

Housing price statistics in Finland - Monthly prices of dwellings, regional division in housing price statistics and usage of administrative data

Authors:

Tomi Martikainen

Petteri Rautaporras

Table of contents

1. INTRODUCTION.....	3
2. MONTHLY PRICES OF DWELLINGS	3
2.1 Monthly index	3
2.2 Hedonic quality adjustment	4
3. CLASSIFICATION AND REGIONAL DIVISIONS.....	6
3.1 Classification	6
3.2 Regional division	6
4. USAGE OF ADMINISTRATIVE DATA	7
4.1 Challenges of using administrative data.....	8
5. CONCLUSIONS	9
References.....	10

1. Introduction

Statistics Finland has produced quite extensive statistics on dwelling prices for a number of years already. The quarterly house price index for existing (old) dwellings in housing companies has been published since 1985 retrospectively. Historical series are available for some regions from 1970 onwards. Prior to year 2002 prices of dwellings index was formed using data from major real estate agents. Since 2002 index has been calculated using asset transfer tax data gathered by Finnish Tax Administration. Statistics Finland has obtained information on transaction prices of apartments in housing companies on quarterly basis since 2001 and monthly basis from the beginning of 2009. Before 2001 price data was received more sparsely- or on yearly basis. This change in availability of the data has given the opportunity to develop and publish prices of dwellings in housing companies on monthly basis.

Statistics Finland also publish information on single-family homes (statistics on real estate prices) and new dwellings in housing companies which are published quarterly together with information on old dwellings in housing companies. At summer 2013 Statistics Finland starts to release real estate prices for new single-family homes and also indices of owner-occupied housing. Indices of owner-occupied housing prices describe the costs of purchasing and owning owner-occupied dwellings, as well as the development of purchase prices of dwellings. The index of owner-occupied housing prices describes the development of transaction prices of dwellings and of other goods and services related to owner-occupied dwellings. The price index of dwellings describes the price development of dwellings bought by households (HPI). The Housing Price Index consists of the price indices of new and old dwellings in housing companies, the price index on real estate prices, and the price index of new detached houses.

This paper describes development made at Statistics Finland on compiling short-term housing statistics. The second chapter explains the monthly index in more detailed level. There is information on the method that is used on compiling the index and some information on data accumulation. Statistics Finland has formed specific regional divisions based on price and geographical location. Regional divisions are explained on third chapter. The fourth chapter presents information on how administrative data are used at Statistics Finland and which are the challenges of using these extensive data sets.

2. Monthly prices of dwellings

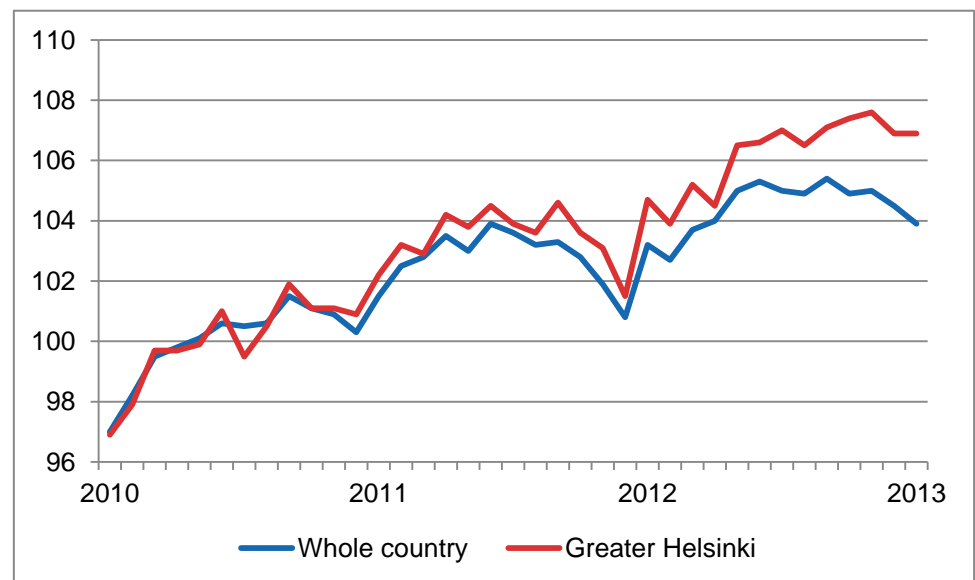
2.1 Monthly index

The reason in developing monthly index for prices of dwellings was that the demand for more current and faster information rose from the users. In addition, the growing interest in housing prices at EU-level has given the impulse to develop monthly series. Monthly index is produced

