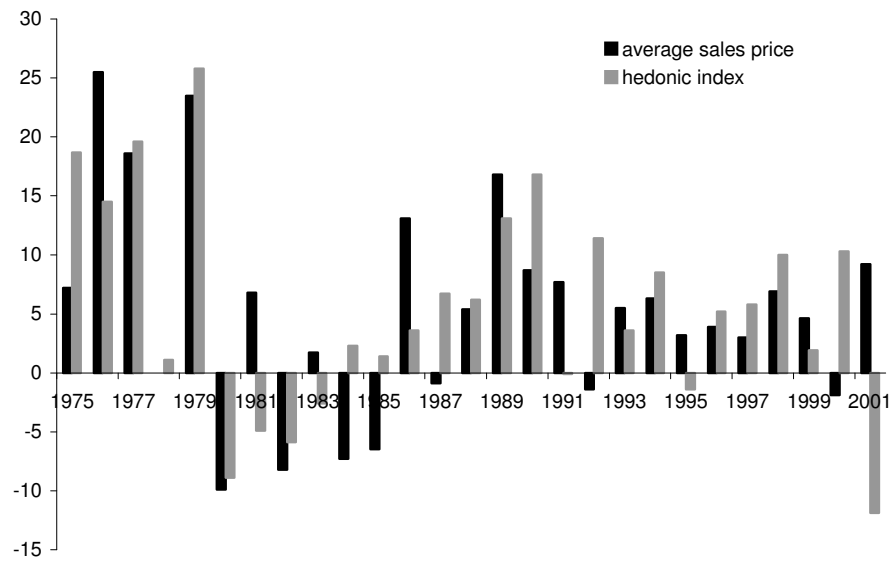


Figure 4: House price inflation in the hedonic index as compared to the average sales prices in the budget surveys

values and characteristics.

## 5 Conclusion

In this paper we have tried to show how indices based on average sales prices of, usually, different dwellings, can present quite a different account of house price evolution than an index that takes into account differences in (levels of) characteristics and their individual contributions to the price. We have argued that the sales value is the product of the price *per unit of housing* and the *quantity* consumed. Differences in sales values thus come from two different sources, price and quantity. We have disentangled these effects by use of a hedonic regression on budget survey data for Belgium. Due to the difference in the moments of observation between sales values and corresponding characteristics, only a subset of the latter was used, i.e. those that might be thought to have remained unchanged between the time of purchase and the time of survey. A semi-logarithmic specification was used to estimate the model. More sophisticated models have been used in the literature, both (semi-)parametric and nonparametric (see Meese and Wallace, 1991; Knight et.al., 1995; Malpezzi, 1999; Anglin and Gencay, 1996; Gilley and Pace, 1995 amongst others). However, since the main intention was to show the discrepancy that might arise between an average sales price index and a hedonic price index, and since the exact contribution (or price) of each characteristic was of no concern here, we did not go too far in the model specification (Börsch-Supan, 1987, more specific studies can be found in e.g. Quang Do and Grudnitsky, 1993; Malpezzi et.al., 1987 and Clapp and Giaccotto, 1998).

The results clearly indicate that caution is indeed warranted when drawing general conclusions from average sales price indices. Although the comparison should be interpreted with caution, the exercise clearly shows a considerable discrepancy between an average sales price index and a hedonic index, even if we compare the hedonic index with an index based on average sales values calculated from the same data. Partly because of the ‘recollection bias’ and the – for this kind of exercise – limited amount of data (both in number of observations and number of characteristics) the results reported here probably underestimate the potential bias in the house price index. They do however indicate the usefulness of collecting data that are suitable for more grounded and elaborate hedonic price regressions and hedonic price indices for the housing sector. Data that come to mind include data on neighbourhood characteristics, proximity to city and business centres, access to public transportation, etc.

