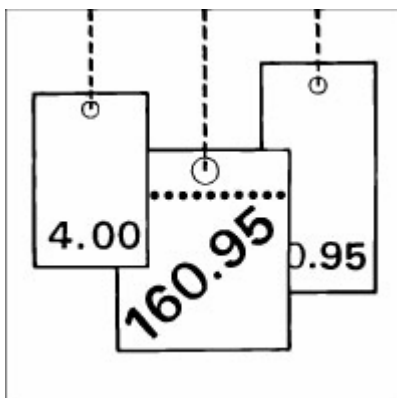


Using Hedonic Pricing for the German House Price Index



15. Juni 2004

Dr. Stefan Linz, Timm Behrmann

Federal Statistical Office of Germany, V A Prices
phone 049 - 06 11 - 75 2659
stefan.linz@destatis.de

© Statistisches Bundesamt, Wiesbaden 2004

Using Hedonic Pricing for the German House Price Index

1. Introduction

The objective of official price statistics is to measure what we call “pure” price changes, purged of the adulterating influence of changes in consumption patterns, types of goods and quality features. This essentially reflects the Laspeyres Principle of once defining a basket of goods and keeping it as constant as possible over a defined period of time.

Especially, the price of an item at two separate times can only be usefully compared if the *quality* of the item remains constant. If this is not the case, quality adjustment is undertaken in order to introduce the monetary value of an item’s quality change into price observation.

Hedonic methods constitute a specific quality adjustment technique. The hedonic method uses regression analysis to measure the influence of product features on the sale price. Thus price changes due to qualitative changes in certain features can be distinguished mathematically and purged from the pure price change which the price index is actually called upon to measure.¹

In 2002 the Federal Statistical Office in Germany began an extensive programme for introducing hedonic techniques of quality adjustment. Figure 1 provides an overview of the stages in this schedule.

As a first step the hedonic method was introduced in June 2002 to the regular monitoring of prices for home computers.²

The second step for the German Federal Statistical Office was to evaluate the quality adjustment procedure hitherto applied to the consumer price index for motor vehicles. A hedonic price index was calculated parallel to the existing price index for new cars. Analysis demonstrated that, for new cars sold in Germany, quality changes due to technological progress are adequately indicated by the “traditional” method of quality adjustment which is so called “Option Pricing” in this case. No systematic deviations between the two indices were observed. The Federal Statistical Office has, therefore, not incorporated the hedonic method into its quality adjustment techniques for new cars, retaining instead the well-proven and significantly cheaper traditional approach.

The third step was to design a hedonic price index for used cars. In May 2003 a used car price index based on the hedonic approach was included in the consumer price index³. One year later,

¹ On hedonic price determination cf. inter alia: Griliches, Zvi, Price Indexes and Quality Change, Studies in New Methods of Measurement, edited by Zvi Griliches for the Price Statistics Committee Federal Reserve Board, Harvard University Press, Cambridge, Massachusetts 1971. Brachinger, H. W. (2002) Statistical Theory of Hedonic Price Indices, Working Paper, available in English at: <http://www.unifr.ch/stat/en-home.php>.

in May 2004, hedonic producer, import and export price indexes for selected data processing equipment have been implemented in the respective indices. Furthermore, work is currently progressing on hedonic price indexes for the categories “electrical home appliances” and “consumer electronics”.

Figure 1: The Federal Statistical Office programme for implementing hedonic methods

<i>Index position</i>	<i>Project status</i>
Consumer prices for home computers	Hedonic index introduced in June 2002
Consumer prices for new cars	Evaluation completed in January 2003. Hedonic methods are not used to calculate this price index as no systematic deviations were observed between the techniques.
Consumer prices for used cars	Hedonic price index implemented in May 2003
Producer, import and export price indexes for electronic data processing equipment	Hedonic price indices implemented in May 2004
Consumer prices for owner occupied housing	Preliminary results in 2004
Consumer prices for electrical home appliances and consumer electronics	Work in progress

The present paper deals with the development of a hedonic price index for “Owner Occupied Housing” in Germany. It is based on the findings and experiences the Federal Statistical Office of Germany has gained within a pilot study initiated by the Statistical Office of the European Communities (Eurostat).

Primary aim of the ‘Harmonized Index of Consumer Prices (HICP)’ is to provide a measure of consumer price inflation on a comparable basis. Nevertheless, the coverage of the HICP is still different in most national Consumer Price Indices. Thus, for example, the HICP includes information on shelter prices that refer to actual rentals for housing paid by tenants, but actual prices for owner occupied residential property are still excluded up to now. The exclusion of Owner Occupied Housing is – mainly from the European Commission’s as well as the European Central Bank’s point of view – not satisfactory and leaves a gap in the coverage of household final consumption expenditure.

