



CoreLogic®

Residential Property Index Series

January 2018



This page intentionally left blank

Contents

Overview	4
Index Series	4
Index Maintenance	6
Daily Valuation	6
Selection Rules	6
Base Date	6
Index Rebalancing	6
Eligibility Factors	6
Data Assumptions and Limitations	6
Index Construction	7
Data Cleaning and Standardization	7
Outlier Detection	7
Price Filtering	7
Hedonic Imputation Index	7
Accumulation Index	13
Repeat Sales Index	13
Stratified Median Index	14
Index Governance	15
CoreLogic Research and Analytics Teams	15
Index audit	15
Index Announcements	15
Recalculation Policy	15
Index Dissemination	15
CoreLogic Model Governance	15
References	16

Overview

The CoreLogic Australia Residential Property Index Series is designed to provide a reliable and consistent set of benchmarks for the Australian residential real estate market. The market includes all properties that are defined as residential units (apartments) and houses.

The hedonic imputation approach (based on hedonic regression) to calculating house value indices is the preferred method for compiling a constant quality dwelling value index in Australia. Hedonic regression is a statistical technique that measures the relationship between values of residential real estate and the observed values of its characteristics, for example attributes that encompass its geographic location and property features, such as number of bedrooms and bathrooms etc. [1]. The hedonic imputation index model essentially estimates the value of every property each day, using the coefficients derived using hedonic regression, and describes the daily movements in the value of the residential real estate portfolio.

The hedonic imputation model operated by CoreLogic International utilises comprehensive information on the attributes and characteristics of residential properties to measure quality adjusted changes in property values over time and also to impute the value of all dwellings based on observed characteristics. The hedonic imputation method has the advantage of controlling for compositional bias that may exist in other indices (for example in a median sales index) [2].

In addition to the hedonic index series a set of additional indices are produced, for example stratified median, repeat sales and median types of indices.

In order to establish a summary of best practices for house price index calculation, the Statistical Office of the European Union (Eurostat) coordinated in 2013 the development of the Residential Property Price Indices Handbook [1]. The handbook proposes a comprehensive presentation and analysis of the methodologies to date, and has been widely used as a reference document by financial institutions and legislators [3, 4, 5]. Whilst the Eurostat paper is not prescriptive, CoreLogic Australia have adhered to its guidelines when implementing indices methods appropriate to the Australian market.

Index Series

The CoreLogic index series are produced at a variety of geographical regions within Australia. The regions have been defined using the Australian Statistical Geography Standard (ASGS) 2016 standard [6]. Each index can be subcategorised according to any region and property attribute combination and additional subindices can be generated on request.

A broad suite of indices are produced for listings, sales and rents using different methodologies such as simple median/mean, repeat observation and hedonic imputation. A summary of the indices that are produced are shown in Table 1.

Table 1. Index Series

Index Method	Index Name	Area			
		Sales	Rents	Listings	Properties
Hedonic	Hedonic Index (sales)				●
	Hedonic Index (asking price)				●
	Hedonic Index (rents)				●
	Tiered High/Mid/Low Hedonic Index (sales)				●
	Rental Yield (rent-to-sales)				●
	Rental Yield (rent-to-asking price)				●
	Total Return (sales plus rents)				●
	Deciles Hedonic Index (sales)				●
Stratified Median	Stratified Median Index	●	●	●	
Repeat Observation	Repeat Sales Index	●	●		
	Repeat Rents Index	●	●		
Median	1 Month Median Index	●	●	●	
	3 Month Rolling Median Index	●	●	●	
	6 Month Rolling Median Index	●	●	●	
	12 Month Rolling Median Index	●	●	●	
Mean	1 Month Mean Index	●	●	●	
	3 Month Rolling Mean Index	●	●	●	
	6 Month Rolling Mean Index	●	●	●	
	12 Month Rolling Mean Index	●	●	●	

Index Maintenance

Daily Valuation

Residential property value indices and accumulation indices based on the hedonic imputation method are calculated daily. The property value indices measure the daily movement in imputed values. The accumulation index series measure the daily movement in residential property values with the full reinvestment of rental income (which are also calculated based on the hedonic imputation method).

