# Can list prices accurately capture housing price trends?

# Insights from extreme markets conditions*

**Ronan C. Lyons,** Trinity College Dublin#

**September 2018**

## Abstract

Housing matters but measuring housing prices is not straightforward. Using population-level datasets of listings and mortgage drawdowns for the Irish housing market 2006-2012, and a novel decomposition of the gap between the two, I examine whether listings accurately capture sale price trends, even in volatile market conditions. I find that, accounting for time-to-sell, an index based on listed prices is a very accurate barometer of, and lead indicator for, true transaction prices. With current and historical listings readily available online, the findings are very relevant for both academics and policymakers interested in measuring and understanding living costs and wealth trends.

*Keywords: hedonic price regression; list prices; Irish housing market; price indices.*

* Acknowledgements: The research presented here would not be possible without the extensive work carried out by, and cooperation of, members of the Financial Stability Unit at the Central Bank of Ireland, in particular Tara McIndoe-Calder, in constructing the CBI dataset, and also Distilled SCH for maintaining and providing access to the daft.ie dataset. Sincere thanks are also due to John Muellbauer for his advice, and I would like to thank three anonymous referees and participants at the following annual conferences for helpful comments: AREUEA (National), European Real Estate Society, and the Irish Economic Association.

# Department of Economics, Trinity College Dublin & Spatial Economics Research Centre, LSE

## 1. Introduction

Measuring the price of housing is central to understanding economic conditions. Housing forms the single largest component of consumer expenditure, across income deciles, and for homeowners, their dwelling is typically by far the largest asset in their household balance sheet. However, the accurate measurement of housing prices is not trivial. The lack of fungibility and liquidity in housing means that statistics relating to the market are patchy and rarely comparable across countries (Malpezzi, 2003; Warnock, 2012). For that reason, a number of methods have been employed to measure housing prices, both sale and rental, over time. These range from simple averages and stratification methods to hedonics, the repeat sales method, and the ratio of sale prices to assessed values.

This paper investigates whether hedonic indices based on listed prices accurately reflect underlying transaction prices, even in turbulent market conditions. To do this, it uses two population-level datasets for the Irish housing market over the period 2006-2012, one for online listings and the other for mortgage drawdowns. Using simple theory, it decomposes the gap between the two into four components. Of these, the counteroffer spread reflects the true difference between list prices and the ultimate transaction prices.

The analysis finds that, in particular once time between listing and sale is accounted for, an index based on listed prices is a very accurate barometer of true transaction prices. In doing so, it also reconciles two pieces of conventional wisdom about the housing market, which would appear to conflict in a downturn. Firstly, in a downturn, sale prices are expected to be below list prices, while secondly, list prices are expected to lead sale prices. The analysis here shows that both were true for Ireland 2006-2012.

The findings are relevant for two audiences. One is academic. A number of papers use listed prices, in particular from newspapers, to measure trends in housing costs over time. For example, they are integral to long-run series for the US on CPI (Officer and Williamson, 2018) and the sale price of housing (Shiller 2005). Some existing research, such as Knight et al (1994), highlights the role for list price indices as a leading market indicator but cautions that this role is weakest at market inflection points. More recently, Hayunga and Pace (2017) explore the role of non-market factors – such as a seller age, race and income – in determining list prices, which may suggest limits to the value of list prices as a gauge of market conditions.

The second audience is policymakers and statisticians charged with measuring living standards and costs today. A majority of countries worldwide do not have mix-adjusted residential property price indices (RPPIs) based on comprehensive administrative data. For that reason, alternatives are required and online listings are easy to access and, typically, comprehensive in breadth and depth.

The rest of this paper is structured as follows. The following section outlines the decomposition of housing prices from time of first listing through the sale process to valuation and drawdown. This four-part decomposition provides the basis for the empirical analysis. Section 3 describes the data used and Section 4 outlines the empirical strategy. Section 5 presents the results, while the final section concludes.

**2. Theory**

Consider the following stylised transaction from the housing market. In month 1, a property is listed by its seller. At this point, a list price ($p_L$) is revealed. This is the price (and timing) used in a list price index. In month 4, the property finds a buyer whose offer ($p_M$) is accepted. The property is now "sale agreed" and the bank makes its valuation, upon which the mortgage amount is based. In month 6, the transaction is completed, when the mortgage is drawn down. The price agreed in month 4 is entered into a sale price index in month 6.

