Land & Property Services®

Northern Ireland House Price Index Methodology 2016

Northern Ireland Statistics & Research Agency

Introduction

This paper describes the background to the development of the standardised House Price Index, which was launched by Land & Property Services/Northern Ireland Statistics & Research Agency (LPS/NISRA) in May 2012. It also provides an account of the enhanced methodology following the development, by all producers of official house price statistics, of a single official UK House Price Index. All NI data released from May 2016 onwards will employ this methodology and from June 2016 the NI house price index will be used as the NI component in the single official UK HPI.

Details on the previous methodology can be found at https://www.finance-ni.gov.uk/articles/about-northern-ireland-house-price-index

The main differences to the methodology are:

- prices for all dwellings are predicted using the same characteristics (no longer a composite index)
- building status (new/old) has been added to the characteristics used to predict price
- a price index is now available for new and old properties
- prices are now comparable between years
- the index reference period[1] is Q1 2015 = 100

1 A reference period is a quarter which is used as a benchmark for other periods, to allow calculation of inflation.

1. Background

The decision by the Minister for Finance and Personnel in December 2002 to commence work on a capital value domestic revaluation sparked Valuation & Lands Agency (now part of Land & Property Services) interest in the concept of a House Price Index (HPI).

Exploratory discussions took place with Ulster University, Northern Ireland Housing Executive (NIHE) and Propertynews.com to advise of LPS intention to produce what LPS term as "the definitive house price index" for Northern Ireland. LPS Statisticians researched the technicalities involved, examined the methodology employed in existing HPIs and confirmed their capability to produce ground breaking local analysis.

The National Statistician's Review of House Price Statistics

The National Statistician's review of House Price Statistics which was published in December 2010 recommended the production of a reliable mix-adjusted House Price Index for Northern Ireland based on actual sale prices. Until May 2012, there was no other reliable HPI based on actual sales evidence available within the province.
http://www.statisticsauthority.gov.uk/national-statistician/ns-guidance-and-reports/national-statistician-s-reports/index.html.

The housing market analysis is based on data relating to house sales on the open market in Northern Ireland from 1 January 2005 onwards.

Housing is a major source of individual wealth and changes in housing costs affects the disposable income of households, therefore influencing consumer spending and saving decisions and the overall economy. It is important to have an accurate measure of aggregate housing prices, however it is difficult to develop such a measure as housing is extremely heterogeneous and sales of a particular house are relatively infrequent.

Heterogeneity

No two residential properties are alike, they differ according to various characteristics relating to the physical attributes of the properties themselves or to their locations. Naturally, observed differences in characteristics between two houses will be reflected by differences in price.

Infrequency of Sales

A transaction on any specific house will occur relatively infrequently, therefore it is difficult to gauge the price at which the specific house will sell today.

The main method of calculating sample prices for a given time period is to collect information on properties sold in that period. However, the quality (where quality refers to the combination of all characteristics of a residential property, both physical and locational) of the properties that are sold may differ from one time period to another. For example, sales in one quarter could be disproportionately skewed towards low-quality properties, therefore producing a biased estimate of average price.

It is necessary to 'standardise' to ensure that the varying mix of properties sold in each quarter does not give a false impression of the actual change in prices. In order to calculate the pure property price change over time it is necessary to compare like with like, and this can only be achieved if the 'characteristics-mix' of properties traded is standardised.
There are a variety of methods that can be used to calculate an average house price and track trends in the property market

1. Simple Mean/Median
2. Repeat Sales
3. Hedonic Price Measurement.

Each one deals differently with the issues of heterogeneity and infrequency of sales.

Simple Mean/Median

This method calculates a simple average price (Mean or Median) of all sale prices observed within the time period.

As the distribution of sale prices of properties is positively skewed with some very high value properties sold, a simple mean will be greatly influenced by the few high value properties and will therefore not give a true reflection of the average price of a standardised property. The simple median is not affected by outliers to the same extent as the mean.

