

# **Property Price Statistics for Ireland - Data sources and compilation issues**

Niall O'Hanlon

Central Statistics Office and Centre for Policy Studies, University College, Cork,  
Ireland

[niall.ohanlon@csso.ie](mailto:niall.ohanlon@csso.ie)

## **Draft Version**

Robust transaction based property price measures require suitable source data and these are not always readily available. Compilers may have to rely on sub-optimal data, lacking in respect of coverage, quality, timeliness or detail for mix-adjustment. Alternatively compilers may seek to overcome such weaknesses by combining data from different sources. Where property price measures are compiled using sub-optimal data compilers should, where possible, assess the extent to which they deviate from the target measure. The Central Statistics Office (CSO) currently uses mortgage data to compile the Residential Property Price Index (RPPI) for Ireland and this paper will describe some of the advantages and disadvantages of these data. It will detail how the CSO is assessing the potential for measurement error resulting from the use of mortgage data by compiling an alternative price index using a new dataset it is developing. This dataset combines taxation returns (property prices) and building energy rating assessment data (property characteristics). It will illustrate the challenges presented by this approach. Finally, the paper will describe how the CSO is using a combination of taxation returns and valuation data to explore the potential for commercial property price indices.

Key Words; Real Estate, measurement error, data-matching, mix-adjustment,

## **1. Introduction**

There is a significant body of literature on the compilation of residential property price indices and this has recently been supplemented by the Handbook on Residential Property Prices Indices (RPPIs), Eurostat (2013). Compilation methods and their strengths and weaknesses are well described. Very often the significant challenge for index compilers is securing access to a suitable database that contains accurate and detailed property characteristics and price information.

Section 2 of this paper describes the data used in the compilation of the CSO's RPPI, with a focus on quality and appropriateness. Section 3 will discuss current work on developing an alternative index, using a new dataset compiled from different administrative data sources.

There is an increasing focus on the compilation of commercial property price indices (CPPIs). The CSO has made some progress in gaining access to data sources which could potentially be used for CPPI compilation. These data sources will be discussed under Section 4.

## **2. The Residential Property Price Index**

The Central Statistics Office Ireland (CSO) published the first set of results for its RPPI in May 2011 and it is currently published between three and four weeks after each reference month. The RPPI is a hedonically mix-adjusted index compiled using data on mortgage drawdowns provided by the main mortgage lenders in Ireland

The evaluation framework for house price indices contained in the Eurostat “Owner Occupied Housing Technical Manual” (2012) classifies the full transaction price as the most appropriate price concept, and expert estimates or asking prices as “methods which should not be used” The price of a property should only be observed therefore when it is sold. In order to mix adjust the index (that is, to allow the index to account for the different types of properties sold in each period) descriptors of the characteristics of a property (such as location, building type and size) as well as the transaction price must be available.

At the time of the initial compilation of the RPPI the CSO found the mortgage data to be the only suitable data available to the CSO. Irish mortgage lenders are required under legislation to submit monthly mortgage returns to the Department of Environment, Community and Local Government (DCLG)<sup>1</sup>. These returns contain data on both mortgage approvals (occurring where a formal letter of mortgage offer has issued) and mortgage drawdowns (where the loan has been drawn down – i.e. the transaction completed). The lenders voluntarily transmit these data directly to the CSO for use in the compilation of the RPPI. The data are anonymised – neither the individual borrower nor the property to be purchased is identifiable. Each record contains 67 variables of which; 2 relate to the financial institution, 32 to the borrower(s), 18 to the loan details and status and 15 to details of the property to be mortgaged. Those variables relating to the property to be purchased, which could be used directly in the compilation of a House Price Index are;