Selection Rules

All residential properties defined as a house or unit and that are known to exist on the index date will be eligible to be included in the index from the initial month after when the property was added to the CoreLogic database. Properties that no longer meet eligibility criteria will be removed from the index once the record is removed, which may occur if it is known that a property has been demolished.

Base Date

All of CoreLogic's Hedonic Indices have a baseline value of 100 as at 31/12/2009. The starting date for all hedonic indices are nominally set at 31/01/1980, however this will vary greatly for different series based on the availability of data.

Index Rebalancing

The portfolio is rebalanced monthly to account for the addition of or removal of stock from the residential property portfolio over the preceding month. Changes in property attributes (for example a change in the recorded number of bedrooms for a property) are handled daily to ensure consistent quality inter-day.

In cases where a property cannot be valued using the hedonic imputation method, for example if there are no observed sales during the period, it will be removed from the index until a value can be imputed from the arrival of new sales data.

Eligibility Factors

Approximately 98% of Australian properties have been captured in the CoreLogic databases with sales time series data that cover a period of more than 40 years (as at June 2017). There is a high degree of accuracy in recent sales data as 60% of sales are captured before being communicated by the Valuer General within each state.

These features make the data repository the timeliest in Australia and thus allows for an equally timely daily hedonic index. These are the primary sources of the CoreLogic Australia Residential Property Hedonic Index series. All Australian residential properties that are defined as a house or unit are eligible to be included in the index.

The entire known residential property (classified as units and houses) data set is used as the population for the index model for a given index run date. Similarly, all sales and listings are used. Appropriate filters are in place for incomplete records, missing values and extreme values.

A sale record is excluded if the transfer code indicates that the sale was a non-arm's length transaction, for example a transfer due to a death, divorce, bankruptcy etc.. In addition to transfer exclusions, the land use code at the time of sale is used to determine eligibility for inclusion. The land use code is provided by the Valuer General at the time of sale.

Data Assumptions and Limitations

There are a number of limitations with the data available for the production of a hedonic index

- 1. It can be reasonably expected that data quality issues will be present in the property databases as the data is derived from a variety of sources and in different formats.**
- 2. Whilst sophisticated techniques exist to ensure that addresses can be matched across multiple data sources, it is recognised that errors can arise.**
- 3. Latitude and longitude are required for matching properties to a statistical region. Latitude and Longitude are present for approximately 98% of the properties and the coverage is continually improving due to data quality initiatives.**
- 4. There is no point in time property type identifier and therefore there may be instances where it is not possible to tell when a property identifier changed, for example when a house was built on vacant land. This is sometimes possible through the known sale data obtained from the Valuer General but is not always reliable.**
- 5. It is expected that the population of properties and their attributes can change from one day to the next, either due to data quality improvements or due to the arrival of new data on sales and listings.**

Index Construction

The CoreLogic hedonic index is specifically designed to track daily value changes in the residential property market [2]. It is a constant quality value index for the stock of residential housing at a particular time. It measures the wealth associated with the ownership of residential property.

Repeat observation and stratified median methodologies are also available as part of the indices suite. Indices based on simpler aggregation methods such as those based on a median or mean are also calculated.

The steps taken in calculating the hedonic index are:

1. Obtain historical sales over the preceding 360 days;
2. Apply filters and obtain attributes and variables to estimate value function;
3. Estimate value function;
4. Apply value function estimate to population of properties in the valuation portfolio;
5. Identify the composition adjusted portfolio;
6. Calculate and weight the index.

The hedonic regression coefficients are calculated at the Statistical Area 3 level, defined by the Australian Bureau of Statistics, although all other index metrics are produced at all levels of geography.