Therefore a pilot study on the coverage of Owner Occupied Housing (development of ‘House Price Indices’) was launched by Eurostat in a limited number of countries (including Germany) in early 2002. The hedonic index computed for this project is the main focus of this paper. Overall, in spite of some difficulties, the main finding is that it seems feasible to compile a hedonic price index for owner occupied housing in Germany at least on a quarterly basis.

² Cf. Linz, Stefan and Eckert, Gudrun, Introducing hedonic methods in price statistics, available at: <http://www.destatis.de/download/e/preise/hedonic-engl.pdf>.

³ Cf. Linz, Stefan and Dexheimer, Verena, Hedonic Methods of Price Measurement for Used Cars, available at http://www.destatis.de/download/e/preise/hed_used_cars.pdf

2. Data sources

In order to avoid the implementation of an extensive new survey on purchases of turnkey-ready houses, existing data has been used for this pilot study. For Germany, mainly the data of the construction price statistics as well as the building activity statistics come into consideration. Nevertheless, there is a lack of data in the field of “turnkey ready buildings” in Germany. In order to close this gap, appropriate data has been acquired from so-called advisory committees of land value (“Gutachterausschüsse für Grundstückswerte”).

The main tasks of these advisory committees consist on the one hand in collecting data on the purchase of buildings and dwellings and on the other hand in estimating the value of houses and land. The advisory committees are mainly organised on a NUTS-3-level (district area level), partly on NUTS-4-level. All of them are independent committees and based on federal (not statistical) law. Furthermore, at present, they are not obliged to provide statistical data on the purchase of buildings. The main problem when using the data of the advisory committees was to coordinate and harmonise the different definitions of variables and concepts used by the committees for data collection. Nevertheless, it was possible to get a lot of detailed data which are at most part appropriate for an Owner Occupied Housing Index. A collaboration within several federal states has successfully been initiated and will be intensified in order to cover additional regions and federal states in future work.

The prices collected and reported by the advisory committees are real transaction prices (and not offer prices) at the time of the conclusion of the contract between the builder (building promoter) and his client. For the pilot study, we could use data from five federal states out of 16 in Germany, which refer to the years 2000 to 2002.

The amount of data reported is quite different: The largest federal state supplies nearly 400 quotations per quarter of a year. The other states report much less data, sometimes only 20 quotations.

As characteristics are reported:

Physical characteristics

- Type of dwelling/property type
 - Single-family house
 - Two-family/semidetached house
 - Terraced house
 - Free-standing house
 - Freehold flat in multi-storey buildings
- Vintage
 - 2000, 2001, 2002; if possible 1999 and 1998
 - Date of purchase (month, day)
- Type of construction
 - Conventional (“Self builders”; not prefabricated)
 - Prefabricated

- Furnishing and luxury elements (Sauna or swimming-pool included)
- Cellar
- Storing position for car (s)

Locational characteristics

- State (“Bundesland”)
 - County
 - City
 - Municipality
 - Part of municipality
- Type of quarter
 - Downtown district
 - Outskirts
 - Rural regions
- Location of building/dwelling in general
 - Simple/plain
 - Medium/average
 - Good/very good

Generally price variables

- Purchase price (real transaction price)
- Size of the (developed) real estate/land/plot of land (in square metres)
- Size of the living area (in square metres)
- Proportionate price of the plot of land in relation to the total purchase price
 - If the proportionate price is not available: specification of the ‘guiding private value’ (in German: Bodenrichtwert”; in EUR/m²). The ‘guiding private value’ is approximately equivalent to the current market price of the location/plot of land.
- Building/dwelling ready for occupancy at the point of time of purchase
- Building/dwelling in stage/phase of building or planning