utilising the old statistical production process, and its management and architecture so that the process is consistent with the quarterly statistics production process. This change has made it possible to carry out data validation and some other parts of the editing process monthly. The calculation of monthly indices takes place approximately one month from the end of the examined month. The quarterly index is compiled with same schedule, one month from end of examined quarter.

The index is compiled by combining the monthly data (i.e. asset transfer taxation records) with the register data (i.e. buildings and dwellings register and real estate register) and the data are updated by using the municipal classification of the year concerned. The obtained data is validated by adding to them data on state subsidised housing, setting price thresholds and by supplementing missing data on e.g. number of rooms, year of completion and type of building.

Picture 1. Monthly index of prices of dwellings in housing companies, 2010=100



When the monthly index is published for the first time, it covers around two-fifths of the transactions in the month concerned. The rest of the data are obtained later as the transactions are entered into the asset transfer taxation records. One of the reasons for the lag is that the buyer of the share in a housing company has two months to pay the asset transfer tax. Monthly and quarterly data are revised retrospectively in connection with each release so that the final data for the statistical year are published with the data for the first quarter of the year following it. Appendix 1 shows how the number of made transactions accumulates in the asset transfer taxation records.

2.2 Hedonic quality adjustment

In Statistics Finland's housing price statistics the repeat sales problem has been solved by utilising a calculation method that combines conventional data classification and regression analysis (hedonic regression). The method helps to distinguish the real price development

from price changes caused by dwelling characteristics in different time periods. The method is also practicable in other index calculations and a respective method has been applied in statistics on real estate prices.

The used classification does not necessarily homogenise the data sufficiently. Factors affecting price, such as micro-location, floor area, year of completion, and so on, are not controlled for by the classification. The available data contains information on these characteristics, which can be used for adjusting the average price of a given category in the comparison period so that the obtained average price adjusted for quality takes into account compositional changes within the category in the base and comparison periods.

The regression models are of the following format:

$$(1) \quad \ln(p_{ij}) = \beta_0 + \beta_1 \times x_{ij1} + \dots + \beta_n \times x_{ijn} + \varepsilon_{ij}.$$

The notation of the model is a standard, half-logarithmic function.

$\ln(p_{ij})$ is the dependent variable: the logarithmic price of observation j in class i . Variables $x_{ij1} \dots x_{ijn}$ are explanatory variables of the model, which are selected according to the available data on the basis of statistical analysis. Numeric variables or classified dummy variables can be used as the explanatory variables. $\beta_0, \beta_1, \dots, \beta_n$ are the parameters to be estimated for the model. ε_{ij} is the error term of the model.

The overall index for prices of dwellings is obtained via the log-Laspeyres formula:

$$(2) \quad index_0^1 = \exp\left(\sum_{i=1}^N w_0^i \ln\left(\frac{\bar{p}_1^i}{\bar{p}_0^i}\right)\right) \exp\left(\sum_{i=1}^N w_0^i \beta_0^i (\bar{x}_0^i - \bar{x}_1^i)\right)$$

The first term of the equation is a pure classification log-Laspeyres index. The second term is explicit within-cell quality adjustment at aggregate level (Koev, 2003). In the second term, vector β_0^i denotes the estimates of the regression model parameters in the elementary level class i . \bar{x}_1^i and \bar{x}_0^i denote the averages of the corresponding sample characteristics of the dwellings in the base and comparison periods respectively. Then within each elementary level class the quality adjustment, due to differences in location and dwelling floor area, can be written as $\beta_0^i (\bar{x}_0^i - \bar{x}_1^i)$.

Index weights are value-share weights from the base year and are derived as follows:

$$(3) \quad w_0^i = \frac{(\bar{A}_0^i \times n_0^i) \times \bar{p}_0^i}{\sum_{i=1}^N (\bar{A}_0^i \times n_0^i) \times \bar{p}_0^i},$$

where \bar{A}_0^i is geometric average of dwelling floor area in elementary level class i in base year and n is the number of dwellings in class i and

p is the geometric average price in class i . Source of the weights varies between Statistics Finland's house price indices. It depends on what we want to portray with our index and what kind of data is available. An expenditure weights is used in price index for new dwellings and stock-weights are used in real estate prices for instance.

