MLS® Home Price Index Methodology

The trademarks MLS®, Multiple Listing Service® and the associated logos are owned by the Canadian Real Estate Association (CREA) and identify the quality of services provided by real estate professionals who are members of CREA.

Table of Contents

Introduction 4
Partnership 4
Highlights 4
MLS® HPI 4
Benchmark prices 5
Market segmentation 5
Data inclusions and exclusions 6
MLS® HPI Methodology 6
Data 6
Subareas and market segmentation 7
Modeling approach 8
Model specification 8
Variables 10
Regression 10
Benchmark prices and sub-indexes 12
Aggregate and composite benchmark prices 12
Aggregate and composite indexes 13
Index maintenance 14
Governance 15
HPI contact information 15
Disclaimer 15
Appendix A 16
Appendix B 18
Appendix C 19
Benchmark home definitions 19
1-storey single family homes 19
2-storey single family homes 19
Single family homes 20
1-storey attached single family homes 20
2-storey attached single family homes 20
Attached single family homes 21

1-storey detached single family homes 21
2-storey detached single family homes 21
Detached single family homes 22
Townhouse/row units 22
Attached townhouse 22
Apartment units 22
Attached single family homes 23
Attached category 23
Detached category 23
Semi-detached category 24
Row category 24
Composite 24

Introduction

The MLS® Home Price Index (HPI) is designed to be a reliable, consistent, and timely way of gauging changes in home prices. It's calculated each month and covers the vast majority of housing markets in Canada. The MLS® HPI is also aggregated for the collection of these markets.

The MLS® HPI tracks price levels at a point in time relative to price levels in a base (reference) period for 1- and 2-storey single family homes, townhouse/row units, and apartment units. A composite MLS® HPI is also calculated for the collection of these housing categories in each of the housing markets tracked by the index.

Partnership

The MLS® HPI is generated and published under agreements between the Canadian Real Estate Association (CREA), participating boards and associations, and Altus Group.

The MLS® HPI model was developed by a design team at Altus Group that includes Professor François Des Rosiers, the 2011 recipient of the International Real Estate Society Achievement Award. He's been teaching Urban and Real Estate Management since 1976 within the Faculty of Business Administration of Laval University in Quebec City, Canada.

Representatives from Statistics Canada, Canada Mortgage and Housing Corporation (CMHC), the Bank of Canada, Finance Canada and Central 1 Credit Union have also reviewed and endorsed the MLS® HPI methodology and provided valuable contributions in support of its development.

Highlights

MLS® HPI

The MLS® HPI is available for single family homes (which are further split into 1-storey, and 2-storey single family homes), townhouse/row units, and apartment units. These sub-indices are used to calculate a composite or overall MLS® HPI in each market being tracked. The MLS® HPI for each market is also used to calculate an aggregate MLS® HPI for the collection of metropolitan markets.

MLS® HPI values track relative price levels by comparing price levels at a point in time to price levels in a base (reference) period. Because the base (reference) period has a value of 100, it's possible to quickly infer the extent to which prices have changed relative to the base period. For example, if the base (reference) period for the MLS® HPI is the month of January 2005, and the MLS® HPI value for apartment units in September 2011 is 135.1, this indicates that apartment units in September 2011 were up 35.1% compared to January 2005.

The MLS® HPI is calculated using multivariate regression analysis, a commonly used statistical technique. Using a hybrid modeling approach that merges the Repeat-Sales and Hedonic Price approaches, the MLS® HPI model reflects contributions made by various quantitative and qualitative housing features toward the home price, including:

- number of rooms above the basement level;
- number of bathrooms and half-bathrooms;
- square footage for main living and basement areas;
- whether it has a fireplace and/or finished basement;
- lot size;
- the age of the property;
- parking;
- how the home is heated;
- foundation, flooring, siding, and roofing types;
- whether the property has waterfront or panoramic view;
- whether the property has been sold previously (newly constructed and previously unsold, or repeat sale); and
- proximity to shopping, schools, hospitals, police stations, churches, sports centres, golf courses, parks, and transportation (including the train station, railways, and airports).

Details on MLS® HPI calculations appear in the MLS® HPI Methodology section below.