3. Imputation Method

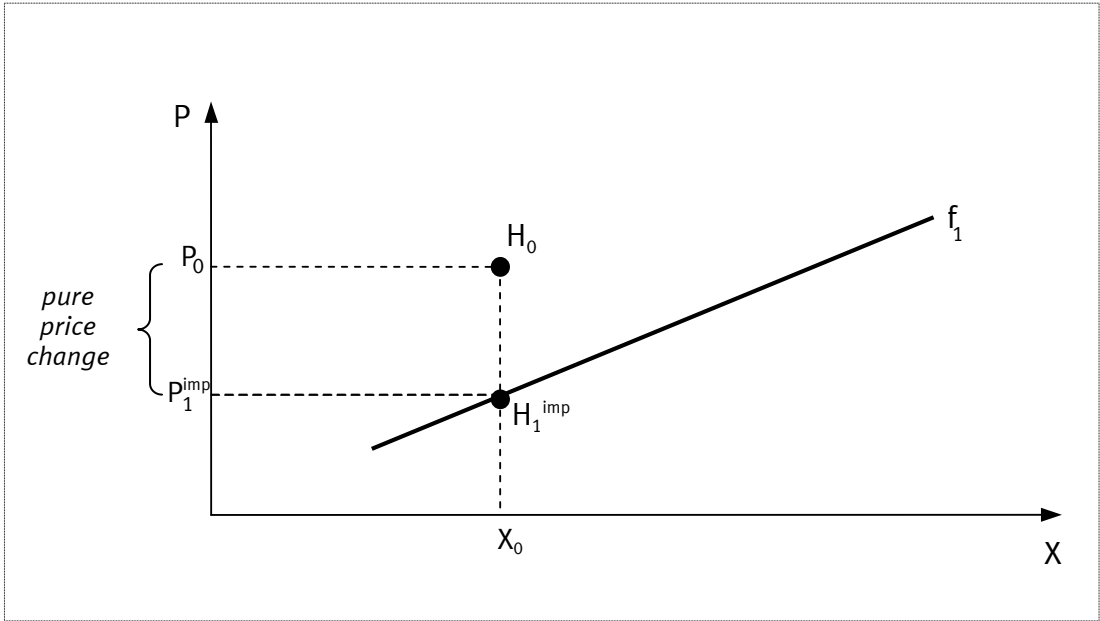
Hedonic methods are well-suited to account for quality changes regarding the heterogeneous product ‘house’ respectively ‘dwelling’. Each house or dwelling can be interpreted as unique, appearing only once on the market. In contrast to other products, a house observed in the base period, will never be exactly the same to be observed in a comparison period. Thus, the application of direct price comparison in order to calculate price indices is not possible.

As hedonic method, the “Imputation Method” is used and missing prices are imputed by means

of regression analysis. In figure 2 below, an example is shown for one House H_0 . The house H_0 has been sold in the base period with a quality of X_0 . The quality X is here meant as quality-vector, which include the values of *several* quality features. The principle is in the picture nevertheless shown for only one dimension of the quality.

We use the regression function f_1 to estimate the price, which the consumer have to pay under the current market conditions for a House with the old quality of X_0 . It is important, that the regression f_1 is calculated by using data of the current period exclusively. Then the difference between P_0 and the estimated (or imputed) price P_1^{imp} gives us the pure price change in the time period considered for the house H_0 .

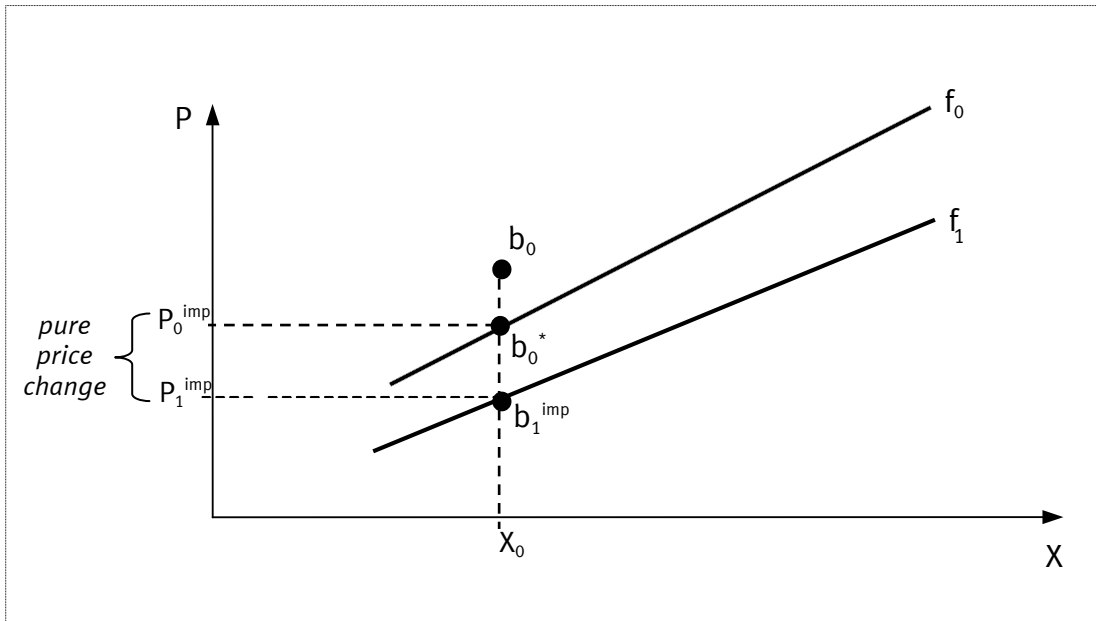
Figure 2: Imputing a current price for an old product