The issue in assessing valuation accuracy is immediately apparent, as this typically involves comparing the valuation from month 1 with the transaction in month 6. What is needed instead is a comparison of what the seller's valuation would have been in month 4, when the price was agreed, with the buyer's valuation at the same time. More formally, there are four distinct market processes that take place over three time periods, t=0 (the time of listing), t=υ (when a sale and price are agreed) and t=τ (when the mortgage is drawn down). The price spread between listed and transaction prices, $p_L(0)/p_M(\tau)$, can be decomposed as follows:

1. "selection spread", $p_L(0)/p_{L^*}(0)$: point-in-time list-price difference between all properties that are listed and those subsequently sale-agreed
2. "matching spread", $p_{L^*}(0)/p_{L^*}(\upsilon)$: list-price difference between newly listed properties that go on to be sale-agreed and those sale-agreed
3. "counteroffer spread", $p_{L^*}(\upsilon)/p_M(\upsilon)$: difference between the list price of sale-agreed properties and the transaction price at the time of valuation
4. "drawdown spread", $p_M(\upsilon)/p_M(\tau)$: transaction-price difference between properties being valued and those whose mortgage is being drawn down

## 3. Data

In Ireland's housing market, almost all transactions are facilitated by an estate agent working on behalf of the seller. List prices are not legally privileged in any way and are typically for information only, after being set by the seller and their agent. List price information used here was collated by online accommodation portal, daft.ie. The dataset comprises all properties advertised online between 1 January 2006 and 31 December 2012. Each listing contains the dwelling's listed price, location, type and size, and the date of listing and where relevant dates marked as sale agreed and taken off the site.

Sale price information comes from the Central Bank of Ireland (CBI)'s 2011 Prudential Capital Assessment Review (PCAR) process (Kennedy, 2011). Under PCAR, Irish banks covered by a government guarantee were recapitalized by CBI in return for equity. Such action required loan-level analysis of the mortgage portfolio, which involved detailed information for over 600,000 loans on 475,000 properties being made available to CBI. The analysis here is of loans associated with a housing market transaction, where certain information criteria are met. The dataset includes all observations 2006Q1-2011Q4. Each transaction includes the bank's valuation price, location (county), type, and dates of valuation and mortgage drawdown.

While location is known at a very granular level in the daft.ie dataset, only modest location information is recorded in the CBI dataset. Similarly, no information on a property's size is available for all properties in the CBI dataset. As size (in bedrooms and bathrooms) is recorded in the daft.ie dataset, these are used as a robustness check in the empirical analysis.

## 4. Specification

Hedonic price regressions are used, to ensure a like-for-like comparison over time and to ensure consistency with how housing price indices in Ireland are officially measured (O'Hanlon, 2011). This includes the exclusion of excessively influential observations by using Cook's distance. Similar to the vast bulk of the housing literature, a log-linear model is applied, allowing coefficients to be interpreted, to an approximation, as percentage differentials.

Quarterly categorical variables are included, to capture changes in housing prices over time. To keep the mix of locations constant, county-specific categorical variables are used; thus, the treatment of location is very similar to the official CSO Residential Property Price Index. Regarding property type information, vectors of categorical variables for the five property types were interacted with each of four regions

(Dublin, Leinster, Munster, and Connacht-Ulster), to allow for heterogeneity across markets.

The empirical model is given below, where each vector of $Q$, $X$ and $Y$ omits one category as control and where $s$ refers to the quarter within the year, $t$ to the year, $c$ to county where $C$=25, $r$ to regions 1-5 and $n$ to the property type ($N$=4):

$$ln(hp)_i = \alpha_0 + \sum_{t=2006}^{2012}\sum_{s=1}^{4}\alpha_{ts}Q_i^{ts} + \sum_{c=1}^{C}\beta_c X_i^c + \sum_{n=1}^{N}\sum_{r=1}^{4}\beta_{nr}Y_i^{nr} + \epsilon_i$$

This is applied to the following datasets, to calculate the four spreads outlined:

1. The population of listings, where date refers to the initial date of listing.
2. The subset of listings that subsequently go on to be ``sale agreed'', where date refers to the date of listing.
3. The set of listings that are ``sale agreed'' before being withdrawn, where date refers to the date the property was marked sale agreed.
4. The population of mortgage-backed transactions, where date refers to the date at which the property is valued.
5. The population of mortgage-backed transactions, where date refers to the date of the transaction.

Individual properties are not matched across datasets. Instead, hedonic weighted average prices are used. For each dataset, national average prices for each quarter are calculated using the mix of counties/markets and within each county, the mix of property types, to weight the given average prices for each property type.