This method essentially ignores the problems of heterogeneity and infrequent sales. No attempt is made to ensure that the sample of house sales used is comparable over time. It is hoped that with a very large number of transactions, the mix of properties in the sample will be sufficiently similar over time to give a reasonably accurate gauge of the change in property prices.

It has the advantage that the calculation is simple and simple average prices can be produced to summarise a large number of transactions on a timely basis.

Repeat Sales

The repeat sales methodology analyses only residential properties that have sold more than once. The idea behind this methodology is that the quality of a residential property remains approximately the same over time. If this is the case, then any observed change in sale price must relate to a change in aggregate prices or random "noise". Analysing price changes across a large number of properties filters out the "noise".

The most obvious problem with the repeat sales method is the constant quality requirement for properties that are included in the analysis. In reality, the quality of most properties change over time. In the cases where properties age and become run down, the methodology may underestimate the change in prices, and when owners improve their property, the repeat sales method may overestimate the change in prices.

Another problem, due to infrequency of sales, is that repeat sales methodology is subject to transaction bias. Residential properties that are repeatedly sold may not be a representative sample of properties more generally.

The House Price Index produced by Land Registry England & Wales between 1995 and June 2016 was based on the Repeat Sales Regression technique. It would not have been possible to produce a House Price Index for Northern Ireland using this method as the number of observed repeat transactions is extremely small compared to the total number of property transactions. Repeat house sales in Northern Ireland represent approximately 25% of the total number of house sales available for analysis.

Hedonic Price Measurement

The hedonic method recognises that heterogeneous goods can be described by their characteristics. A residential property can be viewed, mathematically speaking, as a collection of parameters such as (size = 800m^2, Build Class = Private Sector, House Type = Detached, Housing Market Area = Greater Belfast, and so forth). With enough data points, over enough time, a regression model can be used to estimate prices for the characteristics that determine residential property 'quality', where quality refers to the combination of all characteristics of a residential property, both physical and locational. A hypothetical constant-quality residential property, i.e. one with the same attributes over time, can then be "constructed" and priced.

The proposed hedonic modelling approach uses Multiple Regression Analysis (MRA) to produce a multiplicative model that calculates average values for each property type/geographical location. It ensures that although the actual properties sold may be skewed

in favour of one property type or a certain property characteristic, the value of a “standardised” house can still be calculated.

The hedonic methodology does an excellent job of controlling for variations in the quality of residential property.

The method’s main strength is that it can be used to estimate values based on actual choices. Property records are typically reliable as they are collected for financial purposes. The hedonic modelling approach also allows many different characteristics to be considered and included in the analysis if it is found that they significantly contribute to house prices.

The main drawback of this approach is that it is data intensive. Large amounts of data must be gathered and manipulated. The method is relatively complex to implement and interpret, requiring a high degree of statistical expertise and the results depend heavily on model specification. The resource required to produce a hedonic model depends on the availability and accessibility of data and the statistical expertise available.

The method also assumes that people have the opportunity to select the combination of features they prefer, given their income. However, the housing market may be affected by outside influences, like taxes, interest rates, or other factors.

LPS have adopted the hedonic price methodology to produce the Northern Ireland Residential Property House Price Index. Specifically, a Characteristics method of hedonic modelling will be used.

The Characteristics approach to compiling a hedonic price index is defined by Eurostat within the Handbook on Residential Property Price Indices. It involves separate regressions being run for each time period and the index is constructed by making use of the predicted prices based on the regression co-efficients.

The EuroStat handbook on Residential Property Price Indices can be viewed at http://ec.europa.eu/eurostat/web/hicp/methodology/housing-price-statistics/residential-property-handbook

The simple mean, median and mix-adjusted price paid for properties in Northern Ireland over the period January 2005 to current period can be accessed in Table 9 of the Detailed Statistics at https://www.finance-ni.gov.uk/articles/northern-ireland-residential-property-price-index

The data is described fully in section 2. This is followed in Section 3 by an explanation of the methodology and its application.