Benchmark prices

The MLS® HPI model is used to calculate benchmark prices. A "benchmark home" is one whose attributes are typical of homes traded in the area where it's located, one being generated for each supported subarea. Benchmark property descriptions are based on median values for quantitative property attributes (for example, above ground living area in square feet), and the most commonly occurring value (that is, modal value) for qualitative attributes (such as if the basement is unfinished).

Benchmark prices are available for each housing category tracked by the MLS® HPI in each market. Composite and aggregate benchmark prices are also available, representing an aggregation of benchmark categories and metropolitan markets tracked by the Index. This enables benchmark prices and their price changes to be compared across areas, and to the overall market.

Details on aggregate and composite benchmark home price calculations appear in the MLS® HPI Methodology section below.

Market segmentation

To generate consistent indices, markets are divided into areas and subareas for which sales in MLS® HPI categories have similar attributes (homogenous). Subareas have the same geographical boundaries as

those used by real estate boards/associations, which are well known as neighbourhoods. They are used to set MLS® HPI sub-indices, benchmark properties, and benchmark home prices. Each subarea is tested to confirm it's small enough to ensure homogeneity and large enough to ensure there are sufficient sales volumes to model the MLS® HPI throughout housing market cycles.

Details on market segmentation appear in the MLS® HPI Methodology section below.

Data inclusions and exclusions

The MLS® HPI includes transactional data for home sales via MLS® Systems at participating Canadian real estate boards and associations. This data includes sale price and additional information that's added to support the MLS® HPI model, including information from a geographical information system (GIS) to capture additional neighbourhood characteristics (proximity factors) relating to schools, main streets, water, and others.

To maintain data consistency, transactional data are filtered to include records above 0.5% and below 99.5% of cumulative normal distributions for sale price, age, living area, land area, number of rooms, and number of bathrooms. Should a transaction record appear to include internally inconsistent data, it's manually reviewed and amended (scrubbed).

Transactions for which data discrepancies cannot be reconciled without a field visit are excluded. The scrubbing process results in exclusion of less than 5% of transaction records.

Details on data appear in the MLS® HPI Methodology section below.

MLS® HPI Methodology

Data

MLS® HPI data use sold listings (transactional data) based on the date of sale (acceptance of unconditional offer or removal of non-title conditions). Transactional data collected and used in the MLS® HPI must first be reformatted, analysed, sorted, and in some cases, amended; this process is commonly referred to as "scrubbing."

Transactional data are reformatted to include additional fields necessary to support the MLS® HPI. These new fields include calculated, estimated, or inferred attributes from other available information. For example, Floor Area Above Main and Floor Area Main are created in the database and are more useful than a unique Global Living Area field. Detailed living areas by floor are aggregated and compared to the Global Living Area in MLS® HPI regressions. For markets where transactional data includes detailed living area information, it's prioritized over the single Global Living Area in modeling tests.

In keeping with best practices, results are filtered to include records with values above 2.5% and below 97.5% of cumulative normal distributions; other results are treated as outliers and automatically removed. This differs from the 1% exclusion rule in the section *Data inclusions and exclusions* in that this 5% exclusion rule applies to additional fields calculated, estimated, or inferred. To mitigate volatility, a moving five-year sample period is used, since the use of a shorter sample horizon may result in an insufficient number of sales over the period and cause index inaccuracies.

Cook's Distance is used to estimate the influence of an observation when doing least squares regressions and helps detect outliers or identify a subarea where it would be recommended to have more data points. Cook's Distance is also used to discard outliers that may exert a significantly detrimental impact on the MLS® HPI. When the Cook's Distance for an observation is high, the observation is redirected to the scrubbing process for manual validation. To ensure the full potential to extract knowledge from outliers, observations with a high measurement of Cook's Distance are manually reviewed and validated before being removed.

Subareas and market segmentation

Subarea boundaries are initially provided by boards and associations based on their knowledge of their territory and how they want to deliver the index geographically. Subareas are reviewed to ensure they have provided enough current and historical sales activity to support an index. After validation is complete, the board/association gives final approval to draw the final boundaries stored into a database. The final subareas must be small enough to be homogenous but large enough to have enough sales activity. When it comes to model creation, given the high number of available subareas and the possibility of low activity in some of them, a market segmentation using a clustering analysis is implemented to regroup subareas into homogenous groups to create a common model including dummy variables representing every participating subarea to the model cluster. Should significant differences exist between clusters they'll be captured by the coefficients of those binary variables.