3. Classification and regional divisions

3.1 Classification

Location is one of the most important characteristics of the dwelling. The regional classification is formed to be geographically meaningful and as homogenous as possible in respect of price levels. On the other hand, it is found that the regional stratification will not necessarily homogenize the data sufficiently, and some variation will remain within classes. For example, in the housing price statistics the flats within a class may differ relative to their micro location and floor area.

The calculation method in prices of dwellings is based on a well-constructed classification. The classification must meet following three criteria:

- homogeneity,
- adequate number of observations and
- meaningful aggregability.

In other words, we should aim for as detailed a classification as possible, but at the same time each class must contain enough observations to make the use of regression models appropriate. In addition, the micro classification (the lowest classification level) should be formed so that the needed upper level classes can be aggregated from it. The regional stratification is determined to form interpretable geographical entities with relatively similar price-level.

3.2 Regional division

The original idea for forming regional divisions is based on two needs. First, regional classification must meet the needs of the users, i.e. the chosen areas should have some relevance to users. Secondly, regional division should classify the data in a way that each region has similar price levels. If these two needs do not match, it can create problems into the index compilation or into usage of indices. For instance, lower or higher price areas can be situated inside otherwise unified areas

The prices of existing dwellings statistics use diverse area combinations, such as Greater Helsinki Area, satellite municipalities around the Greater Helsinki Area, regions and urban sub-areas. The Greater Helsinki Area comprises Helsinki, Espoo, Vantaa and Kauniainen, which in statistics is included in Espoo. The satellite municipalities are Hyvinkää, Järvenpää, Kerava, Kirkkonummi, Nurmijärvi, Riihimäki, Sipoo, Tuusula and Vihti. Regions are defined according to the decision of the Council of State.

The urban sub-areas are formed of postal code areas using price level and geographical location as the criteria. Appendix 2 shows a graph how Helsinki is divided into four sub-areas.

Due to the relatively low number of transactions, the statistics on the prices of new dwellings in housing companies and the real estate prices are compiled according to a less detailed regional division than the quarterly statistics on the prices of old dwellings. The classification used in the statistics on the prices of new dwellings and the real estate prices also takes into consideration the needs of the Consumer Price Index, hence the regional classification uses the division into major regions. The area categories are Whole country, Greater Helsinki, Rest of Finland (the same as with old dwellings, i.e. Whole country exclusive of Greater Helsinki), Rest of Uusimaa (exclusive of Greater Helsinki) and major regions Southern Finland, Western Finland, Eastern Finland, and Northern Finland. Additionally, the real estate prices include the classification according to the population.

The monthly index of existing dwellings uses same regional division as the prices of new dwellings, with the difference, that six of the biggest cities by population are included into regional division. These cities are Helsinki, Espoo, Vantaa, Tampere, Turku and Oulu. All regional classes are formed in a way that lowest classification level can be aggregated to region and onwards to NUTS2 regions, and finally ending up to highest level, whole country.

4. Usage of administrative data

For old dwellings, the data of the statistics is based on the price information gathered by the Finnish Tax Administration for asset transfer tax calculation purposes. The asset transfer taxation data have been available for statistical purposes from 1987. Previously Statistics Finland received the data yearly and with a delay of seven months, so compiling quarterly index was impossible. In 2001, Statistics Finland and Finnish Tax Administration made a contract that information on prices of dwellings will be provided on a quarterly basis. For new dwellings, statistics are based on data which Statistics Finland receives via a private price monitoring service about transactions in new dwellings made by the largest real estate agents and building contractors.

