4

Stratification or Mix Adjustment Methods

Simple Mean or Median Indices

4.1 The simplest measures of house price change are based on some measure of central tendency from the distribution of house prices sold in a period, in particular the mean or the median. Since house price distributions are generally positively skewed (predominantly reflecting the heterogeneous nature of housing, the positive skew in income distributions and the zero lower bound on transaction prices), the median is typically used rather than the mean. As no data on housing characteristics are required to calculate the median, a price index that tracks changes in the price of the median house sold from one period to the next can be easily constructed. Another attraction of median indices is that they are easy to understand.

4.2 An important drawback of simple median indices is that they will provide noisy estimates of price change. The set of houses actually traded in a period, or a sample thereof, is typically small and not necessarily representative of the total stock of housing. Changes in the mix of properties sold will therefore affect the sample median price much more than the median price of the housing stock. For example, think of a city with two regions, A and B, and that region A has more expensive houses than region B. Suppose that the median house sold in 2006 and 2008 comes from region A, while the median house in 2007 comes from region B. It follows that the median index could record a large rise from 2006 to 2007 and then a large fall from 2007 to 2008. Such an index would be a very poor indicator of what is actually happening in the housing market. Thus, a median (or mean) index will be a very inaccurate guide to price change when there is substantial change in the composition of houses sold between periods. If there is a correlation between turning points in house price cycles and compositional change, then a median could be especially misleading in periods when the premium on accuracy is highest.

4.3 A perhaps bigger problem than short-term noise is systematic error, or bias. A simple median index will be subject to bias when the quality of the housing stock changes over time. The median index will be upward biased if the average quality improves over the years. Bias can also arise if certain types of houses are sold more frequently than other types of houses and at the same time exhibit different price changes. For example, when higher quality houses sell more frequently and also rise in price faster than lower quality houses, a downward bias may result if the number of sales per type of house does not properly reflect the number of houses in stock. This is sometimes referred to as a sample selection problem. The fact that houses traded are usually a small and not necessarily representative part of the total housing stock can bias other property price index methods as well, including hedonic and repeat sales methods (to be discussed in Chapters 5 and 6).

Stratification

4.4 Post-stratification of a sample is a general technique for reducing sample selection bias. In the case of residential property price indices, stratification is the simplest tool for controlling for changes in the composition or "quality mix" of the properties sold. The method is therefore also known as mix adjustment. Stratification is also needed if users desire price indices for different housing market segments.

4.5 Stratification is nothing else than separating the total sample of houses into a number of sub-samples or strata. After constructing a measure of the change in the central tendency for each stratum, such as a mean or median price index, the aggregate mix-adjusted RPPI is typically calculated as a weighted average of indices for each stratum. With M different strata, the mix-adjusted index, as calculated in practice in various countries, can be written in mathematical form as follows:

$$P^{0t} = \sum_{m=1}^{M} w_m^0 P_m^{0t} \qquad (4.1)$$

where P_m^{0t} is the index for stratum m which compares the mean (median) price in the current or comparison period t with the mean (median) price in an earlier or base period 0, and where w_m^0 denotes the weight of stratum m. The weights are value shares pertaining to the strata. They refer to the base period, which is usually a year (whereas the comparison periods may be months or quarters). For practical reasons, the weights are often kept fixed for several years, but keeping weights fixed for a long time is generally not good practice. More details on aggregation and weighting issues in this context are provided below.

4.6 Which type of value weights is used, depends on the target index that the RPPI is supposed to estimate. If the purpose is to track the price change of the housing stock then obviously stock-weights – the stock value shares of the strata – should be used. If, on the other hand, the target is a sales or acquisitions RPPI, then sales (expenditure) weights should be applied. [1]

4.7 The effectiveness of stratification will depend upon the stratification variables used because a mix-adjusted measure only controls for compositional change across the various groups. For example, if house sales are separated solely according to their location, a mix-adjusted index will control for changes in the mix of property types across the defined locations. But the mix-adjusted measure will not

[1] The house price indices compiled in the EU as part of a Eurostat pilot study are examples of such acquisitions indices (see Makaronidis and Hayes, 2006 or Eurostat, 2010).

account for any changes in the mix of property types sold that are unrelated to location. Also, a mix-adjusted index does not account for changes in the mix of properties sold within each subgroup, in this case changes in the mix of properties sold within the boundaries of each location.

4.8 Very detailed stratification according to housing characteristics such as size of the structure, plot size, type of dwelling, location and amenities will increase homogeneity and thus reduce the quality-mix problem, although some quality mix changes will most likely remain. There is, however, a tradeoff to be considered. Increasing the number of strata reduces the average number of observations per stratum, and a very detailed stratification might raise the standard error of the overall RPPI. Needless to say, a detailed stratification scheme can be constructed only if the strata-defining characteristics are available for all sample data. Another potential practical problem is that it might be difficult to obtain accurate data on the (stock) weights for small subgroups.

4.9 When using only physical and locational stratification variables, like those mentioned above, then the stratification method does not control for quality changes of the individual properties. By quality changes we mean the effect of renovations and remodeling done to the properties in combination with depreciation of the structures. This can also be called "net depreciation". Depreciation obviously depends on the age of the structure, although depreciation rates may differ across different types of dwellings or even across different locations. This is why age of the structure was listed in Chapter 3 as one of the most important price determining quality attributes. Consequently, stratifying according to age class may help reduce the problem of quality change.

4.10 Introducing age class as another stratification variable will further reduce the average number of observations per stratum and may give rise to unreliable estimates of price changes. Under these circumstances, hedonic regression techniques – which are discussed in Chapter 4 – will generally work better than stratification. As mentioned earlier, some sort of hedonic regression method will also be needed to decompose the overall RPPI into land and structures components if this is required for any of the purposes discussed in Chapter 2. Such a decomposition cannot be provided by stratification methods.

