

How to quickly adapt to new policy needs? The experience of the Central Statistics Office, Ireland in developing Housing Price indicators.

Padraig Dalton,

Director General, Central Statistics Office, Ireland

Ken Moore,

Head of Prices Division, Central Statistics Office, Ireland

This paper raises some questions on the ability of CSO as a National Statistical Institute to respond to and meet the needs of our national and international users, all at a time when providers of unofficial statistics are more flexible/timely in filling the data gap with at times sub-optimal data. It highlights the challenges faced in bridging this gap through the use of administrative data sources and details the national experience in improving the quality and timeliness of Irish National House Price Index. The paper also poses some questions for the ESS on what our role is in addressing these challenges.

Context

The demand for house price data in Ireland is significant and growing. The interest in this part of the economy has stemmed from the property bubble that was created during the early to mid-part of the “noughties” and the subsequent crash in the property market which when combined with the global economic downturn also led to significant instability in our financial and banking systems. Data has never been so important to policy makers and the public in general, particularly in the context of the residential property sector as they try to understand the dynamics of the housing market in guiding policy and decision making for the Irish property market.

For many years trends in Irish house prices were measured by the Department of the Environment Housing Statistics Bulletin that presented an average new and second-hand house price for which loans were approved in a particular quarter. However, no adjustment was made for any change in the mix of properties sold or indeed the characteristics of the property. This left a data gap with market needs for a comprehensive index on House Prices not being met. At a time of reducing budgets for the NSI it was private non-official producers who filled this gap with the development of House Price Indices (One index by a financial institution which was based on limited mortgage data followed later by an online property portal where the index was calculated based on asking prices). In the Irish context CSO and official statistics were playing catch up and we could not respond as quickly as desired to our user’s needs as we negotiated access to data sources and built sound methodologies to compile our House Price Index. Indeed it took the then impending legislative requirement to produce Owner Occupied Housing indices for an official measure of House Prices to be introduced in 2011.

However for policy makers this was too late. It is our experience that the first players into the market gains traction. Even though the private producers were producing indices using less robust methodologies and sub-optimal data, users were still referencing this data as no official alternative existed. The challenge we faced was making up the ground on the non-official providers.

Initial CSO Response

Work began on producing House Price data the CSO in 2009 by identifying suitable, data sources, negotiating access to this data and building sound, high quality methodologies to support the compilation of the index. The first set of results for the CSO Residential Property Price Index was published in May, 2011, 5 years after the first publication by non-official producers. The CSO monthly index is a transaction based measure of property prices in Ireland and is compiled using data on mortgage drawdowns supplied by the main mortgage lenders (approx. 85% of all mortgages covered). While the index was in our opinion an improvement on what was previously being provided by non-official providers due to its increased coverage and stronger methodology, there was one key limitation however as cash based transactions were excluded from the index. Mortgage drawdowns accounted for over 80% of residential property transactions in the period 2005-2009, but in the period post 2009 the level of cash based transactions had increased substantially and in particular post 2011. The absence of cash based transactions in the index led to criticism from some commentators who validly questioned the coverage of the index yet the same level of scrutiny was not being applied to the unofficial measures.

Table 1. Number of stamp duty returns by mortgage/cash based market transactions

Year	Stamp Duty Returns for market transactions of residential property	Mortgage drawdown returns	% coverage mortgages	Cash Based Transactions	% Coverage cash
2005	80,685	71,138	88%	9,547	12%
2006	86,086	66,749	78%	19,337	22%
2007	66,090	51,852	78%	14,238	22%
2008	42,682	29,763	70%	12,919	30%
2009	19,719	18,596	94%	1,123	6%
2010	15,175	14,661	97%	514	3%
2011	14,151	8,907	63%	5,244	37%
2012	20,405	11,865	58%	8,540	42%
2013	24,266	11,465	47%	12,801	53%
2014	35,758	17,940	50%	17,818	50%

Improvement

These limitations motivated the CSO to continue to explore additional official data sources to improve the level of coverage and ultimately develop a more robust Residential Property Price Index. New administrative sources were identified as a potential source of cash based transactions but we faced the all too common challenge of having to use a source that was not designed for statistical purposes and trying to link multiple data sources to try and address the problem without a unique identifier. We also faced the challenge of negotiating access with the public bodies to use these sources, even though our national statistical legislation allows for such access. Resource constraints were the primary constraint cited. Despite these challenges our work on a new index, the alternative Residential Property Price Index, began in early 2013 when we identified 5 suitable data sources to support the project:

- Revenue Commissioners (Tax Authority) - Stamp Duty Returns
- Sustainable Energy Ireland - Building Energy ratings
- Census of Population data 2011 – Small area statistics
- GeoDirectory - Location details (post-box granularity)
- Pobal HP Deprivation Index

The sources were primarily linked by direct matching of address strings (see appendix 1 for linking methodology). This proved very challenging and labour intensive because of the lack of a national postcode system, particularly when matching rural addresses where it is common for households to have the same address in local townlands.