When a share-type apartment is bought in Finland, the buyer has to pay asset transfer tax 2,0 per cent of the sale price (until 28.2.2013 the asset transfer tax was 1,6 per cent). The Tax Administration maintains a database from all the paid asset transfer taxes and certain properties of the sold apartments. Statistics Finland receives the information about the sold apartments around the fifteenth day of the following month. Not all transactions of old housing company dwellings are included immediately in the statistics, because the purchaser is allowed two months time to pay the asset transfer tax. Many purchasers pay the tax faster than this and in transactions intermediated by real estate agents the tax is paid at the time of transaction. When quarterly statistics are published they cover

approximately two-thirds of all transactions made in the latest quarter. The latest monthly statistics contain around two-fifths of all transactions. Monthly and quarterly data are updated retrospectively in connection with each release so that the final data for the statistical year are published with the data for the first quarter of the year following it. On the average, the revision in monthly statistics on prices of dwellings amounts to 0.3 per cent either way for the whole country. The revision is larger for smaller geographical areas. The average revision in quarterly statistics amounts to 0.2 per cent either way for the whole country.

From the data from Tax Administration Statistics Finland receives following information: price paid by buyer, share in a debt, size of dwelling, name of the housing company, identification number of the housing company, date of purchase, whether the buyer is first-time home buyer (in Finland you are freed for paying asset transfer tax when you buy your first dwelling or house and use it as a permanent place of residence) and whether the buyer is a company. In Finland it is possible that part of the price of the dwelling is debt taken by the housing company. Before 1.3.2012 you did not have to pay taxes from this so called share in a debt of the price.

In addition, Statistics Finland uses different registers on forming housing price indices. The information provided by tax authorities contains only few characteristics of the dwelling, so supplementary variables are needed. The sources for this information are the Tax Administration's Register of Real Estate Property and Statistics Finland's data on the dwelling stock, which is based on the Population Register Centre's Register of Buildings and Dwellings. From these two registers Statistics Finland get important information for classification and important variables for hedonic regression. Registers contain following information: size of dwelling (in calculations we use size information from tax data. If these two data sources give information that differs too much then the observation is deleted.), location of dwelling (municipality, postal code), year of completion of construction work (to determine the age of dwelling), number of rooms and type of dwelling (terraced house, detached house, blocks of flats).

4.1 Challenges of using administrative data

It is definitely a great asset to be able to use so extensive administrative data as we do in Statistic Finland. However, there are few challenges in using registers. Registers are not necessarily completely up-to-date. This can cause some bias in estimation results. For instance if information about floor area is changed greatly, there is a risk that when adjusting prices with hedonic regression, prices can be corrected incorrectly.

A bigger risk is connected to the annual updating of the registers. The Population Register Centre's Register of Buildings and Dwellings is updated once a year. There can be case that some important variable is updated after longer period or update means that content of some variable changes. In this case it is possible that because of changes in

registers it seems that index changes. These sorts of bigger changes in registers should be dated to changes of base year of index.

5. Conclusions

Statistics Finland has produced statistics on housing for quite some time now. Recent years have brought various new opportunities to improve our statistics concerning dwelling prices and prices of detached houses. The monthly index for prices of dwellings has attracted more attention to statistics on housing prices generally. In addition, the range of publications concerning housing prices will be supplemented in the near future with indices of owner-occupied housing. The index for new detached houses will be published for the first time nationally completing the coverage of the Finnish House Price Indices (HPI). There is still some coverage issues on statistics of real estate prices in relation to detached houses on rented plots. There is no administrative data source where this kind of information could be obtained. On the other hand the price development of houses on rented plots is very similar to detached houses portrayed by the statistics of real estate prices.

At the moment, there is extensive development project in progress, which main goal is to develop consistent production process for the statistics on housing (prices of dwellings and real estate prices). During the course of this project regional divisions will be updated so that new postal code areas are taken into account and conceivable changes in valuations of areas have been studied. The hedonic method will be revised as well.

References

Koev, Eugen (2003) Combining Classification and Hedonic Quality Adjustment in Constructing a House Price Index. Helsinki, University of Helsinki (2003)

Official Statistics of Finland (OSF): Prices of dwellings in housing companies [e-publication].
ISSN=2323-8801. Helsinki: Statistics Finland [referred: 21.3.2013].
Access method: http://tilastokeskus.fi/til/ashi/index_en.html.

APPENDIX 2

Graph 1. Helsinki divided into for sub-areas by postal code areas. Blue represents central areas of Helsinki with higher price levels and areas marked with yellow are the areas farthest from city centre or otherwise undervalued and therefore at the lower end of price levels.