4.11 Mix-adjusted RPPIs have been compiled by numerous statistical offices and other government agencies, including the UK Department of the Environment (1982) and the Australian Bureau of Statistics (ABS, 2006). While mix adjustment has received relatively little attention in the academic literature, (2) there is a growing body of work on market segmentation using statistical techniques like cluster analysis and factor analysis; see e.g. Dale-Johnson (1982), Goodman and Thibodeau (2003), and Thibodeau (2003). These techniques could in principle be used to define housing sub-markets, which could subsequently be used as strata for the construction of a mix-adjusted RPPI. The Australian Bureau of Statistics experimented with this approach (ABS, 2005).

4.12 Prasad and Richards (2006) (2008) proposed a novel stratification method and tested it on an Australian data set. They grouped together suburbs according to the long-term average price level of dwellings in those regions, rather than just clustering smaller geographic regions into larger regions. Their method of stratification was specifically designed to control for what may be the most important form of compositional change, namely changes in the proportion of houses sold in higher- and lower-priced regions in any period. (3) Note that they used median price indices at the stratum level. McDonald and Smith (2009) followed-up on this study and constructed a similar stratified median house price measure for New Zealand.

Aggregation and Weighting Issues

First-stage aggregation

4.13 Stratification involves a two-stage procedure: price indices are compiled at the stratum level, which are then aggregated across the various strata. As was mentioned above, median strata indices have typically been used, in particular because they will often be more stable than the corresponding mean indices. Yet, we will focus on means rather than medians. Conventional index number theory deals with aggregation issues, in this case aggregation of house price observations within strata. Unlike the median, means are aggregator functions, which link up with index number theory. The question then arises: what kind of mean should be taken?

4.14 The CPI Manual (2004) makes recommendations about how to construct price indices at the first stage of aggregation if information on quantities is unavailable and then at the second stage of aggregation when both price and value (or quantity) information is available. At the first stage of aggregation, Chapter 20 in the CPI Manual generally recommends using the unweighted geometric mean or

(2) However, stratified median house price indices have been used by several researchers, mostly for comparison purposes; see e.g. Mark and Goldberg (1984), Crone and Voith (1992), Gatzlaff and Ling (1994), and Wang and Zorn (1997).

(3) A general rule is that stratification according to the variable of interest should not be used since that can lead to biased results. The study variable used by Prasad and Richards (2006) (2008) is (long-term) house price change, not house price level, so their stratification method could perhaps be defended. However, little is known about the statistical properties of this type of stratification index and it would be advisable to investigate the issue of potential bias before producing such an index.

Jevons index to aggregate individual price quotations into an index. However, this general advice is not applicable in the present context.

4.15 If the aim is to construct a price index for the sales of residential properties, the appropriate concept of (elementary) price in some time period t for a homogeneous stratum or cell in the stratification scheme is a *unit value*. Because each sale of a residential property comes with its own quantity, which is equal to one, the corresponding quantity for that cell is the simple *sum* of the properties transacted in period t. We can formally describe this as follows. Suppose that in period t there are $N(t,m)$ property sales observed in a particular cell m, with the selling price (value) of property n equal to V_n^t for $n = 1,...,N(t,m)$. Then the appropriate price and quantity for cell m in period t are:

$$P_m^t \equiv \sum_{n=1}^{N(t,m)} V_n^t / N(t,m) \qquad (4.2)$$

$$Q_m^t \equiv N(t,m) \qquad (4.3)$$

This narrowly defined unit value concept is actually recommended in the CPI Manual (2004; 356). If the stratification scheme leads to cells that are not sufficiently narrow defined, then of course some unit value bias may arise, which is equivalent to saying that some quality mix bias may remain. ([4])

Second-stage aggregation

4.16 The next issue to be resolved is: what index number formula should be used to aggregate the elementary prices and quantities into one overall RPPI? The CPI Manual discusses this choice of formula issue at great length. A number of index number formulae are recommended but a good overall choice appears to be the Fisher ideal index since this index can be justified from several different perspectives. ([5]) The Fisher index is the geometric mean of the Laspeyres and Paasche indices.

4.17 To illustrate this point, let $P^t \equiv [P_1^t,...,P_M^t]$ and $Q^t \equiv [Q_1^t,...,Q_M^t]$ denote the period t vectors of cell prices and quantities. The Laspeyres price index, P_L^{st}, going from (the base) period s to (the comparison) period t can be defined as follows:

$$P_L^{st}(P^s, P^t, Q^s) \equiv \frac{\sum_{m=1}^{M} P_m^t Q_m^s}{\sum_{m=1}^{M} P_m^s Q_m^s} \qquad (4.4)$$

Note that equation (4.4) can be rewritten in the form of (4.1) if $s = 0$ with cell price indices $P_m^{0t} = P_m^t / P_m^0$ and value shares $w_m^0 = P_m^0 Q_m^0 / \sum_{m=1}^{M} P_m^0 Q_m^0$. The Paasche price index going from period s to t, P_P^{st}, is defined as follows:

$$P_P^{st}(P^s, P^t, Q^t) \equiv \frac{\sum_{m=1}^{M} P_m^t Q_m^t}{\sum_{m=1}^{M} P_m^s Q_m^t} \qquad (4.5)$$

The Fisher price index for period t relative to period s, P_F^{st}, can be defined as the geometric mean of (4.4) and (4.5):

$$P_F^{st}(P^s, P^t, Q^s, Q^t) \equiv [P_L^{st}(P^s, P^t, Q^s) \times P_P^{st}(P^s, P^t, Q^t)]^{1/2} \qquad (4.6)$$

Recall that all the quantities occurring in these three formulas are numbers of transactions; that is, numbers of observed prices. Thus, for calculating a Laspeyres, Paasche, or Fisher price index one needs the same information.