- Transaction type – private purchase or with government subsidy
- Agreed purchase price of the property
- County of location (26 administrative regions)
- City indicator (for 4 cities excluding Dublin)
- Postcode, where relevant (for Dublin only)
- Newly built property indicator
- Year of Build
- Dwelling Type (detached, semi-detached, terraced, flat or bungalow)
- Construction type (brick/block, timber frame or pre-cast concrete)
- Floor area
- Plot size (land)
- Number of rooms
- Number of bedrooms
- Use of Property
- Price at drawdown (only 0.04% of records have a different value for this where populated)

Data from Stamp Duty returns (a tax on the transfer of property) was deemed unsuitable as the characteristics descriptors were limited to address and a single size threshold measure. A national property price register was introduced in 2012 but it is compiled from Stamp Duty returns and variables are limited to date and monetary amount of transaction, an indicator for below market value transactions, property address, new or old property and size threshold. These variables in themselves are not sufficient to compile a mix-adjusted index and allow only for the calculation of an index of average prices.

---

<sup>1</sup> Section 13 of the Housing (Miscellaneous Provisions) Act 2002

Until October 2012 the RPPI was the sole transactions based house price index (HPI) for Ireland and it is generally accepted as the preeminent or “national” measure<sup>2</sup>. The data underlying the index has a number of advantages for the compiler. It is timely, received within 12 working days of the end of each reference month allowing for publication with a relatively short lag. The data are provided to the CSO for no charge and provide price and characteristics measures within a single source, allowing the index to be compiled cheaply and efficiently.

However, we must ensure that the index is of the highest possible quality especially as it is high profile and widely used indicator. Furthermore, it will be subject to rigorous quality assessment in the context of both the Harmonised Index of Consumer Prices (HICP) and the Macroeconomic Imbalance Procedure (MIP) Scoreboard, of which national House Price indicators are one of nine measures.

The primary weaknesses in the index lay in the data source rather than the methodologies employed which we consider to be sufficiently robust. The Residential Property Price Index Handbook (Eurostat 2013) sets out 4 criteria for assessing the suitability of data for index compilation; coverage, quality, timeliness and detail. We can assess the suitability of the mortgage data under each of these criteria.

#### Coverage

The data cover the mortgage based element of market only. During the initial development of the index we estimated that mortgage based transactions accounted for approximately 75 per cent of the total residential property market, in volume terms, based on a comparison with Stamp Duty returns between 2005 and 2009. However, the quality of Stamp Duty records was poor, making estimation of market size difficult. Since 2010 Stamp Duty returns are made online and this has significantly improved the quality of data accruing. It is widely accepted that the proportion of cash transactions has increased in recent years as property prices have fallen by some 50 per cent nationally and by almost 60 per cent in Dublin from peak prices recorded in 2007. When estimating the proportion of transactions that are mortgage based it is necessary to exclude non-monetary transfers of property and non-open market sales (such as those between family members) for small monetary amounts prices that would not be financed by mortgages as well as non-household residential properties (such as convents and boarding schools). We estimate that between 2010 and 2012 mortgage based transactions accounted for some 61 per cent of the volume of all in-scope transactions. While this does not in itself imply a bias in the index, the share of cash based transactions is large enough to merit investigation of potential measurement error.

#### Quality

As described by O’Hanlon (2011) the monthly mortgage dataset has a high incidence of erroneous/incomplete returns. Between 2010 and 2012 some 20 per cent of returns were excluded from the index calculation as they were not of sufficiently robust quality. While compilation of the index involves a comprehensive process of data cleaning and editing the high level of excluded transactions remains a concern. Ongoing engagement with mortgage lenders is complicated by the necessity to cooperate with each of them separately and by the fact that data custodians within the lending institutions do not themselves generate the data.