Models of emerging communities within sales territories are taken into account from the date the number of transactional data property records achieves a minimum bound (20 over a period of 36 months). Analysis of these sales must also satisfy various diagnostic testing criteria.

In the initial configuration of subareas, new communities are identified and modeled accordingly. The treatment of new communities is also taken into account as part of annual review of the MLS® HPI. As part of the annual review, changes to names and boundaries for market segments in use by the real estate board/association are also taken into account, together with identification of new subareas that come into being.

Should transactional data in subareas fall below the exclusionary threshold (12 over a period of 36 months), those subareas will be excluded from the models upon the next annual review. These subareas will remain deactivated for a minimum period of one year to minimize the volatility of subareas dropping out and rejoining the MLS® HPI.

Modeling approach

The MLS® HPI is based on a hybrid model that merges Repeat-Sales and Hedonic Price approaches. Using multivariate regression analysis, a commonly used statistical technique, the MLS® HPI model reflects the contribution various housing features make toward the home price and includes a dummy variable in the hedonic model specification to distinguish single and repeat sales.

The MLS® HPI is conceptually similar to the Consumer Price Index (CPI), which measures the value of a "basket" of common goods and services. Similarly, the MLS® HPI measures the contribution toward a home's price that each attribute or feature makes as part of a "basket" of housing features.

The approach used to construct the MLS® HPI is superior to the Repeat-Sales approach that has previously gained media attention in Canada and the United States for the following reasons:

- The Repeat-Sales approach omits useful information and sample size is reduced because only homes that have been sold at least twice are used.
- The Repeat-Sales approach may be incapable of reliably tracking home prices for subareas within a market.
- Price indices calculated using the Repeat-Sales approach may be produced with a considerable time lag due to data collection and availability.
- The Repeat-Sales approach assumes qualitative and quantitative attributes of homes remain constant; however, the significance of Canadian home renovation expenditure each year makes this assumption unrealistic.

Model specification

Designing a reliable MLS® HPI requires the regression model be adequately specified. Model misspecification can arise in a number of ways. A rigorous set of statistical tests is used to identify and resolve potential problems arising from model misspecification.

In a linear regression, one of the main assumptions is there are no remaining multicollinearity[1] phenomena. Stepwise regression is employed to remove excessive multicollinearity by selecting only those explanatory variables that contribute significantly to explaining price variations. As a diagnostic test, variance inflation factors (VIF) are used to highlight and remove variables with a high degree of multicollinearity.

The Akaike Information Criterion (AIC) allows comparing models that differ with regard to their functional form, variable specification, or both; as such, it can aid in model selection based on how close values predicted by the model are to the real data. The AIC is used to test which of the linear or semi-log functional forms provides the best fit. To accommodate nonlinearities, the living area, lot size and age of properties are transformed into non-linear forms. Results of the AIC suggest the use of the semi-log form over the linear form.

Additionally, the Ramsey RESET Test is used to determine if some form of non-linear transformation is required within the model specification (without indicating how to amend the specification).

The RESET test estimates an auxiliary regression using the estimated Yi from the original regression:

$$\hat{Y}_i = \hat{\beta} \quad + \hat{\beta}_1 X_{1i} + \ldots + \hat{\beta}_{ni} X_{ni} + \gamma \hat{Y}_i^2 + \delta \hat{Y}_i^3 + \omega \hat{Y}_i^4 \qquad \text{i=1, 2, … N}$$

Where $\hat{Y}_i$ is raised to the second, third and fourth powers and re-inserted in the initial hedonic equation as additional independent variables. The test then compares the original and the auxiliary regressions via F statistic test. The hedonic function is shown to be non-linear if at least one of these $\hat{Y}_i^n$ added terms emerge as statistically significant.

In cases where the equation fails the Ramsey RESET test, the AIC confirms the functional form. That the age of a property cannot be non-linearly transformed may explain the failure at the third and fourth degree for markets where property age is modelled as a binary variable denoting age range.

Demand for 1- and 2-storey single family homes is significantly different, as reflected in their sales prices. Accordingly, they are modelled separately, with sufficient sales activity to maintain separate and statistically valid categories. Details on how the single family home sub-index is calculated appear in the Aggregates and Composites section below.