In Figure 2 is shown (for reason of explanation) that the imputed price P_1^{imp} can be compared to the actually observed price P_0 . For the praxis, it is useful to estimate also a price for the base period and to compare the imputed price P_1^{imp} with an also estimated price for $t=0$, which is P_0^* . This procedure is called *Double Imputation*.

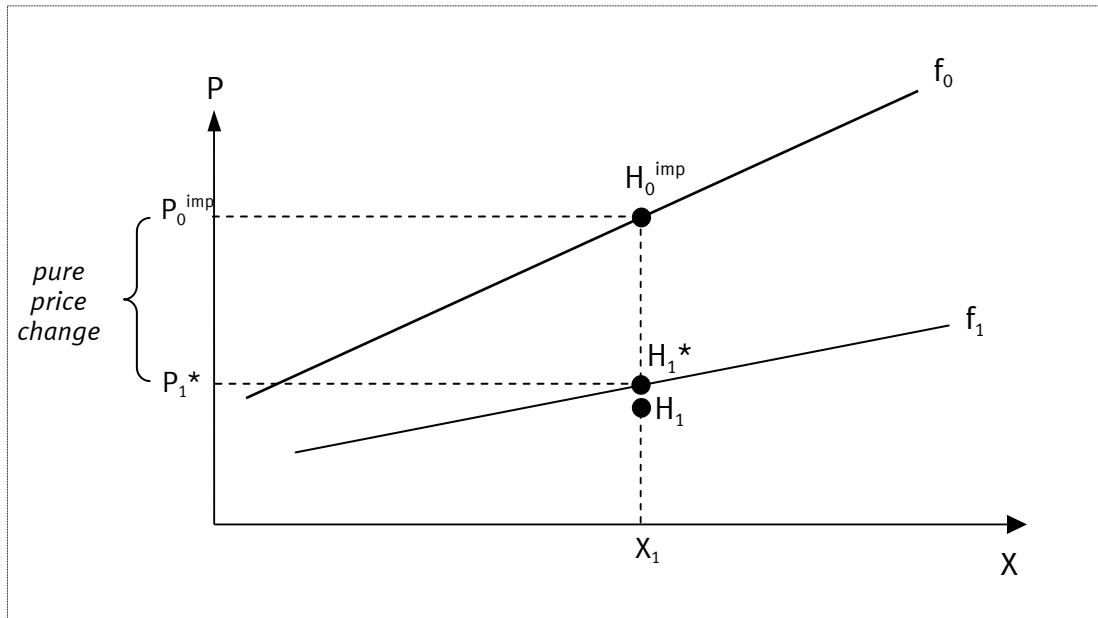
The Double Imputation has the advantage, that outlier do not go into the index calculation. The directly observed price P_0 may be an outlier whereas the estimated price P_0^* is the conditional expected value for this model. Thus, the robustness of the index is increased. (However, the regression functions are still calculated on the directly observed prices.)

Figure 3: Double Imputation



The same procedure can then be used the other way round to impute missing prices for currently observed houses, which have been sold in the current period. Here the regression of the *base period* is used to impute a price which the consumer had to pay in the past period for a house with the *current quality* X_1 . The regression f_0 must here be calculated by using data of the previous period exclusively (see figure 4).

Figure 4: imputation method for new products



To calculate the index, a geometric mean value of prices is computed in each period. In order to be comparable, the two mean values must refer to the same *quality level*. This can be reached by imputing prices for the missing houses: The imputed prices must indicate the price, the average

consumer would have paid in the respective period, when a house with a comparable quality would have existed.