The severity of erroneously relying on such indices depends on the situation. If one is interested in the evolution of sales values of properties that are or become available on the market, the reliance might be less problematic. But since important aspects of housing policy inevitably start from a sound and reliable account of house price evolution, this data collection is a necessary first step in the design and development of tools for policy making and evaluation in the housing sector. Hedonic price estimates can then be used as input to social housing policy, as a tool to evaluate rent subsidy programs and even for real estate appraisal in general (Malpezzi, 2003).

## References

- [1] Anglin, P. M. and Gencay, R., 1996, Semiparametric estimation of hedonic price function, *Journal of Applied Econometrics* 11(6), 633-648.
- [2] Börsch-Supan, A., 1987, *Econometric Analysis of Discrete Choice. With Applications on the Demand for Housing in the U.S. and West-Germany*, (Springer-Verlag, Berlin Heidelberg)
- [3] Butler, R. V., 1982, The specification of hedonic indices for urban housing, *Land Economics* 58(1), 96-108.
- [4] Capéau, B., 2003, De prijs(evolutie) van woningen op de secundaire markt, mimeo, Centre for Economic Studies, University of Leuven, Belgium.
- [5] Capéau, B., Decoster A., and Vermeulen F., 2003, Homeownership and the life cycle: an ordered logit approach, Discussion Paper Series 03.08, Leuven, Centre for Economic Studies, University of Leuven, Belgium
- [6] Clapp, J. M. and Giaccotto, C., 1998, Residential hedonic models: a rational expectations approach to age effects, *Journal of Urban Economics* 44, 415-437.
- [7] Gilley, O. W. and Pace, R. K., 1995, Improving hedonic estimation with an inequality restricted estimator, *The Review of Economics and Statistics* 77(4), 609-621.
- [8] Goodman, A. C., 1995, A dynamic equilibrium model of housing demand and mobility with transactions costs, *Journal of Housing Economics* 4, 307-327
- [9] Goodman, A. C., 1998, Andrew Court and the invention of hedonic price analysis, *Journal of Urban Economics* 44, 291-298.
- [10] Haurin, D. R. and Chung, E.-C., 1998, The demand for owner-occupied housing: implications from intertemporal analysis, *Journal of Housing Economics* 7, 49-68
- [11] Janssens P. and De Wael P., 2005, 50 jaar Belgische vastgoedmarkt 1953-2003: waar geschiedenis tot toekomst vergroeit, (Roularta Books, Roeselare, Belgium)
- [12] Kan, K., 2000, Dynamic modelling of housing tenure choice, *Journal of Urban Economics* 48, 46-69
- [13] Kendig, H. L., 1984, Life cycle and residential mobility: implications for the housing market, *Urban Studies* 21, 271-283
- [14] Knight, J. R., Dombrow, J. and Sirmans, C. F., 1995, A varying parameters approach to constructing house price indices, *Real Estate Economics* 23(2), 187-205.
- [15] Laferrère, A., 2003, Hedonic housing price indices: the French experience, Paper Prepared for the IMF and BIS conference on Real Estate Indicators and Financial Stability, October 27-28, 2003, Washington D.C.



- [16] Malpezzi, S., 1999, A simple error correction model of house prices, *Journal of Housing Economics* 8, 27-62.
- [17] Malpezzi, S., Ozanne, L. and Thibodeau, T. G., 1987, Microeconomic estimates of housing depreciation, *Land Economics* 63(4), 372-385.
- [18] Malpezzi, S., 2003, Hedonic pricing models: a selective and applied review, in O'Sullivan T. and Gibb, K., eds., *Housing Economics and Public Policy*, (Blackwell Science Ltd. Oxford), 67-89.
- [19] Mason, C. and Quigley, J. M., 1996, Non-parametric housing prices, *Housing Studies*.11(3), 373-385.
- [20] McCarthy, J. and Peach, R. W., 2004, Are home prices the next "bubble"?, *FRBNY Economic Policy Review*, December 2004
- [21] Meese, R. and Wallace, N. E., 1991, Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices, *AREUEA Journal* 19(3), 308-332.
- [22] Mills, E. and Simenauer, R., 1996, New hedonic estimates of regional constant quality house prices, *Journal of Urban Economics* 39, 209-215.
- [23] Quang Do, A. and Grudnitski, G., 1993, A neural network analysis of the effect of age on housing values, *The Journal of Real Estate Research* 8(2), 253-264.
- [24] Quigley, J. M., 1995, A simple hybrid model for estimating real estate price indices, *Journal of Housing Economics* 4, 1-12.
- [25] Verbeek, M., 2004, *A guide to Modern Econometrics*, Second Edition, (John Wiley & Sons Ltd., Chichester, England)
- [26] Wallace, N. E., 1996, Hedonic-based price indices for housing: theory, estimation and index construction, *FRBSF Economic Review* 3, 34-48.
- [27] Wang, F. T. and Zorn, P. M., 1997, Estimating house price growth with repeated sales data: what's the aim of the game?, *Journal of Housing Economics* 6, 93-118.