4.18 The Laspeyres, Paasche and Fisher price indices defined by equations (4.4), (4.5) and 4.6) are *fixed base indices*. For example, if there are 3 periods of sales data, including the base period 0, then the Fisher formula (4.6) would generate the following index number series for those 3 periods:

$$1; P_F^{01}(P^0, P^1, Q^0, Q^1); P_F^{02}(P^0, P^2, Q^0, Q^2) \qquad (4.7)$$

Chaining

4.19 An alternative to the fixed base method is the use of chaining. The *chain method* uses the data of the last two periods to calculate a period to period chain link index which is used to update the index level from the previous period. Chaining would, for example, generate the following Fisher index number series for the 3 periods:

$$1; P_F^{01}(P^0, P^1, Q^0, Q^1); P_F^{01}(P^0, P^1, Q^1, Q^1) P_F^{12}(P^1, P^2, Q^1, Q^2) \qquad (4.8)$$

4.20 The next issue to be discussed is whether RPPIs should be constructed by using fixed base or chain indices. Both the System of National Accounts and the CPI Manual recommend the use of chain indices provided that the underlying price data have reasonably smooth trends. ([6]) On the other hand, if there is a great deal of variability in the data, particularly when prices bounce erratically around a trend, the use of fixed base indices is recommended. Property price changes tend to be fairly smooth, ([7]) so it is likely that chained indices will work well in many cases. However, more experimentation with actual data is

([4]) In practice, crude stratification according to region and type of dwelling is often used. The stratification method according to price bands proposed by Prasad and Richards (2008), could be useful to militate against unit value bias. See Balk (1998) (2008; 72-74), Silver (2009a) (2009b) (2010), and Diewert and von der Lippe (2010) for more general discussions of unit value bias.

([5]) See CPI Manual (2004; Chapters 15-18) for alternative justifications for the use of the Fisher formula.

([6]) See SNA (2008) and CPI Manual (2004; 349).

([7]) Although prices do not bounce around erratically in the real estate context, quantities do exhibit considerable variability, particularly if there are a large number of cells in the stratification setup with a limited number of observations in each cell. There is also a considerable amount of seasonal variation in quantities; i.e., sales of residential properties fall off dramatically during the winter months of the year.

required in order to give definitive advice on this issue. There may also be seasonal variation in house prices as the example for the Dutch town of "A", presented below, suggests. In such cases too, one should be careful with using chain indices.

Stock RPPIs

4.21 The above discussion was on the construction of a price index for the *sales* of residential properties when using a stratification method. But how should an RPPI be constructed for the *stock* of residential properties? Assuming that, for each cell *m*, the properties sold are random (or 'representative') selections from the stock of dwelling units defined by cell *m*, the period *t* unit value prices P_m^t defined by (4.2) can still be used as (estimates of the) cell prices for a stock RPPI. The quantities Q_m^t defined by (4.3) are, however, no longer appropriate; they need to be replaced by (estimates of) the number of dwelling units of the type defined by cell *m* that are in the reference stock at time *t*, say Q_m^{t*}, for $m = 1,\dots,M$. With these *population* quantity weights, the rest of the details of the index construction are the same as was the case for the sales RPPI.

4.22 To compile stock weights, it will be necessary to have a periodic census of the housing stock with enough details on the properties so that it can be decomposed into the appropriate cells in the stratification scheme for a base period. If information on new house construction and on demolitions is available in a timely manner, then the census information can be updated and estimates for the housing stock by cell (the Q_m^{t*}) can be made in a timely manner. The stock RPPI can be constructed using a (chained) Fisher index as was the case for the sales RPPI. On the other hand, if timely data on new construction and demolitions is lacking, it will only be possible to construct a fixed base Laspeyres index using quantity data from the last available housing census (in say period 0), $Q^{0*} = [Q_1^{0*},\dots,Q_M^{0*}]$, until information from a new housing census is made available (in say period *T*). The Laspeyres stock RPPI thus is

$$P_L^{0t}(P^0, P^t, Q^{0*}) \equiv \frac{\sum_{m=1}^{M} P_m^t Q_m^{0*}}{\sum_{m=1}^{M} P_m^0 Q_m^{0*}} \qquad (4.9)$$

$$t = 0,\dots,T$$

4.23 In Chapter 3 it was mentioned that for some purposes it is useful to have a stock RPPI for Owner Occupied Housing, i.e. excluding rented homes. The construction of such an index proceeds in the same way as for the construction of an RPPI for the entire housing stock except that the cells in the stratification scheme are now restricted to owner occupied dwellings. This will be possible if the periodic housing census collects information on whether each dwelling unit is owned or rented.

4.24 It should be noted that the construction of a stratified (stock or sales) RPPI becomes more complex when some of the cells in the stratification scheme are empty for some periods. At the end of this chapter, where an empirical example using data on housing sales for the Dutch town of "A" is presented, a matched-model approach will be outlined that can be used in case some cells are empty.

Main Advantages and Disadvantages

4.25 We will summarize the main advantages and disadvantages of the stratified median or mean approach. The main advantages are:

- Depending on the choice of stratification variables, the method adjusts for compositional change of the dwellings.
- The method is reproducible, conditional on an agreed list of stratification variables.
- Price indices can be constructed for different types and locations of housing.
- The method is relatively easy to explain to users.

4.26 The main disadvantages of the stratified median or mean method are:

- The method cannot deal adequately with depreciation of the dwelling units unless age of the structure is a stratification variable.
- The method cannot deal adequately with units that have undergone major repairs or renovations (unless renovations are a stratification variable).
- The method requires information on housing characteristics so that sales transactions can be allocated to the correct strata.
- If the classification scheme is very coarse, compositional changes will affect the indices, i.e., there may be some unit value bias in the indices.
- If the classification scheme is very fine, the cell indices may be subject to a considerable amount of sampling variability due to small sample sizes or some cells may be empty for some periods causing index number difficulties.