2. Data

The Commissioner of Valuation for Northern Ireland, situated within Land & Property Services is entitled, under statute, to have access to the Land Transaction returns which are used by solicitors to notify HM Revenue & Customs (HMRC) of the details of taxable transactions.

This should include all domestic property sales as well as non-domestic property sales, land sales and property rentals in Northern Ireland. There are a small number of sales which do not require notification to HMRC, these are:

- transactions where no money changes hands
- property that's left in a will
- transfers of property in a divorce or when a civil partnership is dissolved

The following transaction information is obtained from HMRC:

- Purchase Price
- Property Address
- Effective Date of Sale (Transfer of Funds)
- Vendor Name
- Purchaser Name

In order to "construct" and price a hypothetical constant-quality residential property, it is necessary to have data on property characteristics as well as transaction price. Property characteristics for each house in Northern Ireland are held on the Valuation List which is maintained by the Commissioner of Valuation.

Information is obtained about the following property characteristics:

- Property Address
- Primary Classification of Property (Private/Public sector build)
- Sub-Classification of Property (Detached, Semi-Detached, Terrace, Apartment)
- Era (Age band based on Year Built)
- Number of Rooms (Habitable rooms, Bedrooms, bathrooms)
- Location (Local Government District, Ward, Market Area, Neighbourhood)

The database established by LPS contains all domestic transaction records which could be matched to a corresponding property on the NI Valuation List, thus it contains transaction information and physical characteristic data for the properties sold.

There are a number of transactions which are omitted from the analysis because they are atypical or may distort the series. The following criteria were applied to identify sales to be excluded:

- Sales at a discounted price (including Right to Buy/Co-Ownership etc)
- Sales with a Sale Price less than £20,000
- Sales where Floor space outside identified limits (< 30m^2, > 1000m^2)

3. Methodology

The hedonic method recognises that heterogeneous goods can be described as a function of their characteristics. The starting point for the modelling is the assumption that the price P_i^t of property *i* in period *t*, is a function of a fixed number of characteristics, *j*. The characteristics can be qualitative (such as the type of the property, location of the property etc) and quantitative (such as the size of the property, the number of bedrooms etc). Each property (*i)* sold will differ in price due to differences in its characteristics. The characteristics are measured by "quantities" X_{ij}^t. Therefore, the function for each property sold P_i can be written as:

$$P_i^t = f(X_{i_1}^t, \dots, X_{i_j}^t, \varepsilon_j^t)$$

Where ε_i are a group of unmeasured factors (assumed to be randomly distributed) which are specific to each property but for which data are not available.

It is necessary to determine the appropriate form of the functional relationship between the variables. As the distribution of sale prices is positively skewed, it is necessary to transform the dependent variable to be measured in natural logarithms. The function can be expressed as a parametric model as follows:

$$Ln\ P_i^t = \beta_0^t + \sum_{j=1}^{J} \beta_j^t X_{ji}^t + \varepsilon_i^t$$

Where β_0^t and β_i^t are the intercept term and the parameters to be estimated corresponding to the qualitative and quantitative variables (X_j).

Multiple Regression Analysis (MRA) and the technique of Ordinary Least Squares allows the estimation of the relative contribution of each of the variables, on average, to the measurement of house prices. This relative contribution is the co-efficient of the variables in the regression equation (β), which indicates the relative importance of the variables in explaining the variation of house prices.

The equation uses the values of the quantitative characteristics e.g. size, however the qualitative characteristics can only be represented by categorical or 'dummy' variables which can take the value of one or zero depending upon the presence or absence of a particular attribute.

Definitions of the variables for the set of characteristics and the codes used are listed in Table 1 overleaf.