Data Cleaning and Standardization

It is often the case that historical sales data have missing values for some of the hedonic attributes, especially for earlier periods. In these cases the earliest known non-missing variable is used.

There have been cases where a large number of sales have been recorded where the contract date, sale amount and Statistical Area 1 is the same. These cases typically occur where an apartment block is sold as one transaction and the observed sale price of entire block is attributed to each apartment. This type of data error does not happen often although it is prudent to handle any unusual cases of multi-unit sales that have not been identified correctly. Any instances where 5 or more sales have occurred with the same contract price and date within the same Statistical Area 1 are removed.

Extreme values may be present in some cases for property attributes, sales or rental values. These cases may represent instances of incorrect data entry or invalid sales, particularly if the event related to a non-arm's length

transaction, for example transmission of ownership by death, and the event has not been allocated the correct transfer code. These transfers do not represent market value.

Observed sales that fall outside thresholds determined through expert assessment are also removed (for example less than \$1,000 for and greater than \$100,000,000 for sales).

In the case of sales or listings the following steps are performed:

1. Partition the sales or listings time series by Statistical Area 2 region (or Statistical Area 3 if there are not enough sales at the lower region) [6];
2. Create a date schedule of 30 day time periods ending on the current index calculation date. Note that the first period is likely to be shorter than 30 days and this represents a 'short first' convention;
3. Discard observed sales or listing prices that are identified as extreme values within each 30 day period.

In the case of property attributes the following steps are performed:

1. Partition the properties data set by Statistical Area 3 region [6];
2. Winsorize the data at the 2.5th and 95th percentiles to reduce the effect of extreme outliers.

Every record is processed programmatically to ensure that missing or extreme variables are handled appropriately. In some cases records are removed if there are null values for some variables, however for others the variables are imputed.

Outlier Detection

Outliers are detected dynamically on every index run date in order to exclude them from the regression functions. A daily outlier detection routine was chosen so as to ensure that the index did not move suddenly as a result of a change in the distribution of properties due to less frequent outlier filtering.

The outlier detection method is consistent for transfers, sales listings and rental listings and is determined at the Statistical Area 2 geography level. The outlier detection process is split into two stages (true outliers and price filters) because it was found during testing that outlier

detection and price filtering methods are affected by the outliers themselves. If there are not enough sales at the Statistical Area 3 level then the hedonic regression is carried out at the Statistical Area 4 level (the region above) and the outliers are therefore detected at the Statistical Area 3 level. The outlier detection stage is described in this section and the price filtering stage is described in the following section.

The outlier detection stage is designed to remove records which are obvious errors in the data (true outliers). The time series is split into periods of 30 days. For the start of each 30 day period, all records are selected from the period of 30 days before and after. If there are less than 100 records the window is extended by 30 days on either side until at least 100 records are found. The value of 100 was chosen as the minimum number of observations required in order to correctly identify the required percentiles. Additionally, a constant minimum window length of 30 days was used in order to identify monthly outlier levels as a lot of the indices reporting is based on monthly data. The 2.5th and 97.5th percentiles are identified and any records that are less than 80% of the 2.5th percentile or more than 120% the 97.5th percentile are discarded. Any observations that are less than \$1,000 or greater than \$100,000,000 are also discarded. This step is required so as not to skew the results of the continuous Gaussian kernel smoother in the subsequent price filtering stage.

Price Filtering

The price filtering stage fits the upper and lower price filter bands around the time series. The time series is again split into periods of 30 days. For the start of each period, all records (which now exclude outliers) are selected from the period of 30 days before and after. If there are less than 100 records the window is extended by 30 days on either side until at least 100 records are found. The 2.5th and 97.5th percentiles are identified. It is worth highlighting that for regions with low sales volumes the same sale could be included in the estimation for the price filtering levels for consecutive sales periods. If less than 100 sales are identified in total, no price filter level can be detected for that period.