4.27 An overall evaluation of the stratification method is that it can be satisfactory if:

- an appropriate level of detail is chosen;
- age of the structure is one of the stratification variables, and
- a decomposition of the index into structure and land components is not required.

Stratification can be interpreted as a special case of regression.(8) Chapter 5 discusses this more general technique, known as hedonic regression when applied to price index construction and quality adjustment.

An Example Using Dutch Data for the Town of "A"

4.28 This chapter will be concluded by a worked example for the construction of a stratified index using data on sales of detached houses for a small town (the population is around 60 000) in the Netherlands, town "A", for 14 quarters, starting in the first quarter of 2005 and ending in the second quarter of 2008. The same data set will be exploited in Chapters 5, 6, 7 and 8 to illustrate the other methods for constructing house price indices and the numerical differences that can arise in practice.(9)

4.29 A dwelling unit has a number of important *price determining characteristics*:

- The land area of the property;
- The floor space area of the structure; i.e., the size of the structure that sits on the land underneath and surrounding the structure;
- The age of the structure; this determines (on average) how much physical deterioration or depreciation the structure has experienced;
- The amount of renovations that have been undertaken for the structure;
- The location of the structure; i.e., its distance from amenities such as shopping centers, schools, restaurants and work place locations;
- The type of structure; i.e., single detached dwelling unit, row house, low rise apartment or high rise apartment or condominium;
- The type of construction used to build the structure;
- Other special price determining characteristics that are different from "average" dwelling units in the same general location such as swimming pools, air conditioning, elaborate landscaping, the height of the structure or views of oceans or rivers.

The variables used in this study can be described as follows:

- V_n^t is the selling price of property n in quarter t in Euros;
- L_n^t is the area of the plot for the sale of property n in quarter t in meters squared;
- S_n^t is the living space area of the structure for the sale of property n in quarter t in meters squared;
- A_n^t is the approximate age (in decades) of the structure on property n in quarter t.

4.30 It can be seen that not all of the price determining characteristics listed above were used in the present study. In particular, the last five sets of characteristics of the property were neglected. There is an implicit assumption that quarter to quarter changes in the amount of renovations that have been undertaken for the structures, the location of the house, the type of structure, the type of construction and any other price determining characteristics of the properties sold in the quarter did not change enough to be a significant determinant of the average price for the properties sold once changes in land size, structure size and the age of the structures were taken into account.(10)

4.31 The determination of the values for the age variable A_n^t needs some explanation. The original data were coded as follows: if the structure was built in 1960-1970, then the observation was assigned the decade indicator variable BP = 5; 1971-1980, BP=6; 1981-1990, BP=7; 1991-2000, BP=8; 2001-2008, BP=9. The age variable in this study was set equal to 9 - BP. For a recently built structure n in quarter t, $A_n^t = 0$. Thus, the age variable gives the (approximate) age of the structure in decades.

4.32 Houses which were older than 50 years at the time of sale were deleted from the data set. Two observations which had unusually low selling prices (36 000 and 40 000 Euros) were deleted as were 28 observations which had land areas greater than 1200 m^2. No other outliers were deleted from the sample. After this cleaning of the data, we were left with 2289 observations over the 14 quarters in the sample, or an average of 163.5 sales of detached dwelling units per quarter. The overall sample mean selling price was 190 130 Euros, whereas the median price was 167 500 Euros. The average plot size was 257.6 m^2 and the average size of the structure (living space area) was 127.2 m^2. The average age of the properties sold was approximately 18.5 years.

4.33 The stratification approach to constructing a house price index is conceptually very simple: for each of the important price explaining characteristic, divide up the sales into relatively homogeneous groups. Thus in the present case, sales were classified into 45 groups or cells, consisting of 3 groupings for the land area L, 3 groupings

(8) See Diewert (2003a) who showed that stratification techniques or the use of dummy variables can be viewed as a nonparametric regression technique. In the statistics literature, these partitioning or stratification techniques are known as analysis of variance models; see Scheffé (1959).

(9) This material is drawn from Diewert (2010).

(10) To support this assumption, it should be noted that the hedonic regression models discussed in later chapters consistently explained 80-90% of the variation in the price data using just the three main explanatory variables: L, S and A. The R^2 between the actual and predicted selling prices ranged from .83 to .89. The fact that it was not necessary to introduce more price determining characteristics for this particular data set can perhaps be explained by the nature of the location of the town of "A" on a flat, featureless plain and the relatively small size of the town; i.e., location was not a big price determining factor since all locations have more or less the same access to amenities.

for the structure area S and 5 groups for the age A (in decades) of the structure (3´3´5 = 45 separate cells). Once quarterly sales were classified into the 45 groupings of sales, the sales within each cell in each quarter were summed and then divided by the number of units sold in that cell in order to obtain unit value prices, the cell prices P_m^t. These unit values were then combined with the number of units sold in each cell, the Q_m^t, to form the usual p's and q's that can be inserted into a bilateral index number formula, like the Laspeyres, Paasche and Fisher ideal formulae defined by (4.4)-(4.6) above, ([11]) yielding a stratified index of house prices of each of these types. However, since there are only 163 or so observations for each quarter and 45 cells to fill, each cell had only an average of 3 or so observations in each quarter, and some cells were empty for some quarters. This problem will be addressed subsequently.