Single family homes include both attached and detached structures since analysis shows the behaviour of a combined 'detached/attached' index tracks congruently with a 'detached' index (configured by extracting sales records of 'attached' homes while maintaining compliance with test criteria).

[1] Multicollinearity is a statistical phenomenon in which two or more exogenous variables in a multiple regression model are highly correlated.

‘Detached’ and combined ‘detached/attached’ indices are monitored to ensure the congruency of their respective trends supports a combined index.

New communities within a sales territory are considered as part of an annual review of the MLS® HPI. When accumulation of transactional data results in adjustments to market segmentation of a sales territory, MLS® HPI models are re-run to take account of geographic revisions while ensuring homogeneity is maintained for each grouping.

Variables

All available information and data that describes land, buildings and location amenities is considered in the MLS® HPI model specification. Socio-demographic attributes (namely, education level and average income) also contribute to the determination of subareas and their grouping for sampling purposes. Additionally, a GIS is used to capture additional neighbourhood characteristics (proximity factors) such as those relating to schools, main streets, water, and other factors.

Data is validated before being used in the modeling process. Each variable is analyzed (minimum, maximum, distribution, form), resulting measurements are stored, and key variables are monitored on an ongoing basis.

Variables for living area, land area, property characteristics and dummy time variables are included in the model, and key variables (such as living area and land area) are transformed to fit the data (a list of variables used in the MLS® HPI appears in Appendix A).

To maintain homogeneity, outlier records are filtered out so that data include records above 0.5% and below 99.5% of the cumulative normal distributions of sale price, age, living area, land area, number of rooms, and number of bathrooms.

A random control sample is then created using 10% of the remaining transactional data records to run through the same process as the initial model to validate variables.

Regression

Using a stepwise regression procedure, independent variables are successively forced into the model and then removed from the hedonic equation based on their statistical significance via a student’s t-test. Variables kept in the model are fully analyzed and interpreted. It’s ensured that time dummy variables are included, and key variables satisfy logical rules (for example, the number of rooms cannot be negative). Also, variables with data occurrence greater than 5% within the database are included in

the model specification[2], and a random control sample is confirmed as valid. Afterwards, Cook's Distance is applied to identify and discard outliers that may exert a deleterious impact on hedonic coefficient estimates.

Diagnostic statistical tests (as below) are then performed to determine if assumptions underlying ordinary least square (OLS) regression modelling are violated. If test results indicate these assumptions are violated, or the model is mis-specified (for example, the omission of an important variable) or subject to a functional form design flaw, then the results and the sample are analysed, and corrective actions are taken at the data, scrubber, market delineation or functional level as appropriate.

One of the main assumptions for the OLS regression method is errors have the same variance throughout the sample. If true, the model is said to be homoscedastic. If not, the data is said to be heteroskedastic.

As long as the assumption of homoskedasticity is not violated, OLS is considered to be the best linear unbiased estimator (BLUE). When the assumption is violated, OLS regression estimates are deemed inefficient, and OLS is not the best regression method.

One or a combination of additional measures and strategies are used to detect heteroskedasticity, and when required, correct for it (White test, Weighted Least Squares regression technique, and additional data transformations for example).

Moran's Index Test, often referred to as the Moran's I test, is used to measure the degree of spatial dependence among residuals. A model can be considered adequate if its residuals are not related in space. If they are, this is considered to be evidence of spatial autocorrelation. Like heteroskedasticity, the presence of spatial autocorrelation violates the OLS method assumption that residuals are independent from each other.

The presence of spatial autocorrelation is typically marked by unstable regression parameters and unreliable inference tests. Several solutions are available to correct for the presence of spatial autocorrelation, including Casetti's expansion method, spatial autoregressive techniques, and peer effect models.

[2] For example, if the number of properties that have parking is greater than 5% but the parameter 'Parking' is not in the model, the parameter is forced into the model.

Benchmark prices and sub-indexes

Following the generation of regression equations, each subarea's benchmark property attributes are inserted in the equation to calculate their respective benchmark prices. Each property type supported in the said subarea is attributed a benchmark property, ignoring other property types. These individual benchmark prices are calculated each month.

Monthly sub-indexes are calculated using the benchmark price of the reference period (January 2005) as the denominator and prices in other periods as numerators to calculate corresponding monthly sub-indexes.