A continuous Gaussian kernel smoother is generated and fitted to the 2.5th and 97.5th percentiles. In an effort to ensure that a smooth price filter curve was generated and for the Gaussian kernel function to operate within low and high volume transaction environments, a bandwidth of 50 was chosen. Through empirical testing it was found that this was the optimum bandwidth parameter in order to

prevent too many observations from being excluded from the hedonic regression.

Finally, any observations that are outside of the upper and lower bounds specified by the fitted curve are excluded by the regression functions.

Hedonic Imputation Index

There is no uniformity in the application of hedonic regression and it can suffer from omitted variable bias if an important hedonic attribute is missing. It can also be data intensive, although recent advances in distributed computing technologies have helped to solve computationally time consuming and expensive issues. As long as there are sufficient data available for all hedonic attributes, the hedonic imputation method is generally the best technique for constructing a constant quality residential property value index.

The hedonic index requires that the actual prices of all properties are known in order to calculate the movement in values. However, in each period not all properties are sold, therefore their values must be imputed based on sales that were observed for a given Statistical Area, property type and period. The hedonic double imputation method is used to impute estimate the values (or listed values in the case of sales listing and rents indices) of all properties for each period. If the value of property i cannot be imputed, for example if there aren't enough sales in the Statistical Area for that period, the value of property i is imputed from the sample based on the higher level Statistical Area. For example the hedonic data for the Statistical Area 4 region will be used to impute the value of each property instead of the index for the Statistical Area 3 region.

The hedonic regression formula is shown in Equation (1). It gives the least squares estimate of the log price (or asking rental price in the case of the rental index) of a property at a particular time conditional on the information available and controlling for its most statistically significant observable attributes [7]. Listings data are used in the calculation of equivalent rental and accumulation indices, but not in sale price indices. In the case of the rental index, sales are substituted by rental asking prices found in the listings data. The hedonic variables and their transformations are shown in Table 2.

$$(1) \ln P_i^t = C_0^t + \sum_{k=1}^K C_k^t X_{ik}^t + \varepsilon_i^t$$

where:

- P_i^t = sale price (or listing price) of property i ($i=1$ to M) in period t
- X_{ik}^t = transformation of the hedonic variable k for property i in period t
- $C_{0 \text{ to } k}^t$ = numerical coefficients to be estimated
- ε_i^t = (zero mean) residual error term for property i in period t

The calculation of the hedonic index value is shown in Equation (2). The hedonic index is market value weighted and adjusts as new properties enter or exit the index, reflecting when properties were constructed, demolished or added to the residential database. In the case of a daily index, the rebalancing is performed daily to reflect changes in the level of property stock. All properties are revalued for the previous index date. This technique ensures that the population of properties (and their attributes) of properties are consistent between two consecutive index dates and that the effect of time only is captured in the index. This index is non-revisionary.

All properties are revalued for the previous index date. This technique ensures that the population of properties (and their attributes) of properties are consistent between two consecutive index dates and that the effect of time only is captured in the index. This index is non-revisionary.

$$(2) I_{IMP}^t = I_{IMP}^{\hat{t}} \times \frac{\sum_{i=1}^{n^t} \hat{P}_i^t}{\sum_{i=1}^{n^{\hat{t}}} \hat{P}_i^{\hat{t}}}$$

where:

- I_{IMP}^t = sale price (or listing price) imputation index at period t
- $I_{IMP}^{\hat{t}}$ = sale price (or listing price) imputation index at the previous index run date to period t
- \hat{P}_i^t = imputed price (or listing price) of property i at period t
- $\hat{P}_i^{\hat{t}}$ = imputed price (or listing price) of property i at the previous index run date to period t
- $n^{\hat{t}}$ = number properties known for the previous index run date to period t