Table 1: Characteristic Variables

House Characteristic		Code	Definition
Habitable Space		Hab_Space	Gross External area (square metres) as held on NI Valuation List
Privately Built		Class_111	Dummy variable taking the value 1 if a property is built for Private sector. Otherwise 0.
Publicly Built		Class_121	Dummy variable taking the value 1 if a property is built for Public sector. Otherwise 0.
Detached		Subclass_det	4 dummy variables taking the value 1 if a property corresponds to a particular type. Otherwise 0.
Semi-Detached		Subclass_sdt	
Terrace		Subclass_ter	
Apartment		SubClass_apt	
MARKET AREAS	ANTRIM	MA_Antrim	22 dummy variables taking the value of 1 according to the housing market area in which the property is located. Otherwise 0. LPS Housing Market areas were developed for use in the Domestic Capital Valuation Model.
	ARDS PENINSULA	MA_Ards_Peninsula	
	ARMAGH	MA_Armagh	
	BALLYMENA & HINTERLAND	MA_Ballymena	
	BALLYMONEY & HINTERLAND	MA_Ballymoney	
	BANBRIDGE	MA_Banbridge	
	COLERAINE N COAST & GLENS	MA_ColGlen	
	CRAIGAVON	MA_Craigavon	
	DERRY	MA_Derry	
	DOWN	MA_Down	
	DUNGANNON & CLOUGHER VALLEY	MA_Dungannon	
	EAST ANTRIM	MA_East_Antrim	
	FERMANAGH	MA_Fermanagh	
	GREATER BELFAST & CASTLEREAGH	MA_Gtr_Belfast	
	LIMAVADY	MA_Limavady	
	LISBURN	MA_Lisburn	
	MID DOWN	MA_Mid_Down	
	MID ULSTER	MA_Mid_Ulster	
	NEWRY & MOURNES & SOUTH ARMAGH	MA_NewrySA	
	NORTH DOWN	MA_North_Down	
	OMAGH	MA_Omagh	
	STRABANE	MA_Strabane	
Affluent Achievers		Acorn_1	6 dummy variables taking the value 1 according to the geodemographic segmentation of the area in which the property is located. Otherwise 0.
Rising Prosperity		Acorn_2	
Comfortable Communities		Acorn_3	
Financially Stretched		Acorn_4	
Urban Adversity		Acorn_5	
Unassigned		Acorn_6	
New Dwelling		New	Dummy variable taking the value 1 if the sale is the first sale of a new property. Otherwise 0.

The coefficients of the regression function give the implicit price in natural logarithm terms, of the characteristics of the property. The implicit price can be estimated for specific values of the characteristics of a property. To “construct” and price the “standardised” residential property, and allow for the varying mix of characteristics between one time period and another, the average value of each characteristic of residential properties sold in the previous year, are used in the equation.

For quantitative characteristics, the mean value from all sales in the previous year is input and for qualitative characteristics the proportions of that characteristic from all sales in the previous year is used.

These average values can be thought of as a standard ‘representative’ set of weights. The index numbers calculated represent the movement in price paid for a “standardised” residential property possessing the same characteristics as those sold in the previous year. The index numbers themselves are computed by comparing the weighted (i.e. mix-adjusted) prices in each current quarter with the weighted average price in the base period. The base period has been chosen as January – March 2015.

It is not possible to measure all the characteristics that may influence prices. In particular, qualitative factors relating to the condition of properties, attributes of the neighbourhood, amount of traffic, distance to shopping/places of work, etc are not measured. There are also a number of quantitative financial measures such as household income which are not available but are implicitly linked to the ability to purchase a particular property. Consequently it is not possible to explain all of the variation in prices that is observed. However the characteristics used in the equations in this study generally explain around 70% of the variation. During the period 2006 – 2008 the housing market in Northern Ireland behaved in an extraordinary and unprecedented fashion and as a result the amount of variation explained by the model in this period is greatly reduced. It is widely stated that normal market forces were not driving activity during this period and so it would be almost impossible to discover and measure the exact factors which influenced prices during this time.