---

<sup>2</sup> Daft.ie introduced the “Residential Property Price Register Index” in October 2012

### Timeliness

There is a lag between when a price is agreed and a transaction is completed and recorded. The Residential Property Price Index Handbook suggests a gap of approximately 9 weeks based on analysis of the house purchasing process in the UK. In Ireland the first formal recording of an agreed price takes place when a letter of mortgage approval is generated by the lending institution. Mortgage approvals and drawdown records detail the month and year in which they take place only so it is not possible to precisely measure the lag between the two events. We analysed drawdown records generated in the latter half of 2012, finding that 12 per cent of drawdowns occurred in the same month as the mortgage approval issued with a further 55 per cent occurring within the following two calendar months. This means that 33 per cent of drawdowns occur more than 2 calendar months after their respective approvals.

However, there is no other source of transactions data that can provide a reduced lag. Many approvals do not result in a final sale and so an approvals based index does not meet Eurostat criteria for a transactions based index. Changing the reference date of drawdown records from their month of drawdown to month of approval would create the problem of causing frequent revisions to the published series and so does not provide a practical solution. One option might be to include only those drawdowns occurring within one or two months of their approval in the index compilation. This would of course significantly reduce the number of transactions available for analysis but these transactions might better reflect the current market.

### Detail

Perhaps the most significant weakness in the mortgage data is the lack of address detail. National location descriptors are limited to administrative county, city indicator and postcodes where relevant. Ireland does not yet have a national postcode system. Postcodes are limited primarily to Dublin City which is divided into 22 postal districts. This limitation on address detail and the absence of a national system of location specific postcodes in Ireland complicates the compilation of a mix-adjusted index and places limitations on the potential explanatory power of a hedonic model and the level of “granularity” in published results.

### Overall Assessment

We consider the mortgage data, despite some advantages, to be sub-optimal. The exclusion of cash based transactions, persistent data quality issues and the lack of detailed address information all risk introducing measurement error into the calculation of residential property price change. However, the potential for measurement error can be accurately gauged only by comparison to other indices based on alternative data sources. Analysis by Daft.ie (Lyons 2012) suggests that between Q1 2010 and Q 3 2012 there was a correlation of in excess of 99.5% between the CSO RPPI and its own asking price and register price based indices. While this is encouraging we believe it is necessary to evaluate the extent to which the RPPI deviates from the target measure, with reference to an alternative index we can create using alternative data sources.

## **3. Developing an Alternative House Price Index**

The CSO has identified two sources of data, which if combined, should provide an excellent dataset for HPI compilation. In line with the European Communities (Energy Performance of Buildings) Regulations 2006, all residential properties that are offered for sale or rent in Ireland must have undergone a Building Energy Rating

(BER) assessment within the previous ten years. BER is the calculated energy use for space and hot water heating, ventilation and lighting based on standard occupancy and is similar to the energy label attached to household electrical appliances. As part of the BER assessment a large number of physical characteristics of a home are recorded covering build type and quality, and size.

There are a number of relatively recent studies of the impact of energy efficiency on residential property values such as Hyland *et al.* (2012) for Ireland, ABS (2008) for Australia, and Brouen and Kok (2011) for the Netherlands. However, there does not appear to be any example of the use of energy efficiency data in the compilation of a residential property price index, despite the detailed information they can provide on the physical characteristics of properties.

The CSO has access to all Irish BER records and is currently matching these to data from Stamp Duty returns at property level. The resultant data will therefore include detail on the price and physical characteristics of all properties sold in Ireland since 2010. In addition, the CSO expects to supplement these data with the Pobal HP Deprivation Index which will provide a qualitative measure of neighbourhood/micro-location and a size of plot indicator provided by the Property registration Authority of Ireland<sup>3</sup>

Data are being matched by address string as Ireland does not have a national system of postcodes. It is important that we make exact matches only. However, non-unique addresses are common in rural Ireland so it will not be possible to achieve a full match – early indicators suggest a match rate of around 60 per cent is likely. While a higher match rate would be desirable, this alternative index will provide us with, at the very least, a good basis for further assessment of the published RPPI. In particular it should allow us to establish the level of measurement error in the national measure. Ultimately, the CSO will decide, based on the analysis, which index will provide the better national measure.