Table 2. Hedonic variables and their transformations

Hedonic Variable	Description	Transformation
Suburb	The suburb that the property is situated in.	This variable is not transformed.
Land area	Total size of the parcel in square metres.	Transformed to a log value. The land size is reserved for houses, not units. It is winsorized at the 5th and 95th percentile for land area in the property's Statistical Area 3 region.
Street classification	Highway, main road, etc..	This variable is transformed as shown in the first column in Table 3.
Property type	<p>Full category for properties, such as house, unit with additional information such as multi storey, duplex, one storey/lowset etc..</p> <p>The primary land use code which is supplied by the Valuer General's office indicates what usage the government permits for the land (for example, residential, commercial etc.). There are currently around 940 land use types across Australia and 19 residential property types are derived from them.</p>	This variable is not transformed.
Bedrooms	The most recent recorded bathroom count at the time of the recorded sale or listing.	This variable is winsorized at the 5th and 95th percentiles for bedrooms in the property's Statistical Area 3 region and is represented as a factor.
Bathrooms	The most recent recorded bedroom count (inclusive of ensuites) at the time of the recorded sale or listing.	This variable is winsorized at the 5th and 95th percentiles for bathrooms (including ensuites) in the property's Statistical Area 3 region and is represented as a factor.
Bedrooms / Bathrooms	The ratio of the number of bedrooms to bathrooms.	The number of bedrooms divided by the number of bathrooms. This variable is winsorized at the 5th and 95th percentiles.
Car spaces	The most recent recorded car space count at the time of the recorded sale or listing, or the number of garages if the number of car spaces is zero.	This variable is winsorized at the 5th and 95th percentiles for car spaces in the property's Statistical Area 3 region and is represented as a factor.
Pool	A swimming pool is present.	This is either Y or N.
Air conditioning	The property has air conditioning.	This is either Y or N.
View	Property has a scenic view flag.	This is either Y or N.
Recently updated	A flag to indicate that the property has been built or refurbished within the previous 10 years.	This is either Y or N.
Density	A metric which describes the density of the property's mesh block (as defined by the ASGS 2016 standard [6]). It is measured in the number of properties per square kilometre and transformed to a log value.	<p>The density is the total number of properties in the property's mesh block region divided by the total area of the mesh block (in square kilometres), transformed to a log value. The mesh blocks are defined by the Australian Bureau of Statistics [6].</p> <p>(1 is added to the number of properties to safeguard against any errors arising from zero observed properties in a mesh block.)</p>

Table 2. Hedonic variables and their transformations continued

Hedonic Variable	Description	Transformation
Spatial Dependency	A variable that captures the spatial dependency of the property.	The spatial dependency is represented by a 2-D cubic spline function of the X and Y Cartesian coordinates of the property [7]; a thin plate regression spline of the coordinates in the GAM model is used [8].
Time Fraction	Period fraction indicating when the property was sold (or listed in the case of the listing or rent indices) within period t.	<p>The period fraction is shown in the following formula:</p> $F_i^t = \frac{\text{date}(P_i^t) - \text{date}(t_{\text{start}})}{\text{date}(t_{\text{end}}) - \text{date}(t_{\text{start}})}$ <p>(P_i^t) = date of sale (or listing) of property i in period t</p> <p>$\text{date}(t_{\text{start}})$ = start date of period t</p> <p>$\text{date}(t_{\text{end}})$ = end date of period t</p>
Prior Price	Prior sale (or listing) price for the property.	Transformed to a log value. When the prior price is missing then this field is imputed as zero.
Prior Price Days	Number of days since the previous sale.	<p>This variable is not transformed and is calculated as:</p> $\text{days}(P_i^t) = \text{date}(P_i^t) - \text{date_prior}(P_i^t)$ <p>$\text{days}(P_i^t)$ = date of sale (or listing) of property i in period t</p> <p>$\text{date_prior}(P_i^t)$ = previous date of sale (or listing) of property i in period t</p> <p>(An interaction is specified between Prior Price and Prior Price Days).</p>
Prior Price Missing	A dummy variable that indicates whether there is an observed prior sale (or listing) for the property.	This variable is not transformed.