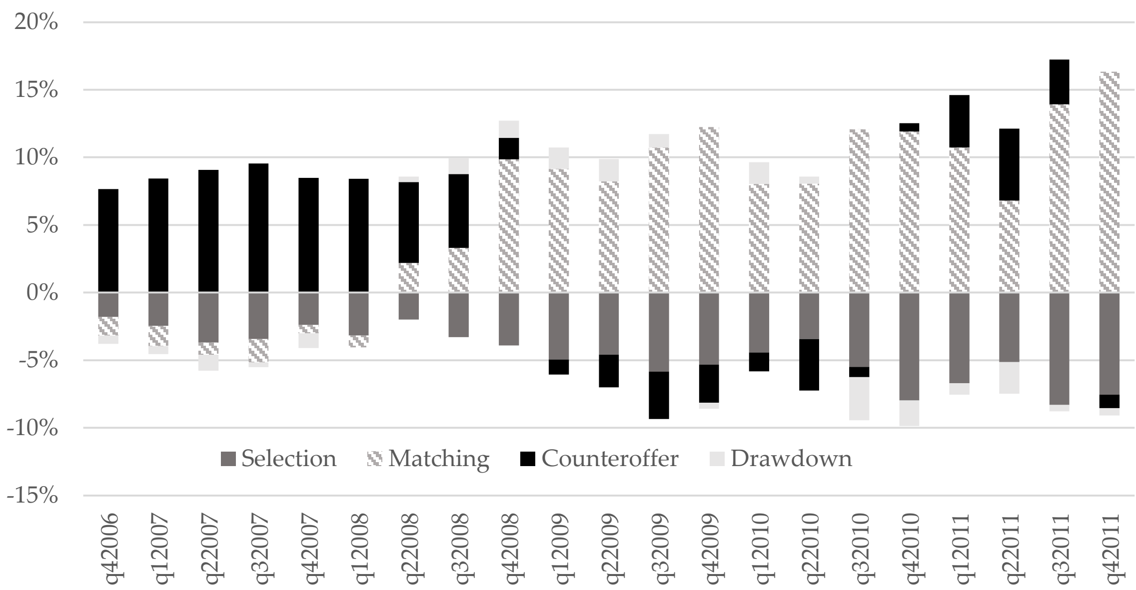

*Figure 1: Decomposition of list price-transaction price gap, 2006-2011*

**5. Results**

Results for each of the four spreads are shown graphically in Figure 1.

1. The drawdown spread is the smallest of the four, averaging just 1 percentage point, reflecting the small and largely constant amount of time between valuation and drawdown.
2. The selection spread is negative throughout, indicating that those properties that subsequently find buyers are listed at systematically cheaper prices. The series is also cyclical (e.g. a larger negative number when prices fell faster in 2009): in a downturn, properties priced more realistically are more likely to sell.
3. The matching spread reflects time-to-sell and how prices change between initial listings and the time a buyer is found. Median and mean time-to-sale-agreed trebled, from two to six months, 2006-2009. Combined with annual price falls of more than 10% 2008-2011, the matching spread is quantitatively the most important of the four. The positive number from 2008 on indicates that properties sale-agreed in a quarter had a higher initial list price than properties listed for the first time in that quarter that ultimately found a buyer.
4. The counteroffer spread, which cross the two datasets, is strongly cyclical. The large positive number in 2007, for example, where the accepted offer was 8% above the list price, even allowing for prices to have changed in the interim. During 2009-2010, the offer was below the list price.

The regression results can be used to calculate a residential property price index (RPPI), adjusting for differences across the two datasets. In particular, for a like-for-like comparison across the two datasets, it is necessary to correct for the selection, matching and drawdown spreads. The year-on-year percentage change in the resulting RPPI is shown in Figure 2 below. As can be seen, the two measures – despite coming from very different datasets – are very highly correlated.