4.34 How should the size limits for the L and S groupings be chosen? One approach would be to divide the range of L and S by three and create three equal size cells. However, this approach leads to a large number of observations in the middle cells. In the present study, size limits were therefore chosen such that roughly 50 % of the observations would fall into the middle sized categories and roughly 25 % would fall into the small and large categories. For the land size variable L, the cutoff points chosen were 160 m^2 and 300 m^2, while for the structure size variable S, the cutoff points chosen were 110 m^2 and 140 m^2. Thus if L < 160 m^2, then the observation fell into the small land size cell; if 160 m^2 ≤ L < 300 m^2, then the observation fell into the medium land size cell and if 300 m^2 ≤ L, then the observation fell into the large land size cell. The resulting sample probabilities for falling into these three L cells over the 14 quarters were .24, .51 and .25 respectively. Similarly, if S < 110 m^2, the observation fell into the small structure size cell; if 110 m^2 ≤ S < 140 m^2, then the observation fell into the medium structure size cell and if 140 m^2 ≤ S, then the observation fell into the large structure size cell. The resulting sample probabilities for falling into these three S cells over the 14 quarters were .21, .52 and .27 respectively.

4.35 As mentioned earlier, the data that were used did not have an exact age for the structure; only the decade when the structure was built was recorded. So there was no possibility of choosing exact cutoff points for the age of the structure. $A = 0$ corresponds to houses that were built during the years 2001-2008; $A = 1$ for houses built in 1991-2000; $A = 2$ for houses built in 1981-1990, $A = 3$ for houses built in 1971-1980; and $A = 4$ for houses built in 1961-1970. The resulting sample probabilities for falling into these five cells over the 14 quarters were .15, .32, .21, .20 and .13 respectively. See Table 4.1 for the sample joint probabilities of a house sale belonging to each of the 45 cells.

4.36 There are several points of interest to note about Table 4.1:

- There were no observations for houses built during the 1960s ($A = 4$) which had a small lot (L = small) and a large structure (S = large), so this cell is entirely empty;
- There are many cells which are almost empty; in particular the probability of a sale of a large plot with a small house is very low as is the probability of a sale of a small plot with a large house; ([12])
- The "most representative model" sold over the sample period corresponds to a medium sized lot, a medium sized structure and a house that was built in the 1990s ($A = 1$). The sample probability of a house sale falling into this highest probability cell is 0.09262.

([11]) The international manuals on price measurement recommend this unit value approach to the construction of price indices at the first stage of aggregation; see CPI Manual (2004), PPI Manual (2004), and XMPI Manual (2009). However, the unit value aggregation should take place over homogeneous items and this assumption may not be fulfilled in the present context, since there is a fair amount of variability in L, S and A within each cell. But since there are only a small number of observations in each cell for the data set under consideration, it would be difficult to introduce more cells to improve homogeneity since this would lead to an increased number of empty cells and a lack of matching for the cells.

([12]) Thus lot size and structure size are positively correlated with a correlation coefficient of .6459. Both L and S are fairly highly correlated with the selling price variable P: the correlation between P and L is .8234 and between P and S is .8100. These high correlations lead to multicollinearity problems in the hedonic regression models to be considered later.

Table 4.1. Sample Probability of a Sale in Each Cell

L	S	A = 0	A = 1	A = 2	A = 3	A = 4
small	small	0.00437	0.02665	0.01660	0.02053	0.02097
medium	small	0.00349	0.02840	0.01966	0.01092	0.03888
large	small	0.00087	0.00175	0.00044	0.00218	0.00612
small	medium	0.01223	0.05242	0.04281	0.02053	0.00699
medium	medium	0.03277	0.09262	0.08869	0.07907	0.02141
large	medium	0.00786	0.02315	0.01005	0.01442	0.01398
small	large	0.00306	0.00218	0.00175	0.00568	0.00000
medium	large	0.03145	0.03495	0.00786	0.02097	0.00306
large	large	0.04893	0.05461	0.02315	0.02490	0.01660

Source: Authors' calculations based on data from the Dutch Land Registry

4.37 The average selling price of the representative house, falling into the medium L, medium S and $A=1$ category, is graphed in Figure 4.1 along with the overall sample mean and median price in each quarter. These average prices have been converted into indices which start at 1 for quarter 1, which is the first quarter of 2005. It should be noted that these three house price indices are rather variable.

4.38 Some additional indices are plotted in Figure 4.1, including a fixed base matched model Fisher index and a chained matched model Fisher price index. It is necessary to explain what a "matched model" index in this context means. If at least one house was sold in each quarter for each of the 45 cells, the ordinary Laspeyres, Paasche and Fisher price indices comparing the prices of quarter t to those of quarter s would be defined by equations (4.4)-(4.6) respectively, where M = 45. This algebra is applicable to the situation where there are transactions in all cells for the two quarters being compared. But for the present data set, on average only about 30 out of the 45 categories can be matched across any two quarter, and the formulae (4.4)-(4.6) need to be modified in order to deal with this *lack of matching problem*. Thus, when considering how to form an index number comparison between quarters s and t, define the set of cells m that have at least one transaction in each of quarters s and t as the set $S(s,t)$. Then the *matched model counterparts*, P_{ML}^{st}, P_{MP}^{st} and P_{MF}^{st}, to the regular Laspeyres, Paasche and Fisher indices between quarters s and t given by (4.4), (4.5) and (4.6) are defined as follows: (13)

$$P_{ML}^{st} \equiv \frac{\sum_{m \in S(s,t)} P_m^t Q_m^s}{\sum_{m \in S(s,t)} P_m^s Q_m^s} \tag{4.10}$$

$$P_{MP}^{st} \equiv \frac{\sum_{m \in S(s,t)} P_m^t Q_m^t}{\sum_{m \in S(s,t)} P_m^s Q_m^t} \tag{4.11}$$