Table 3. Street classifications

Group	Street Classification	State / Territory								
		ACT	NSW	NT	OT	QLD	SA	TAS	VIC	WA
A	VEHICLE TRACK	●	●	●	●	●	●		●	●
A	ACCESS ROAD	●	●	●	●	●	●	●	●	●
A	PROPOSED ROAD		●	●		●	●	●	●	●
B	HIGHWAY	●	●	●	●	●	●	●	●	●
B	RESTRICTED ACCESS ROAD		●	●	●	●	●	●	●	●
C	FREEWAY	●	●			●	●		●	●
C	UNKNOWN	●	●	●	●	●	●	●	●	●
D	CONNECTOR ROAD TUNNEL	●	●							
D	LOCAL ROAD	●	●	●	●	●	●	●	●	●
D	SUB-ARTERIAL ROAD	●	●	●		●	●	●	●	
E	ARTERIAL ROAD	●	●	●		●	●	●	●	●
F	CONNECTOR ROAD	●	●	●		●	●	●	●	●
F	PATH	●	●		●	●	●	●	●	●
F	PRIVATE ROAD	●	●	●	●	●	●	●	●	●
OTHER	Not yet observed based on transfers									

Accumulation Index

The accumulation index assumes that daily rental income is reinvested into the hedonic house value index daily [9]. Transaction, taxation and ownership costs associated with purchasing and owning residential property are not considered (along with individual tax concessions such as negative gearing). A gross rental yield is calculated. It is assumed that all properties are rented; therefore the weekly rental income can be estimated from the hedonic imputation rental model which is based on advertised weekly rents. If the rent of property i cannot be imputed using the hedonic imputation model, for example if there are no listings in period t , the rent of property i is imputed from the hedonic imputation index for the region above. The imputed weekly rents are converted to daily values and reinvested into the daily house value index (assuming an Actual/365 day count convention). In line with the hedonic imputation index it is necessary to ensure that the population of properties (and their attributes) is consistent between consecutive index dates. The accumulation index is shown in Equation (3).

$$(3) \quad I_{TR}^t = I_{TR}^s \times \sum_{i=1}^{N^t} \frac{\hat{P}_i^t + \hat{R}_i^t}{\hat{P}_i^s}$$

where:

I_{TR}^t = sale price total return index at period t

I_{TR}^s = sale price total return index at the previous index run date to period t

\hat{P}_i^t = imputed price (or listing price) of property i at period t

\hat{P}_i^s = imputed price (or listing price) of property i at the previous index run date to period t

\hat{R}_i^t = imputed daily rent of property i at period t

N^t = number properties known for the previous index run date to period t

Repeat Sales Index

In line with the hedonic indices described previously, the purpose of the repeat sales index is to measure the average change in home price in a particular statistical region. The repeat sales method was originally introduced in 1963 by Bailey et al., and there exist a different number of implementations of the methodology, with the S&P / Case-Shiller Home Price Index being the best known. In Australia, Residex (a CoreLogic business) has been producing a repeat sales index since 1991.

The main variable used for index calculation is the price change (sale pair) which occurred between two arms-length transactions of the same property [10]. Thus the repeat sales method aims to capture pure price changes across different time periods, but does not control for a change in quality (which can account for a significant part of the variability between two sales of the same property) [11]. A dummy variable regression model is estimated on the pooled data (sales pairs) across the sample period [1]. So the estimated growth rate in any period is determined not only by the prices of properties sold in that period, but also by the prices of properties for which consecutive transactions occurred on either sides of that period [11]. By design, the method is revisionary and the most recent periods are subject to larger revisions, as the properties which changed hands in a period constitute a relatively small proportion of the total number of sale prices that will eventually affect the index in that period.