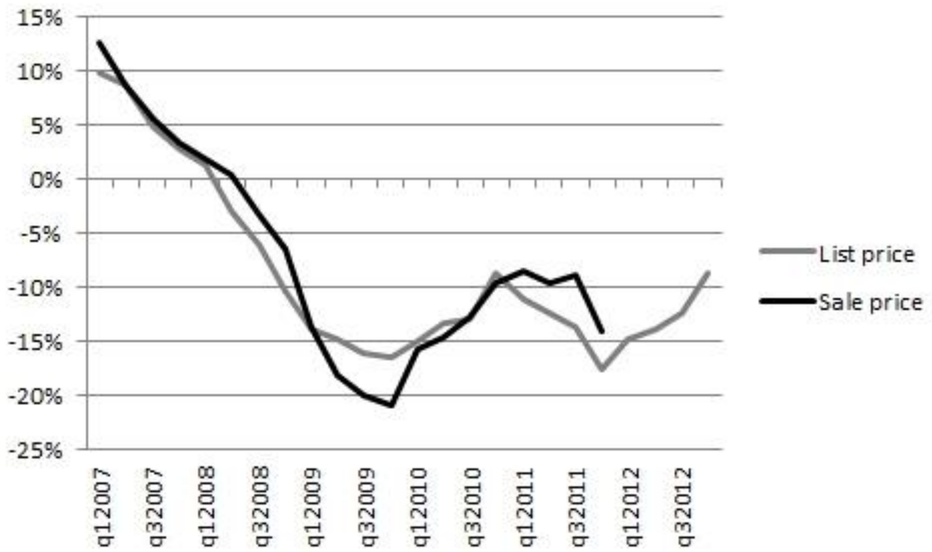


*Figure 2: Year-on-year change in sale and list price indexes, 2007-2012*

The results, over the albeit short span of 21 quarters, suggest that, even in extreme market conditions, list price indices are a good measure of ultimate sale prices. This is confirmed with a simple Granger causality test, a test of whether lagged observations of one variable (here, the quarterly change in list prices) have incremental forecasting power when added to a univariate autoregressive representation of another variable (here, the quarterly change in transaction prices). The $X^2$ test statistic for the change in transaction prices Granger-causing the change in list prices is not statistically significant (3.68, with a p-value of 0.16). However, the test statistic for the reverse, list price changes Granger-causing changes in transaction prices, is strongly statistically significant (20.3, with a p-value of 0.000).

**6. Conclusions**

This paper examined the relationship between listed and transaction prices during Ireland's recent turbulent housing market cycle, using two population-level datasets. After outlining the four constituent elements that make up the relationship between the two – the selection, matching, counteroffer, and drawdown spreads – these were then calculated for Ireland at quarterly frequency, from 2006 to 2011. While the counteroffer spread was large during the end of the price boom, during the falling market, the matching spread was by far the contributor to the gap. Where listings are captured in a timely fashion, therefore, this suggests that they are an adequate substitute where transaction data are not available.

This research has its limits. It is based on two market-level datasets, rather than information at the level of the property, aggregated up. In the CBI dataset, characteristics typically included in hedonic regressions were not available, meaning unobserved heterogeneity may be a factor. There are also issues of external validity: in countries where offers are registered or listed prices enjoy legal privilege, the same results may not hold.

Nonetheless, the results are encouraging for policymakers. Given difficulties in measuring housing price trends, online listings represent a rich potential data source across economies. This paper finds that listed prices are an adequate substitute for transaction prices, even in relatively turbulent market conditions.

The research presented here suggests numerous further avenues for research. These include the investigation of nominal housing price rigidities, as properties stay online at the same price, despite clear evidence that market prices have fallen. Another is a closer analysis of the selection effect – understanding why some properties that are listed find a buyer while others do not. A third avenue is the use of digitized newspapers to revised estimates of living costs and standards over time. Ultimately, handled appropriately, real estate listings can assist policymakers identify conditions in the housing market and establish the need for policy action.

**Declarations of Interest**

None.

**References**

Carillo, P. (2013). To Sell or Not to Sell: Measuring the Heat of the Housing Market. Real Estate Economics, 41(2), 310-346.

Dunse, N., C. Jones and M. White (2010). Valuation accuracy and spatial variations in the efficiency of the property market. Journal of European Real Estate Research, 3(1):24-45.

Hayunga, D. K., & Pace, R. K. (2017). List prices in the US housing market. The Journal of Real Estate Finance and Economics, 55(2), 155-184.

Kennedy, G. and T. McIndoe-Calder (2011). The Irish Mortgage Market: Stylised Facts, Negative Equity and Arrears. Research Technical Paper 12RT11, Central Bank of Ireland.

Knight, J. R., Sirmans, C. F., & Turnbull, G. K. (1994). List price signaling and buyer behavior in the housing market. The Journal of Real Estate Finance and Economics, 9(3), 177-192.

Malpezzi, S. (2003), Hedonic pricing models: A selective and applied review, in T. O'Sullivan and K. Gibb, eds, `Housing Economics and Public Policy: Essays in honour of Duncan Maclennan', Blackwell Science, pp. 67-89.

Miller, N. and M. Sklarz (1986). A Note On Leading Indicators Of Housing Market Price Trends. Journal of Real Estate Research, 1(1):99-109.

Officer, L.H. and Williamson, S.H., 2018. The Annual Consumer Price Index for the United States, 1774-Present, MeasuringWorth.net.

O'Hanlon, N (2011). Constructing a national house price index for Ireland. Journal of the Statistical and Social Inquiry Society of Ireland, 40:167-196.

Shiller, R. (2005), Irrational Exuberance, Princeton University Press.

Warnock, F. E. and V. C. Warnock (2012). Developing Housing Finance Systems. In A. Heath, F. Packer and C. Windsor, editors, Property Markets and Financial Stability. Reserve Bank of Australia and the Bank for International Settlements, 49-67.