#### **4. Commercial Property Price Indices**

There is an increasing focus on the need for Commercial Property Price Indices (CPPIs).<sup>4</sup> Given the highly heterogeneous nature of commercial property and the scarcity of observable transactions taking place in Ireland, it may not be possible to compile pure transactions based price indices. Instead it is likely that price measures will be appraisal based, or at least supplemented by appraisal or valuation data.

The CSO has identified a number of data sources which could be used in the calculation of CPPIs. Stamp Duty returns for commercial properties are also provided to the CSO on a monthly basis by the Irish taxation authorities. These provide detail on the value and conditions of all sale and lease transactions of commercial property in Ireland.

Additionally, the CSO is cooperating with the Irish Valuation Office (VO) in the development of an automated valuation modelling technique to assist the VO in its valuation of Irish commercial properties. These valuations are based on net annual value (rental income net of associated costs).

The VO compiles detailed information on the physical and usage characteristics of properties and then categorises properties into “bulk classes” prior to their valuation.

---

<sup>3</sup> See Haase, T., Pratchke, J. and Gleeson, J. 2012 for detail on the Pobal HP Deprivation Index

<sup>4</sup> See for example the proceedings of the Joint BIS-ECB-Eurostat-IMF-OCED conference on Commercial property price indicators, Frankfurt (2012)

Valuation levels are then produced for properties based on very detailed valuations of a subset of representative “informer” properties.

A pilot project is currently underway to develop an automated valuation model for a single bulk class of commercial property, Georgian office buildings in South Dublin city. Almost 1,700 relatively heterogeneous buildings in close proximity are being revalued, according to 9 different valuation levels produced with reference to almost 100 informer properties.

As part of the pilot project, the CSO will match valuation office data for this single bulk class to Stamp Duty returns for individual buildings. This should provide us with a database, for a heterogeneous set of properties, containing detailed and precise information on their physical, usage and market (sale, lease and net annual valuation) characteristics. We could, in theory, use this database to compile a mix-adjusted transactions based index. However it is highly unlikely that there will have been a sufficient number of transactions recorded in recent years, given the very low levels of commercial property transactions in Ireland during the current economic downturn. In that case an appraisals based index should be considered.

## References

Australian Bureau of Statistics (2008) “Energy efficiency rating and house price in the ACT” Department of the Environment, Water, Heritage and the Arts.

Brounen, D. and Kok, N. (2011) “On the economics of energy labels in the housing market,” *Journal of Environmental Economics and Management*, Elsevier, vol.62(2), pages 166-169

Eurostat (2012) “Owner Occupied Housing Technical Manual”.

Eurostat (2013) “Handbook on Residential Property Prices Indices (RPPIs)”.

Haase, T., Pratschke, J. and Gleeson, J (2012) “All-Island Deprivation Index Towards the Development of Consistent Deprivation for the Island of Ireland”. Available at <http://trutzhaase.eu/wp/wp-content/uploads/R-2012-All-Island-Deprivation-Index.pdf> Accessed on June 4 2013

Hyland, M., Lyons, R. and Lyons, S. (2012) “The value of domestic building energy efficiency - evidence from Ireland,” *Economics Series Working Papers 614*, University of Oxford, Department of Economics.

Lyons, R. (2012) “The first house price index based on the house price register”. Available at <http://www.ronanlyons.com/2012/10/16/the-first-house-price-index-based-on-the-house-price-register/> Accessed on October 16 2012

O’Hanlon, N. (2011) “Constructing a National House Price Index for Ireland”, *Journal of the Statistical and Social Inquiry Society of Ireland*, Vol. XL.

Proceedings of the Joint BIS-ECB-Eurostat-IMF-OCED conference on Commercial property price indicators, Frankfurt (2012). Available at [http://www.ecb.int/events/conferences/html/20120511\\_cpqi.en.html](http://www.ecb.int/events/conferences/html/20120511_cpqi.en.html) Accessed on June 4 2013