Benefits

The new index is nearing completion and will be a huge improvement on the existing model as it will have full market coverage with both cash and mortgage based transactions covered. The variety of data sources will ensure that all relevant property characteristics are built into the model thus ensuring higher quality data and it will allow us to disseminate more property related information (Transaction Values, Volumes, etc.) than is currently being produced. While CSO has invested substantial time in the development of the new

index, the system for processing the administrative data will also allow for improved timeliness in meeting our user's demands. Of course securing the provision of the necessary data sources according to the time line we require to compile this index on a regular basis (monthly) has also taken time. While we had problems initially in securing the data sources, due primarily to resource constraints, we would like to acknowledge the support and cooperation of the Revenue Commissioners (Irish Tax Authorities), Sustainable Energy Ireland, The GeoDirectory and Pobal in making their data available.

Challenges of new index

There are a number challenges associated with the proposed new index that needs to be managed and reviewed including:

1. *Lack of Common Identifier:* At present because of the lack of a common identifier, the model utilises only approximately 75% of market transactions recorded in the administrative data source. The introduction of a national postcode system in July, 2015 will assist in increasing this matching rate once the identifier has "bedded in" but we are still faced with the problem of unmatched historical data.
2. *Delay in filing return with Tax Authorities:* While stamp duty returns are required to be filed with 44 days of execution, we have found that there is often a time lag before the transaction date and the date the stamp duty return is completed. In 2014 an average of 55% of returns related to the current month. This delay raises challenges as to how we deal with the transition for traditional static data to progressive datasets.
3. *Keeping model under constant review:* As new data sources emerge and as the level of administrative data sources improve we will continue to review the new model so as to increase coverage, add additional characteristics and improve the overall quality of the model.

Launch of new index

The model on which the new Residential Property Price Index is based is currently being finalised and it is expected that it will be launched at the beginning of 2016, with the index providing monthly data back to beginning of 2010.

Challenges and strategic considerations for the ESS based on our experiences

Who are the flag-bearers? Can Official statistics be first responders?

As referenced earlier, the early bird catches the worm. In recent times official statistics have been playing catch up with private producers of data who are more agile in identifying data gaps, albeit in a narrow range of domains. With the onset of the “data deluge” we can expect, as discussed many times, to see a significant increase in the number of non-official producers of statistics.

We have to ask ourselves in the first instance if we need to compete with these non-official sources on every front. There are some issues that are just not compatible with official statistics and we need to come to a common view on what these areas are and what criteria we will use to identify them.

Where we should be “in the game”? We need to ask ourselves how we can change our models etc. to pre-empt demands and “get onto the pitch first” as in our experience it is not ideal to be playing “catch-up”.

While our statistics are subject to scrutiny, which is appropriate, can/should we as official compilers question the data produced from non-official producers from a quality perspective? There are problems with this approach as we run the risk of being accused of attempting to take out the competition, irrespective of the validity of our concerns. Is there a danger that we may be perceived as being over critical in our response to competition? How do we address this as a community and what should our approach be into the future? Is there a need for an independent “watchdog”?

How can we improve in anticipating and meeting future policy requirements at a time of reduced resources and increasing demands?

Quality-Timeliness debate

One of the central considerations when responding to new policy needs and demands is maintaining a balance between timeliness and quality. As we compete with non-official data providers, the value of slower, more carefully produced official statistics may become less obvious to some or indeed many users. However we must continue to produce official statistics in a considered, impartial and methodologically sound manner, resisting the temptation to sacrifice quality for time. Nevertheless we must continually strive to improve the timeliness of our data through process efficiency so as to meet the changing needs of our users. How do we strike a balance between managing our user's expectations for more timely, diverse, and customised data while also striving to provide high quality official data?

Branding and Communication of Official Statistics

One of the main challenges for Official Statisticians is that we have a diverse range of users ranging from the expert user to the citizen. Many of the user categories are looking for diverse range of bespoke products to be made available through a diverse range of dissemination channels. In such circumstances we face a significant communications challenge. Should we develop a common or an agreed range of products and dissemination platforms for each user group that we could share across member states? How do we promote our new statistical products effectively so that they reach as wide and diverse an audience as possible? How can we brand official statistics to compete with the non-official sources and inform our users of the merits of using our, perhaps less timely but higher quality, data? In theory we have access to a broader range of sources and perhaps this is an advantage we could exploit by adding value to individual sources and disseminating a more nuanced product.

Data quality of administrative sources

It is critical that we ensure that data quality is not affected in our use of administrative data to meet policy needs. We must address any concerns we have regarding the statistical robustness of these new sources so that the quality of the official statistics we produce

based on administrative sources are not compromised. Very often these administrative sources are not designed for statistical purposes so we must increase the level of engagement with our data providers so that our views are considered when they are designing new data collection/databases. At a national level how do we ensure that a legal framework exists to remove any potential barriers that would limit the use of administrative data while at the same time address any concerns regarding data management and data privacy issues?

Appendix 1: Dataset matching schematic for the Alternative Residential Property Price Index