$$P_{MF}^{st} \equiv [P_{ML}^{st} P_{MP}^{st}]^{1/2} \tag{4.12}$$

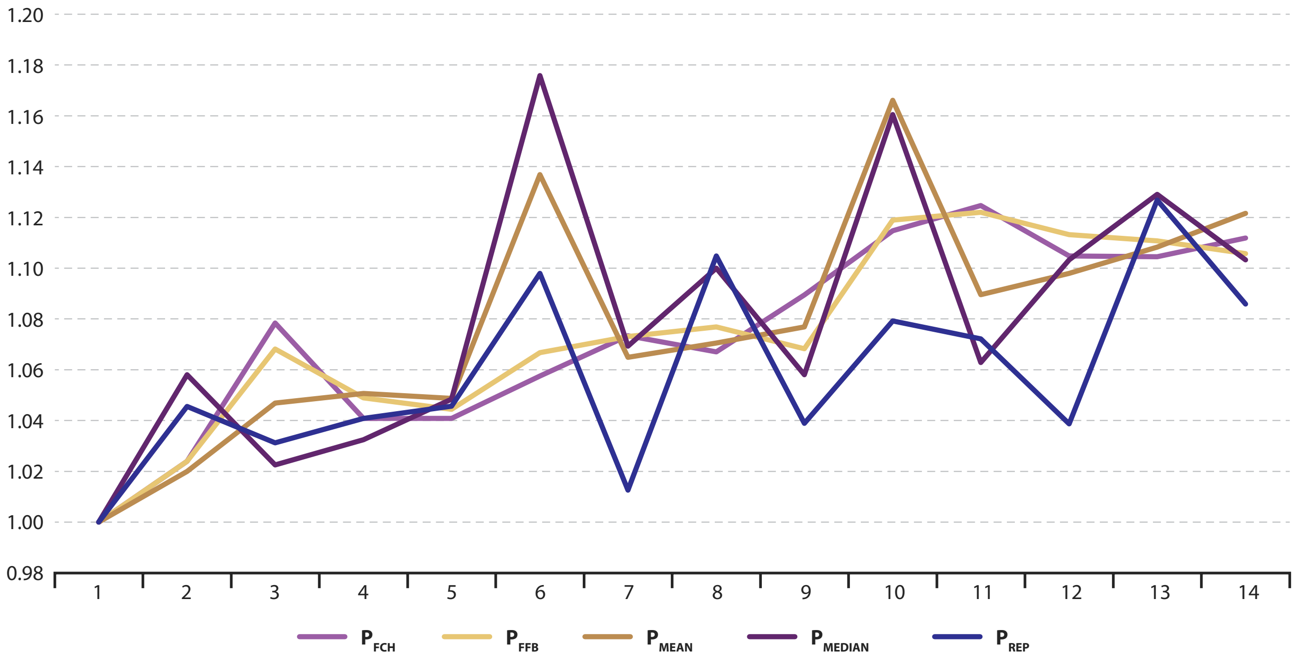
In Figure 4.1, the Fixed Base Fisher index is the matched model Fisher price index defined by (4.12), where the base period s is kept fixed at quarter 1; i.e., the indices $P_{MF}^{1,1}$, $P_{MF}^{1,2}$,…, $P_{MF}^{1,14}$ are calculated and labeled as the Fixed Base Fisher Index, *PFFB*. The index that is labeled the matched model Chained Fisher Index, *PFCH*, is the price index $P_{MF}^{1,1}$, $P_{MF}^{1,1} P_{MF}^{1,2}$, $P_{MF}^{1,1} P_{MF}^{1,2} P_{MF}^{2,3}$, …, $P_{MF}^{1,1} P_{MF}^{1,2} \ldots P_{MF}^{13,14}$ $P_{MF}^{1,14}$. Notice that the Fixed Base and Chained (matched model) Fisher indices are quite close to each other and are much smoother than the corresponding Mean, Median and Representative Model indices. (14) The data for these 5 series plotted in Figure 4.1 are listed in Table 4.2.

4.39 The matched model Fisher indices must be regarded as being more accurate than the other indices which use only a limited amount of the available price and quantity information. As the trend of the Fisher indices is fairly smooth, the chained Fisher index should be preferred over the fixed base Fisher index, following the advice given in Hill (1988) (1993) and in the CPI Manual (2004). Recall also that there is no need to use Laspeyres or Paasche indices in this situation since data on sales of houses contains both value and quantity information. Under these conditions, Fisher indices are preferred over the Laspeyres and Paasche indices (which do not use all of the available price and quantity information for the two periods being compared).

(13) A justification for this approach to dealing with a lack of matching in the context of bilateral index number theory can be found in the discussion by Diewert (1980; 498-501) on the related problem of dealing with new and disappearing goods. Other approaches are also possible. For approaches based on maximum matching over all pairs of periods; see Ivancic, Diewert and Fox (2011) and de Haan and van der Grient (2011) for approaches based on imputation methods; see Alterman, Diewert and Feenstra (1999). A useful imputation approach could be to estimate imputed prices for the empty cells using hedonic regressions. The discussion is left until various hedonic regression methods have been discussed.

(14) The means (and standard deviations) of the 5 series mentioned thus far are as follows: P_{FCH} = 1.0737 (0.0375), P_{FFB} = 1.0737 (0.0370), P_{Mean} = 1.0785 (0.0454), P_{Median} = 1.0785 (0.0510), and $P_{Represent}$ = 1.0586 (0.0366). Thus the representative model price index has a smaller variance than the two matched model Fisher indices but it has a substantial bias relative to the two matched model Fisher indices: the representative model price index is well below the Fisher indices for most of the sample period.