## Appendix

Table A1. Number of purchases per year on the secondary market as recorded in the budget surveys<sup>12</sup>

year	number of purchases	year	number of pruchases
1953	12	1978	91
1954	10	1979	92
1955	20	1980	114
1956	21	1981	130
1957	15	1982	111
1958	22	1983	113
1959	18	1984	132
1960	36	1985	186
1961	24	1986	181
1962	41	1987	165
1963	37	1988	174
1964	31	1989	205
1965	41	1990	197
1966	32	1991	204
1967	50	1992	247
1968	43	1993	215
1969	50	1994	267
1970	66	1995	286
1971	52	1996	295
1972	73	1997	303
1973	67	1998	263
1974	53	1999	192
1975	80	2000	57
1976	87	2001	14
1977	82		

---

<sup>12</sup>Remark that it is not impossible that some purchases are double counted. This occurs when the same household participated in more than one budget survey (and the housing situation did not change in the meantime)

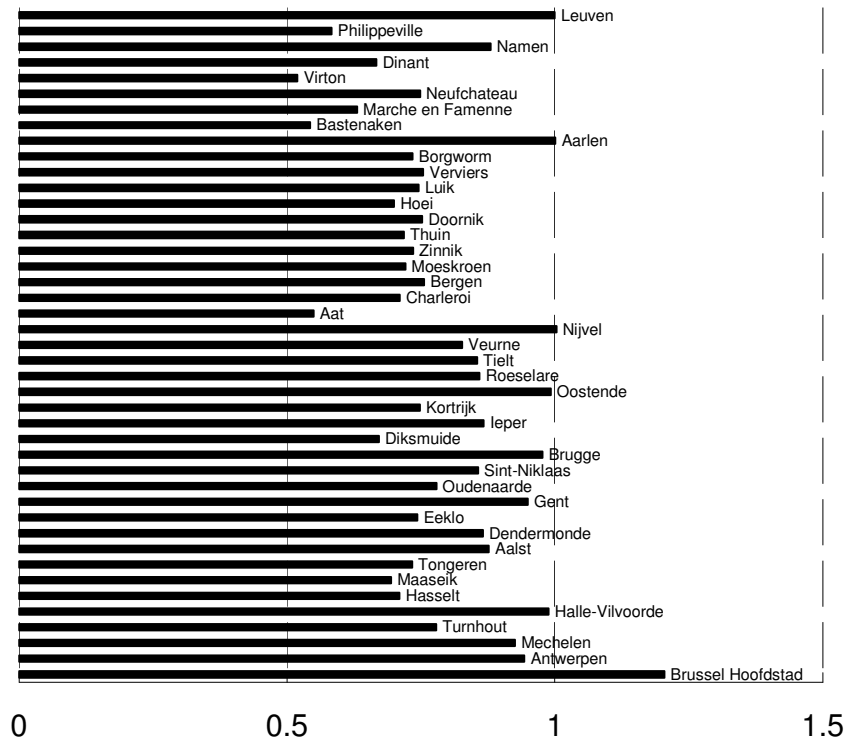


Figure A1: Variation in hedonic prices in different districts in Belgium in the year 2001  
(Leuven is reference category)