Figure 5: whole sample with imputed and directly observed prices

<i>sub-sample</i>	<i>base period</i>	<i>current period</i>
A	$P_0(1)$	$P_1(1)$
	$P_0(2)$	$P_1(2)$
B	$P_0^{imp}(3)$	$P_1^*(3)$
	$P_0^{imp}(4)$	$P_1^*(4)$
	$P_0^*(5)$	$P_1^{imp}(5)$
	$P_0^*(6)$	$P_1^{imp}(6)$
	geometric mean t=0	geometric mean t=1

$P^{imp}(n)$...imputed price to pay in time t for the quality of product n

The geometric means of the two periods are then compared in order to derive the quality adjusted price development.

4. Regression Analysis

A first step regarding regression analysis is to address the issue of the ‘right’ choice of the functional form. Initially, the ‘linear model’ as well as the ‘log-linear form’ and the ‘log-log model’ were taken into consideration. After testing these various models, the log-log-form which is in its basic shape a multiplicative model and therefore taking into account interactions between variables, has proved to be most robust.

To calculate the index for single-family/two-family houses, the following regression function has been used:

$$(1) \quad \ln(p) = \beta_0 + \beta_1 \cdot \ln(\text{grund}) + \beta_2 \cdot \ln(\text{wohn}) + \beta_3 \cdot \ln(\text{BRW}) + \beta_4 \cdot d_haus + \beta_5 \cdot d_keller + \beta_6 \cdot d_stellplatz + \sum_{i=1}^8 \beta_{6+i} \cdot d_kreis_i + \varepsilon$$

- p ... Price including plot of land in €
- $grund$... size of plot of land [square metres]
- $wohn$... size of living area [square metres]
- BRW ... class of ‘guiding private value’
- d_haus ... dummy for the type of dwelling: 1, when free-standing detached house
0, when semidetached/terraced house
- d_keller ... dummy for cellar/basement: 1, when cellar does exist
0, when cellar does not exist
- $d_stellplatz$... dummy for storing position for car(s): 1=storing pos. for car does exist
0= does not exist
- d_kreis_i ... dummy for area type (1-9): 1, if area type = i , 0, if not

The table 1 shows excerpts of outputs of the regression analysis for the first quarter of 2001 referring to detached/semidetached houses within the state “Lower Saxony”.

Figure 6: results of the regression analysis

variable	parameter	standard-error	t -value	variance inflation factor
intercept	9,681	0,31	31,09	0,00
d_haus	0,189	0,03	6,31	1,45
ln(grund)	0,118	0,03	3,62	1,89
ln(wohn)	0,316	0,07	4,63	1,47
d_kreis_1	0,748	0,10	7,27	1,20
d_kreis_6	-0,068	0,03	-2,15	1,35
d_kreis_9	-0,162	0,10	-1,65	1,09
d_BRW_0-50	-0,165	0,05	-3,40	1,27
d_BRW_100-199	0,143	0,02	6,19	1,41
d_BRW_200-299	0,291	0,06	4,62	1,32
d_keller	0,160	0,02	7,10	1,14
d_parkplatz	0,080	0,02	3,84	1,12

The site of the house plays an important role for the price. The site is represented by dummy variables for the area type as well as by the so called ‘guiding private value’ (Bodenrichtwert). The “guiding private value” is approximately equivalent to the current market price of the location/plot of land. It is classified, so that only the cross-sectional information goes into the regression analysis. An overall price increase of the guiding private value would not be measured as increase of quality. As a result, increasing prices due to higher land prices will be identified as “pure price increases” by the index. The house price index is therefore an index including the price of the house as well as the price of the plot of land.

5. Preliminary Results

The following results point out first experimental indices for turnkey ready buildings with regard to different federal states. In particular, for the East German federal states (Saxony, Saxony-Anhalt, Mecklenburg-Western Pomerania) the samples seem to be too small in part, because the results of the regression analysis for each of these federal states are unsatisfactory due to a lack of observations. Therefore, the data of all the Eastern federal states were pooled and aggregated in order to perform regression analyses which lead to suitable results. Thus, the following results show first price indices for Lower Saxony and East Germany (consisting of the three federal states mentioned above) only. The development of the index is quite similar for Lower Saxony and the Eastern Länder. There is a significant price increase in the 4th quarter of 2001 succeeded by a price decrease in the first quarter of the year 2002. This is probably due to the anticipation by the consumer of a new administrative housing regulations introduced in the beginning of 2002. Also, the number of purchases declined in the beginning of 2002.

Figure 7: price indices for housing in Germany



Overall, in spite of some difficulties, the main finding is that it seems feasible to compile a hedonic price index for owner occupied housing in Germany at least on a quarterly basis.