Aggregate and composite benchmark prices

The MLS® HPI calculates a set of price indexes and sub-indexes, benchmark prices and relative benchmark prices.

Aggregate benchmark prices for areas in the base (reference) period of January 2005 are based on the weighted[3] contribution of sales activity in constituent subareas for each benchmark category (1-storey single family, 2-storey single family, townhouse/row unit, and apartment unit), whereby the MLS® HPI model calculates benchmark home prices for each subarea using applicable benchmark home attributes in each subarea:

$$P = \sum_{j} W_{i,j} * P_{i,j}$$

Where 'P' represents the MLS® HPI category benchmark price, 'i' represents benchmark category, 'j' represents constituent subarea, and 'w' represents the proportion of benchmark category activity for the subarea. Please refer to Appendix B for a more detailed example of this calculation.

Several levels of aggregation exist and vary from board to board, depending on their specified requirements. The next level may be municipality and then area, or just area, with the level above this being the sales territory of the real estate board, followed by province and then the aggregate of participating boards in Canada.

Aggregate benchmark prices following the base (reference) period are adjusted by applying corresponding MLS® HPI inflation/deflation. As a result, growth rates for aggregate benchmark prices are equal to those for their corresponding MLS® HPI. However, since published benchmark prices are rounded to the nearest $100, growth rates calculated using rounded aggregate benchmark data may

[3] Weights based on proportional values for a three-year period of sales activity.

differ slightly from those for corresponding aggregate MLS® HPI indices. For that reason, it's recommended users rely on percentage changes based on aggregate MLS® HPI indices.

Composite benchmark prices in each area in the base (reference) period of January 2005 are based on the weighted contribution of sales activity in constituent subareas per benchmark housing category, with the single-family benchmark price analogously calculated based on weighted contributions of just 1- and 2-storey sales activity:

$$\boldsymbol{P} = \sum_{\boldsymbol{i}} \sum_{\boldsymbol{j}} \boldsymbol{W_{i,j}} * \boldsymbol{P_{i,j}}$$

Where 'P' represents the MLS® HPI composite benchmark price, 'i' represents benchmark category, 'j' represents constituent subarea, and 'w' represents the benchmark category's proportion of total sales activity for the subareas.

Similarly, metropolitan composite benchmark prices in the base (reference) period of January 2005 are based on the weighted contribution of sales activity in constituent subareas per benchmark housing category.

Composite prices following the base (reference) period are adjusted by applying corresponding MLS® HPI inflation/deflation. As a result, growth rates for composite benchmark prices are equal to those for their corresponding MLS® HPI. However, since published benchmark prices are rounded to the nearest $100, growth rates calculated using rounded composite benchmark data may differ slightly from those for corresponding composite MLS® HPI indices. For that reason, it's recommended users rely on percentage changes based on composite MLS® HPI indices.

Aggregate and composite indexes

Since benchmarks are the only item in the consumer basket, Paasche and Laspeyres index[4] values don't change while calculating sub-indexes per benchmark category, since quantities cancel themselves out:

$$\Delta P_L = \frac{\sum p_{j,i} q_{0,i}}{\sum p_{0,i} q_{0,i}}$$

[4] Research and Innovative Technology Administration, *Use of the Chained Fisher Ideal Index to produce the Aggregated Transportation Services Index,* Economics and Finance, http://www.bts.gov/programs/economics_and_finance/transportation_services_index/methodology/pdf/methodology_chained_fisher_ideal_index.pdf

$$\Delta P_p = \frac{\sum p_{j,i} q_{j,i}}{\sum p_{0,i} q_{j,i}}$$

Where 'PL' and 'PP' represents Laspeyres and Paasche Index respectively, 'i' represents benchmark category, 'j' represents the subject period, and '0' represents the reference period.

Since the Fisher index 'PF' is obtained by taking the geometric mean of Laspeyres and Paasche, quantities also cancel themselves out. It's important to understand this statement is only true on sub-indexes per type.

$$\Delta P_F = \sqrt{\Delta P_L * \Delta P_P}$$

Unlike the Laspeyres Index which overestimates the variation in prices, and the Paasche Index which underestimates it, the Fischer Price Index is more reliable in the estimation of actual price change over time.