If we have a sample of houses which transacted more than once over the sample period $t = 0, 1, \dots, T$ (with $0 \leq s < t \leq T$), the sales pairs are pooled together and the model is estimated with the standard repeat sales Equation (4).

$$(4) \quad \ln(P_n^t/P_n^s) = \sum_{t=0}^T \gamma^t D_n^t + \mu_n^t$$

where:

P_n^t = Price of property n at the end of period t

P_n^s = Price of property n at the end of period s

D_n^t = Dummy variable equals to 1 in the period that the second sale occurs, -1 in the period that the first sale occurs, and 0 otherwise

γ^t = Coefficient for period t

μ_n^t = error term for property n during period t

Under the classical assumptions that the errors have a zero mean and constant variance, the repeat sales formula shown in Equation (4) can be estimated by the OLS regression.

And finally the repeat sales index from period 0 to period t can be obtained by exponentiation of the coefficients γ^t as in Equation (5).

$$(5) \quad P_{RS}^{0t} = \exp(\gamma^t)$$

Stratified Median Index

Median and stratified median price series are also implemented as part of the CoreLogic Australia Residential Property Index Series. Stratification is a process for creating subsets of dwellings which are qualitatively similar. Unique price series are created for these subsets which are then aggregated to estimate strata adjusted price movements in the overall market. The strata definitions used to classify properties into subsets are based on price, geography and interactions of these variables.

The stratified median index for units and houses groups statistical regions by their long-term median transaction price within a higher level statistical region. Sales transfers are grouped by property type, Statistical Area 2 (defined by the Australian Bureau of Statistics [6]) and five year median price within each transfer's Statistical Area 4 statistical region [6]. The Statistical Area 2 regions are ranked and grouped by deciles of equal numbers (the first decile may contain more Statistical Area 2 regions if equal number of Statistical Area 2 regions cannot be determined for each decile). The stratified median for a particular region (Statistical Area 4 level or above) is the weighted mean of the log return of each strata within that region (weighted by sales volumes).

Median and stratified median indices are simple to calculate and provide the price of the middle ranking property observed to sell in a given measurement period. Further, rather than selecting the 50th percentile (i.e., the median) from whatever properties are observed to sell over a particular measurement period, other percentiles such as the 25th or 75th percentile can equally be selected using the same method so as to create other percentile series.

The major shortcoming of median (or any other selected percentile) price series is that they do not represent changes in the value of the residential property market portfolio. Median prices do not track the same properties through time.

In fact, it is most unlikely that any of the properties observed selling in one period will be the same as the properties selling in the next measurement period when examining quarterly or monthly median price series. By way of example, in one period the median priced house observed may happen to be a three bedroom home and in the next period it may be a four bedroom home. Further, in order to provide useful statistics, one needs to aggregate a sufficiently large number of property sales over what can be long periods of time. As such, the publication frequency of median price series is often limited to quarterly, although CoreLogic do offer a monthly median series [12].

The fact that median or other percentile based series cannot be used to track changes in value of a market portfolio does not make them wrong: it is simply that they have different applications than hedonic indices. For example, median price series are useful in answering economic policy questions relating to housing affordability [13].

Index Governance

CoreLogic Research and Analytics Teams

The CoreLogic Research and Analytics teams maintain the index and have full discretion in the application and modification of the index rules and policies regarding index membership. It is also responsible for the oversight of the rules used to rebalance the index.

Index audit

The consultancy KPMG was engaged to audit the hedonic index. The purpose of the collaboration was to determine that the implementation of the index was in line with the methodology CoreLogic Australia had retained for the daily index.

In addition to the audit of the implementation, the model was subject to an academic review. The purpose of such an academic review was twofold:

- i. **Ensure that the new hedonic imputation model was in line with leading research and industry best practices;**
- ii. **Establish that the variables (features) used in the model were appropriate as intended in the methodology.**

Index Announcements

Announcements with respect to changes in the indices models will be communicated periodically in line with continuous improvement initiatives.