The hedonic modelling process can be summarised as follows:

(i) From the sales and property characteristics database, remove any sales which are
   a. Sales at a discounted price (e.g. Right to buy)
   b. Sales with a Sale Price less than £20,000
   c. Sales where Floor size outside identified limits (< 30m^2, > 1000m^2)

(ii) To obtain the average value of each characteristic of properties sold in the previous year (i.e. standardised weights, X_j), calculate the proportions of the qualitative characteristics and the means of the quantitative characteristics in the appropriate year;

(iii) Record sale price in natural log form and use the technique of ordinary least squares to estimate the regression co-efficients β_j for the *j* explanatory characteristics, in all time periods from January 2005.

(iv) Calculate the price paid for a "standardised" property in period *t* as:

$$P^t = \exp\left(\frac{\beta_0^t + \sum \beta^t{}_j X_j^t}{\beta_0^0 + \sum \beta_j^0 X_j^0}\right)$$

P_h^t = Price paid for a "standardised" property in period *t*

(v) Calculate a weighted (Laspeyre's type) index (Ix) for the current quarter:

$$Ix = \frac{P^t}{P^0}$$

The index is an annual chain-linked Laspeyres type index. Chain-linking involves "*joining together two indices that overlap in one period by rescaling one of them to make its value equal to that of the other in the same period, thus combining them into single time series*".[1] An annual chain-linked index is preferred as it takes better account of changes in the mix of properties being sold as time progresses.

[1] OECD Glossary of Statistical Terms: http://stats.oecd.org/glossary/detail.asp?ID=5605 (2/06/08)

For 2005, the base set of transactions are those that took place during calendar year 2005, as data for 2004 are not available. The average predicted price is calculated for each of the four quarters of 2005 and the first quarter of 2006, using the base set of transactions, and the coefficients generated for each of the quarterly models. A current year index, based on Q1 2005 = 100, is then calculated for each of the five quarters from the average predicted prices.

For 2006 and later years, the base set of transactions are those that took place in the previous calendar year. The calculation then proceeds as for 2005, except that the current year indices are based on Q1 of the relevant years.

The chain linked index can be calculated because there is an overlap period (Q1) where the previous and current year indices are both known. The current year index is scaled up by the value of the index in the overlap quarter for the previous year, to give a chained index. An example is given below.

Table 3: Example of a Chain Linked Index

	Current year index			Chain linked index based on Q1 2005 = 100		Chain linked index based on 2005 = 100
	2005	2006	2007			
Q1 2005	100			100.0		98.5
Q2 2005	101			101.0		99.5
Q3 2005	102			102.0		100.5
Q4 2005	103			103.0		101.5
Q1 2006	104	100		104.0		102.5
Q2 2006		102		106.1	=102*104/100	104.5
Q3 2006		105		109.2	=105*104/100	107.6
Q4 2006		106		110.2	=106*104/100	108.6
Q1 2007		108	100	112.3	=108*104/100	110.7
Q2 2007			99	111.2	=99*112.32/100	109.6
Q3 2007			98	110.1	=98*112.32/100	108.4
Q4 2007			97	109.0	=97*112.32/100	107.3
Q1 2008			96	107.8	=96*112.32/100	106.2

Reference period

The choice of reference period is to a certain extent academic. It is usual for price indices to be referenced to a calendar year, although referencing on a shorter period is known, eg the retail prices index is based on January 1987. The Northern Ireland House Price Index is referenced to January – March 2015.

In the final column of the example above, the index is referenced to 2005=100. This is a simple re-scaling of the index based on Q1 2005 =100, with the latter being scaled down by 101.5/100, where 101.5 is the average of the four quarters of 2005, based on Q1 2005 =100.

The reference period does not affect calculations of inflation rates, as the relative magnitude of the index values for the two points in time is unaffected by the reference period.