The Chained Fisher Index is used to calculate aggregate and composite indexes to conserve the direct month-to-month link that keeps recent sale prices non-obsolete. Accordingly, the results of calculations used in deriving the metropolitan composite and aggregate composite MLS® HPIs also serve in its calculation:

$$\Delta P_{FC} = \sqrt{\frac{\sum p_{0,i} q_{1,i}}{\sum p_{0,i} q_{0,i}} * \frac{\sum p_{1,i} q_{1,i}}{\sum p_{1,i} q_{0,i}}} * \sqrt{\frac{\sum p_{1,i} q_{2,i}}{\sum p_{1,i} q_{1,i}} * \frac{\sum p_{2,i} q_{2,i}}{\sum p_{2,i} q_{1,i}}} * \ldots * \sqrt{\frac{\sum p_{j-1,i} q_{j,i}}{\sum p_{j-1,i} q_{j-1,i}} * \frac{\sum p_{j,i} q_{j,i}}{\sum p_{j,i} q_{j-1,i}}}$$

Where 'PFC' represents the HPI Chained Fisher Index, 'i' represents benchmark category, 'j' represents the subject period, and 'j-1' represents the reference period.

Index maintenance

The MLS® HPI is reviewed annually. The annual review includes e-testing model specifications with a view to potentially strengthening the model. If reviews result in models being re-specified, historical data are revised. Data exclusions are also reviewed and updated, as necessary. With the updated methodology as of June 1, 2022, a moving five-year period of sales is used to calculate the benchmark attributes. For example, five years of sales activity up to and including December 2021 were used to calculate the new benchmarks launched on June 1, 2022. Similarly, for all subsequent annual reviews launched on each June 1, the five years of sales activity up to and including December of the preceding year will be used moving forward. With benchmark attributes updated every year during the annual review, all benchmark prices and historical time series data will also be updated to reflect the most recent changes.

Governance

Policy decisions on the use and circulation of MLS® HPI information is at the purview of the MLS® HPI Steering Group, which consists of representatives of the Canadian Real Estate Association (CREA) and real estate boards and associations taking part in the MLS® HPI.

HPI contact information

For technical enquiries or enquires about index operations or business development regarding the MLS® HPI, please contact admin.stats@crea.ca.

For media enquiries regarding the MLS® HPI, please contact Pierre Leduc, CREA's Media Relations Officer, pleduc@crea.ca.

Disclaimer

Data and reports regarding the MLS® HPI are provided for informational purposes only. Products and information regarding the MLS® HPI are not intended for investment purposes. The information and any statistical data regarding the MLS® HPI are obtained from sources CREA believes to be reliable but does not represent they are accurate or complete. All estimates and opinions expressed by CREA regarding the MLS® HPI constitute judgments as of the date of this report and are subject to change without notice.

© 2022 The Canadian Real Estate Association. All rights reserved. Unauthorized use, distribution, duplication, or disclosure without the prior written permission of CREA is prohibited by law and may result in prosecution.

Appendix A

Variables used in the model

Parking access
Tangible or intangible benefits that increase attractiveness or value
Property is serviced by municipal aqueduct
Property is near a shopping mall
Method of heating
Source of energy for heating
Flooring type
Foundation material
Property is equipped with a fireplace
Garage has two parking spaces
Property is equipped with geothermal energy
Property building is semi-detached
Land size in square feet
Property siding material
Property has undergone major renovations
Only a part of the property is renovated
Property is equipped with a roughed-in fireplace
Basement is finished
Parking lot has a shelter or carport
Garage is located below main floor
Roofing material
Property has a crawlspace
Property has a view of water
Property has a panoramic view
Number of bathrooms
Number of half-bathrooms
Property is in proximity to an elementary school or a high school
Hydro line neighbours the property's lot
Property has a view of power lines
Property is in proximity to a train station
Property is in proximity to a church
Property is in proximity to an airport

Property is in proximity to a boulevard
Property is adjacent to a boulevard
Property is in proximity to a sports centre
Property is in proximity to a railroad
Property is in proximity to a hospital
Property is in proximity to a police station
Property is in proximity to a prison
Property is in proximity to a golf course
Property is in proximity to a park
Property is adjacent to a park
Basement living area in square feet
Time dummy variable for each month and year
Number of rooms above basement level
Main living area in square feet
Number of rooms at basement level
Age of property