Figure 4.1. Matched Model Fisher Chained and Fixed Base Price Indices, Mean, Median and Representative Model Price Indices



Source: Authors' calculations based on data from the Dutch Land Registry

Table 4.2. Matched Model Fisher Chained and Fixed Base Price Indices, Mean, Median and Representative Model Price Indices

Quarter	P_{FCH}	P_{FFB}	P_{Mean}	P_{Median}	$P_{Represent}$
1	1.00000	1.00000	1.00000	1.00000	1.00000
2	1.02396	1.02396	1.02003	1.05806	1.04556
3	1.07840	1.06815	1.04693	1.02258	1.03119
4	1.04081	1.04899	1.05067	1.03242	1.04083
5	1.04083	1.04444	1.04878	1.04839	1.04564
6	1.05754	1.06676	1.13679	1.17581	1.09792
7	1.07340	1.07310	1.06490	1.06935	1.01259
8	1.06706	1.07684	1.07056	1.10000	1.10481
9	1.08950	1.06828	1.07685	1.05806	1.03887
10	1.11476	1.11891	1.16612	1.16048	1.07922
11	1.12471	1.12196	1.08952	1.06290	1.07217
12	1.10483	1.11321	1.09792	1.10323	1.03870
13	1.10450	1.11074	1.10824	1.12903	1.12684
14	1.11189	1.10577	1.12160	1.10323	1.08587

Source: Authors' calculations based on data from the Dutch Land Registry

4.40 Since there is a considerable amount of heterogeneity in each cell of the stratification scheme, there is the strong possibility of some unit value bias in the matched model Fisher indices. However, if a finer stratification were used, the amount of matching would drop dramatically. Already, with the present stratification, only about 2/3 of the cells could be matched across any two quarters. There is a trade-off between having too few cells with the possibility of unit value bias and having a more detailed stratification scheme but with a much smaller degree of matching of the data within cells across the two time periods being compared.

4.41 Looking at Table 4.2 and Figure 4.1, it can be seen that the chained Fisher index shows a drop in house prices during the fourth quarters of 2005, 2006 and 2007. There is a possibility that house prices drop for seasonal reasons in the fourth quarter of a year. In order to deal with this possibility, in the next section a rolling year matched model Fisher index will be constructed.

The Treatment of Seasonality for the Dutch Example

4.42 Assuming that each commodity in each season of the year is a separate "annual" commodity is the simplest and theoretically most satisfactory method for dealing with seasonal goods when the goal is to construct *annual* price and quantity indices. This idea can be traced back to Mudgett in the consumer price context and to Stone in the producer price context:

"The basic index is a yearly index and as a price or quantity index is of the same sort as those about which books and pamphlets have been written in quantity over the years." Bruce D. Mudgett (1955; 97).

"The existence of a regular seasonal pattern in prices which more or less repeats itself year after year suggests very strongly that the varieties of a commodity available at different seasons cannot be transformed into one another without cost and that, accordingly, in all cases where seasonal variations in price are significant, the varieties available at different times of the year should be treated, in principle, as separate commodities." Richard Stone (1956; 74-75).

Diewert (1983) generalized the Mudgett-Stone annual framework to allow for *rolling year comparisons* for 12 consecutive months of data with a base year of 12 months of data or for comparisons of 4 consecutive quarters of data with a base year of 4 consecutive quarters of data; i.e., the basic idea is to compare the current rolling year of price and quantity data to the corresponding data of a base year where the data pertaining to each season is compared.(15) In the present context, we have in principle,(16) price and quantity data for 45 classes of housing commodities in each quarter. If the sale of a house in each season is treated as a separate good, then there are 180 annual commodities.

4.43 For the first index number value, the four quarters of price and quantity data on sales of detached dwellings in the town of "A" (180 series) are compared with the same data using the Fisher ideal formula. Naturally, the resulting index is equal to 1. For the next index number value, the data for the first quarter of 2005 are dropped and the data pertaining to the first quarter of 2006 are appended to the data for quarters 2-4 of 2005. The resulting Fisher index is the second entry in the Rolling Year (RY) Matched Model series that is illustrated in Figure 4.2. However, as was the case with the chained and fixed base Fisher indices that appeared in Figure 4.1, not all cells could be matched using the rolling year methodology; i.e., some cells were empty in the first quarter of 2006 which corresponded to cells in the first quarter of 2005 which were not empty and vice versa. So when constructing the rolling year index P_{RY} plotted in Figure 4.2, the comparison between the rolling year and the data pertaining to 2005 was restricted to the set of cells which were non empty in both years; i.e., the Fisher rolling year indices plotted in Figure 4.2 are matched model indices. Unmatched models are omitted from the index number comparison.(17)

4.44 The results are shown in Figure 4.2. Note that there is a definite downturn at the end of the sample period but that the downturns which showed up in Figure 4.1 for quarters 4 and 8 can be interpreted as seasonal downturns; i.e., the rolling year indices in Figure 4.2 did not turn down until the end of the sample period. Note further that the index value for observation 5 compares the data for calendar year 2006 to the corresponding data for calendar year 2005 and the index value for observation 9 compares the data for calendar year 2007 to the corresponding data for calendar year 2005; i.e., these index values correspond to Mudgett-Stone annual indices.