Recalculation Policy

There may be occasions where the non-revising indices need to be recalculated. This may occur when material changes to the historic data set have been implemented or where errors are detected in the base level data or outputs. Any errors that arise will be corrected immediately and the changes reflected in the relevant index. Any recalculations will be alerted within the Index Announcements and more broadly if required.

Index Dissemination

The CoreLogic indices data are available from the research and analytics teams and will be distributed via a number of portals. Please contact CoreLogic for further information.

CoreLogic Model Governance

In 2012 CoreLogic established a formal governance program. The objective of this program is to (1) establish model risk management policy and governance program over all CoreLogic models; (2) provide guidance to help CoreLogic business units and mitigate risks arising from customer reliance on computer-based models; and (3) define key model development, testing, validation and monitoring processes and best practices. The program is modelled on published Office of the Comptroller of the Currency bulletins, including "Supervisory Guidance on Model Risk Management" (OCC 2011-12) and "Third-Part Relationships" (OCC 2013-29).

The hedonic imputation index method has been reviewed by the CoreLogic Model Governance committee and approved for client use. This committee maintains oversight on all models approved for client use and require regular reviews of these models to ensure continued compliance.

Please refer to CoreLogic Model Governance White Paper published in May 2015 [14] for further information, which is available on request.

References

- [1] Eurostat, "Handbook on Residential Property Price Indices (RPPIs)," Eurostat Methodologies & Working Papers, Luxembourg, 2013.
- [2] CoreLogic, "CoreLogic RP Data Daily Index," 12 2011. [Online]. Available: <http://corelogic.com.au/research/rpdata-corelogic-home-value-index.html>. [Accessed 07 12 2015].
- [3] F. S. B. International Monetary Fund, "The Financial Crisis and Information Gap," September 2016. [Online]. Available: <https://www.imf.org/external/np/g20/pdf/2016/090216.pdf>.
- [4] B. F. I. Settlements, "Analysis of residential property price developments," February 2017. [Online]. Available: http://www.bis.org/statistics/pp_analysis.htm.
- [5] T. E. P. a. Council, "Regulation on harmonised indices of consumer prices and the house price index," May 2016. [Online]. Available: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0792>.
- [6] Australian Bureau of Statistics, "Australian Statistical Geography Standard (ASGS)," [Online]. Available: [http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+\(ASGS\)](http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+(ASGS)). [Accessed 29 02 2016].
- [7] M. Hardman, "Calculating High Frequency Australian Residential Property Price Indices," 12 2011. [Online]. Available: http://www.corelogic.com.au/resources/pdf/rismark/technical_method_paper.pdf. [Accessed 07 12 2015].
- [8] S. Wood, "Package 'mgcv'," 8 February 2017. [Online]. Available: <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.
- [9] S&P Dow Jones Indices, "Index Mathematics Methodology," S&P Global, 2016.
- [10] "S&P CoreLogic Case-Shiller Home Price Indices," [Online]. Available: <http://us.spindices.com/documents/methodologies/methodology-sp-cs-home-price-indices.pdf>.
- [11] RBA, "Reserve Bank of Australia Bulletin," July 2004. [Online]. Available: <http://www.rba.gov.au/publications/bulletin/2004/jul/pdf/bu-0704-1.pdf>.
- [12] CoreLogic, "CoreLogic Daily Index - Type of Indices," 17 March 2017. [Online]. Available: <https://www.corelogic.com.au/research/types-of-indices.html>. [Accessed 17 March 2017].
- [13] P. o. Australia, "Out of reach? The Australian housing affordability challenge," 8 May 2015. [Online]. Available: http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Economics/Affordable_housing_2013/Report.
- [14] CoreLogic, "Model Governance," Available on Request, 2014.



CoreLogic[®]

corelogic.com.au

For more information contact us 1300 734 318
or email ask@corelogic.com.au

© 2018 CoreLogic, Inc. No unauthorised use or disclosure. All rights reserved.

CORELOGIC and the CoreLogic logo are New Zealand and Australian trademarks of CoreLogic, Inc. and/or its subsidiaries.