Appendix B

Benchmark modelling for a 2-storey single-family home

Benchmark attribute	Benchmark value	Model coefficient	Benchmark model value (Benchmark value x model coefficient)
Finished basement? (yes=1 no=0)	1	0.09916109	0.09916109
Hard siding? (yes=1 no=0)	1	0.01822980	0.01822980
Living area (sq. ft.)	1,600	0.00012801	0.20481942
Number of bathrooms	2	0.03689373	0.07378746
Number of fireplaces	1	0.06672047	0.06672047
Number of garage spaces	1	0.01064111	0.01064111
Number of half bathrooms	1	0.07630139	0.07630139
Number of rooms	7	0.00766833	0.05367831
Other factors not modelled (default=1)	1	12.99547005	12.99547005
Square root of lot size (sq. ft.)	57.01	0.00720978	0.41102934

Table 1 – sample benchmark attributes and weightings in a model

To determine the benchmark price, all the benchmark model values are summed. Then the formula $P = e^x$ is used, where P is the benchmark price, e is a mathematical constant (Euler's number), and x is the sum of the benchmark model values.

In this example, the benchmark price is calculated to be $1,214,500 (rounded to the nearest hundred dollars) for a 2-storey single-family home in this neighbourhood in a particular month.

Note: all figures used in the table are for demonstration purposes only and do not represent any specific subarea or period. Every property type and subarea combination in the MLS® HPI will have its own unique set of values, as well as coefficients that vary from one month to the next. Additionally, while only a sample of benchmark attributes have been presented in the table, the set of attributes will also change depending on the type of home and subarea.

Appendix C

Benchmark home definitions

Benchmark homes are representative of standardized homes for specific subareas. Their physical characteristics remain fixed over time. Benchmark property attributes are formulated for each subarea for benchmark housing categories that have a significant presence in a subarea.

The following describes general characteristics for each benchmark housing category and for selected composites of categories:

1-storey single family homes

A property with one floor above ground. This type of property is characterized by the bedrooms, kitchen and dining rooms being on the same floor; the utility room and laundry room are generally located below ground. Special attention is made to raised bungalows, where the basement is partially above ground and where the room distribution provides criteria for its assignment to the appropriate benchmark housing category. This includes property styles submitted by participating real estate boards labeled as: back split, bi-level, bungalow, hillside bungalow, hillside split, 2-storey split and 3 level split. This type of property does not differentiate between attached and detached homes.

2-storey single family homes

A property with two, or more, above ground floors. This type of property is characterized by the distribution of bedrooms on the upper floor and a kitchen, living room and other day-to-day rooms on the main floor. This category includes property styles submitted by participating boards labeled as: 4 level split, 5 level split, 1-and-a-half storey, 2-storey, 2-and-a-half storey,

and 3-storey. This type of property does not differentiate between attached and detached homes.

Single family homes

Benchmarks and indices for “single family homes” are generated as a composite of 1- and 2-storey single family homes described above.

1-storey attached single family homes

A property with one floor above ground sharing at least one wall (or part of a wall) with another home. In addition to sharing a wall, this type of single-family home is characterized by the bedrooms, kitchen and dining rooms being on the same floor; the utility room and laundry room are generally located below ground. Special attention is made to raised bungalows, where the basement is partially above ground and where the room distribution provides criteria for its assignment to the appropriate benchmark housing category. This includes property styles submitted by participating real estate boards labeled as: back split, bi-level, bungalow, hillside bungalow, hillside split, 2-storey split and 3 level split. This type of attached home is distinct from units such as a townhouses, apartments or condominiums which typically share more than one wall with another home being within multi-family dwellings and where the ownership and maintenance costs of exterior walls, land and entrances may be shared.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

2-storey attached single family homes

This category includes properties with two or more above ground floors sharing at least one wall (or part of a wall) with another home. In addition to sharing a wall, this type of home houses one family and is characterized by the distribution of bedrooms on the upper floor and a kitchen, living room and other day-to-day rooms on the main floor. This category includes property styles submitted by participating boards labeled as: 4 level split, 5 level split, 1-and-a-half storey, 2-storey, 2-and-a-half storey, and 3-storey. This type of attached home is distinct from units such as a townhouses, apartments or condominiums which typically share more

than one wall with another home being within multi-family dwellings and where the ownership and maintenance costs of exterior walls, land and entrances may be shared.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Attached single family homes