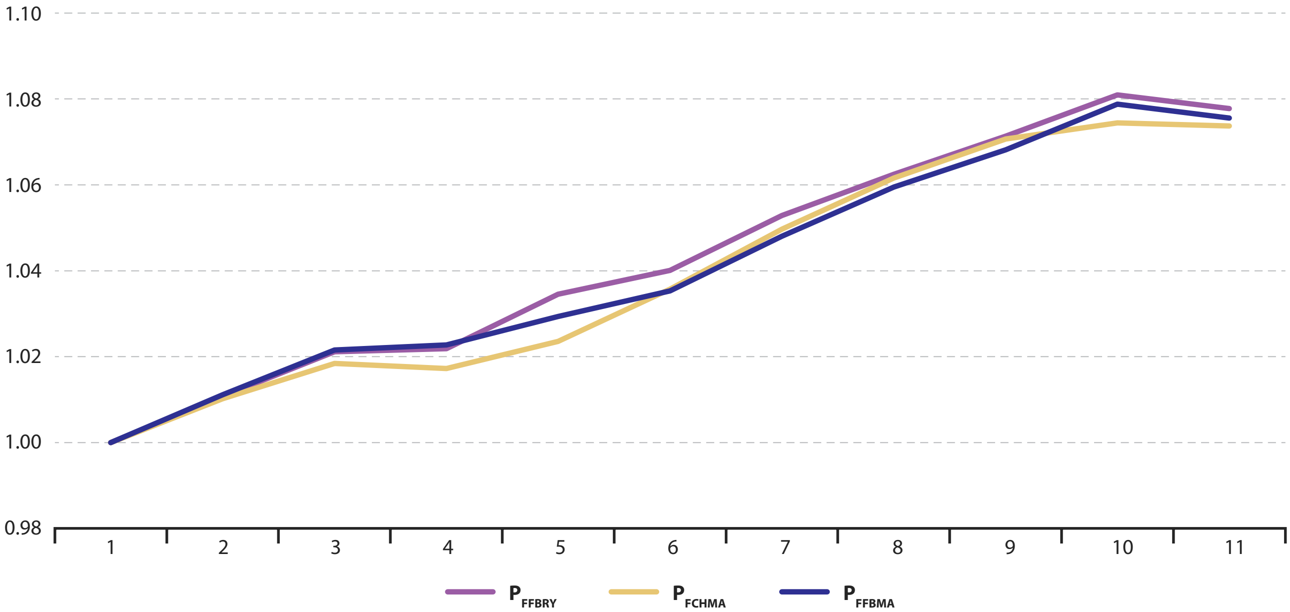
(15) For additional theory and examples of this rolling year approach, see the chapters on seasonality in the CPI Manual (2004) and the PPI Manual (2004), Diewert (1998), and Balk (2008; 151-169). To justify the rolling year indices from the viewpoint of the economic approach to index number theory, some restrictions on preferences are required; details can be found in Diewert (1999; 56-61). It should be noted that weather and the lack of fixity of Easter can cause "seasons" to vary and a breakdown in the approach; see Diewert, Finkel and Artsev (2009). However, with quarterly data, these limitations of the rolling year index are less important.

(16) In practice, as we have seen in the previous section, many of the cells are empty in each period.

(17) There are 11 rolling year comparisons that can be made with the data for 14 quarters that are available. The numbers of unmatched or empty cells for rolling years 2, 3, ..., 11 are as follows: 50, 52, 55, 59, 60, 61, 65, 65, 66, 67. The relatively low number of unmatched or empty cells for rolling years 2, 3 and 4 is due to the fact that for rolling year 2, ¾ of the data are matched, for rolling year 3, ½ of the data are matched and for rolling year 4, ¼ of the data are matched.

4.45 It is a fairly labour intensive job to construct the rolling year matched model Fisher indices because the cells that are matched over any two periods vary with the periods. A short-cut method (which is less accurate) for seasonally adjusting a series, such as the matched model chained Fisher index P_{FCH} and the fixed base Fisher index P_{FFB} listed in Table 4.2, is to simply take a 4 quarter *moving average* of these series. The resulting rolling year series, P_{FCHMA} and P_{FFBMA}, can be compared with the rolling year Mudgett-Stone-Diewert series P_{RY}; see Figure 4.2. The data that corresponds to Figure 4.2 are listed in Table 4.3.

Figure 4.2. Rolling Year Fixed Base Fisher, Fisher Chained Moving Average and Fisher Fixed Base Moving Average Price Indices



Source: Authors' calculations based on data from the Dutch Land Registry

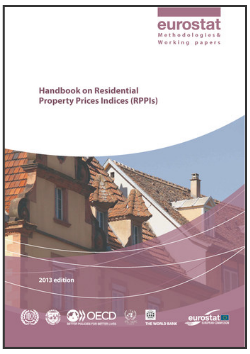
Table 4.3. Rolling Year Fixed Base Fisher, Fisher Chained Moving Average and Fisher Fixed Base Moving Average Price Indices

Rolling Year	P_{FFBRY}	P_{FCHMA}	P_{FFBMA}
1	1.00000	1.00000	1.00000
2	1.01078	1.01021	1.01111
3	1.02111	1.01841	1.02156
4	1.02185	1.01725	1.02272
5	1.03453	1.02355	1.02936
6	1.04008	1.03572	1.03532
7	1.05287	1.04969	1.04805
8	1.06245	1.06159	1.05948
9	1.07135	1.07066	1.06815
10	1.08092	1.07441	1.07877
11	1.07774	1.07371	1.07556

Source: Authors' calculations based on data from the Dutch Land Registry

4.46 It can be seen that a moving average of the chained and fixed base Fisher quarter to quarter indices, P_{FCH} and P_{FFB}, listed in Table 4.2, approximates the theoretically preferred rolling year fixed base Fisher index P_{FFBRY} fairly well. There are differences of up to 1 % between the preferred rolling year index and the moving average index, however. Recall that the fixed base Fisher index compared the data of quarters 1 to 14 with the corresponding data of quarter 1. Thus the observations for, say, quarters 2 and 1, 3 and 1, and 4 and 1 are not as likely to be as comparable as the rolling year indices where data in any one quarter is always lined up with the data in the corresponding quarter of the base year. A similar argument applies to the moving average index P_{FCHMA}; the comparisons that go into the links in this index are from quarter to quarter and they are unlikely to be as accurate as comparisons across the years for the same quarter. ([18])

([18]) The stronger is the seasonality, the stronger will be this argument in favour of the accuracy of the rolling year index. The strength of this argument can be seen if all house price sales for each cell turn out to be strongly seasonal; i.e., the sales for any given cell occur in only one quarter in each year. Quarter to quarter comparisons are obviously impossible in this situation but rolling year indices will be perfectly well defined.



From:

Handbook on Residential Property Price Indices

Access the complete publication at:

https://doi.org/10.1787/9789264197183-en

Please cite this chapter as:

de Haan, Jan and Erwin Diewert (2013), "Stratification or Mix Adjustment Methods", in OECD*, et al.*, *Handbook on Residential Property Price Indices*, Eurostat, Luxembourg.

DOI: https://doi.org/10.1787/9789264197183-6-en