Benchmarks and indices for the category “attached single family homes” are calculated as a composite of 1- and 2-storey attached single family homes described above. The term “attached” describes a home that is part of a larger multi-family building.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

1-storey detached single family homes

A “1-storey detached single family home” is built with one above ground floor on land that exceeds the footprint of the building on each of its sides. Ownership of the surrounding land (typically less than a 40,000 square foot yard), entrance and associated lesser structures is an integral and inseparable part of the home.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

2-storey detached single family homes

A “2-storey detached single family home” is built with two or more above ground floors on land that exceeds the footprint of the building on each of its sides. Ownership of the surrounding land (typically less than a 40,000 square foot yard), entrance and associated lesser structures is an integral and inseparable part of the home.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Detached single family homes

Benchmarks and indices for “Detached single family homes” are generated as a composite of 1- and 2-storey detached single family homes. As described above, these homes are independent structures that are typically built on land that exceeds the footprint of the building on each of its sides. Ownership of the surrounding land (typically less than a 40,000 square foot yard), entrance and associated lesser structures is an integral and inseparable part of the home.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Townhouse/row units

Townhouses have configurations which lay between apartment units and freehold non-strata buildings, such as bungalows and 2-storey houses. Owners typically pay co-ownership fees for maintenance and enjoy exclusive access to a part of the lot. This category includes property styles submitted by participating boards labeled as any of the submitted styles, with a note that the property is a townhouse.

Attached townhouse

Attached townhouses have configurations which lay between apartment units and freehold non-strata buildings, such as bungalows and 2-storey houses. Owners typically pay co-ownership fees for maintenance and enjoy exclusive access to a part of the lot. This category includes property styles submitted by participating boards labeled as any of the submitted styles, with a note that the property is a townhouse. Additionally, the townhouse must be either a semi-detached or an attached property.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Apartment units

Apartment units are characterized by being part of a multi-unit building. Occupants of apartment units may or may not have direct access to the lot from their units. There are also no parts of the lot whereby access is reserved for only one of the co-owners or apartment

occupants. Several type of homes within this category had originally been duplexes and triplexes and are now treated as condominiums or apartments according to a formal declaration of co-property.

This category includes property styles submitted by participating boards labeled as: single level apartment, multi-level apartment, loft, penthouse, duplex, triplex, and studio suite.

Attached single family homes

Benchmarks and indices for the category “Attached single family homes” are calculated as a composite of 1- and 2-storey attached single family homes described above. The term “attached” describes a home that is part of a larger multi-family building.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Attached category

Benchmarks and indices for the category “Attached category” are calculated as a composite of 1-, 2-storey attached single family homes and attached townhouses as described above. The term “attached” describes a home that is part of a larger multi-family building.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Detached category

Benchmarks and indices for “detached category” are generated as a composite of 1- and 2-storey detached single family homes. As described above, these homes are independent structures that are typically built on land that exceeds the footprint of the building on each of its sides.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Semi-detached category

Benchmarks and indices for the category “Semi-detached category” are calculated as a composite of 1-, 2-storey semi-detached single-family homes and semi-detached townhouses as described above. The term “semi-detached” describes a home that is part of a two units multi-family building.

Benchmarks and indices for this particular property type are generated under arrangement only with certain boards.

Row category

Benchmarks and indices for the category “Row category” are calculated as a composite of 1-, 2-storey row single family homes and row townhouses as described above. The term “row” describes a home that is part of a multi-family building with at least three units physically linked.

Benchmarks and Indices for this particular property type are generated under arrangement only with certain boards.

Composite

Benchmarks and indices within the “composite” category represent all homes used in models, including 1- and 2-storey single family homes, townhouses, and apartments. Accordingly, the composite index includes both attached and detached homes within 1- and 2-storey single family homes.

NOTE: MLS® HPI data published by CREA for 1- and 2-storey single family homes (and for the single-family composite) include detached and attached properties. By comparison, MLS® HPI data for single family homes published by some real estate boards in their local market areas include only detached properties. Differences in property coverage may result in MLS® System data variances for single family homes as published by CREA and the aforementioned real estate boards/associations.