Description of Price Index for New One-Family Houses Sold

Introduction

The Price Indexes of New One-Family Houses Sold Including Value of Lot are a group of price indexes designed to show the inflation in new houses built for sale. Contractor-built houses, owner-built houses, and houses built for rent are not included.

The data used for computing these indexes are obtained from the U.S. Census Bureau's Survey of Construction. The survey collects information on the physical characteristics and prices of new one-family houses. This is done through monthly interviews with the builders or owners of a national sample of new houses. The size of the sample is currently about 20,000 observations per year.

Price index design – Laspeyres type indexes

The Price Indexes of New One-Family Houses Sold Including Value of Lot are Laspeyres type indexes. The basic form of a Laspeyres type price index is:

$$\frac{\sum_i (p_{ti} * q_{0i})}{\sum_i (p_{0i} * q_{0i})}$$

Where the p_{0i}'s and p_{ti}'s are the prices in the base and current period, respectively, and the q_{0i}'s are the quantities in the base period. This represents the ratio of the current cost of the quantity of goods purchased in the base year to the cost in base year prices of the same quantity of goods. Notice that the denominator is the price of the average base period house. To compute this index the p's must be derived from a regression model since we only collect a total house and land price.

Experience has shown that regression estimation of the price in the following multiplicative model is superior to estimation for the above additive model:

$$\frac{e^{\sum_i (d_{ti} * q_{0i})}}{e^{\sum_i (d_{0i} * q_{0i})}}$$

Where the d_{0i}'s and d_{ti}'s are price factors and the q_{oi}'s are the quantities in the base period.

It is necessary to obtain the *d's* and *q's* for various commodities to compute this index. We estimate from survey data the *q*'s (quantities) for commodities that we refer to as house characteristics. The *d's* – referred to as price factors – for the characteristics cannot be collected so we estimate them from a regression model. For this reason the sums in the above equation can be thought of as regressed values taken from a regression model. The regression models used to estimate the price factors are described in more detail below.

Forming strata also improves the regression estimates. We use five strata: four defined as the detached units in the four Census regions and the fifth as all attached units. The Laspeyres price index at the national level is then:

$$\frac{e^{\sum\limits_{h} w_h * \sum\limits_{i} (d_{ti} * q_{0i})}}{e^{\sum\limits_{h} w_h \sum\limits_{i} (d_{0i} * q_{0i})}}$$

Where the first sum in each term is over the five strata and w_h is the proportion of the houses sold (activity) in the stratum and the other terms are computed by stratum.

Regional Indexes

Price Indexes of New One-Family Houses Sold Including Value of Lot for the four Census regions are shown annually. These Laspeyres type indexes are computed as follows:

$$\frac{e^{w_r * \sum\limits_{i} (d_{tir} * q_{0ir}) + w_a * \sum\limits_{i} (d_{tia} * q_{0ia})}}{e^{w_r * \sum\limits_{i} (d_{0ir} * q_{0ir}) + w_a * \sum\limits_{i} (d_{0ia} * q_{0ia})}}$$

Where w_r and w_a are the proportion of units sold (activity) in the stratum of detached units and attached units for the region, respectively.

Regression Models

The Survey of Construction collects the price including land and various characteristics for a sample of new houses sold. Available characteristics are evaluated to determine a suitable set to use for index calculations. Since consumer preferences change over time, any set of characteristics is useful for a limited number of years. At irregular intervals, one set of characteristics is replaced by a new set. When a new set of characteristics is introduced into the index computations, the index for prior periods may be recomputed or the index may be re-based. See the section of documentation titled History of One-Family Index Methodology for a complete discussion of re-basing procedures.

To compute an index, price factors ($d's$) must be derived for the characteristics used in the index. Price factors are derived from the price and characteristic data by multiple linear regression.

Presently there are five separate regression models used to calculate the price indexes: one model for detached units in each census region and one model for all attached units. Each of these models is designed to measure the contributions of important physical and geographic characteristics to the prices of new houses.

The characteristics used in each model except for floor area are divided into categories. For example, each house is classified by whether it has less than three bedrooms, three bedrooms, or more than three bedrooms; whether it has no garage or a carport, a one or two car garage, or a garage for three or more cars; etc. Each category is treated qualitatively in that a value of "1" indicates that the house has that characteristic and "0" indicates that the house does not have it. One category from each of the qualitative characteristics must be omitted to avoid an over determined system. The price and floor area are treated quantitatively – the logarithms of the actual values are used directly in the model.

Tables A-1 (detached houses) and A-2 (attached houses) show by stratum the characteristics used in the regression models and the categories associated with them for the Price Indexes of New One-Family Houses Sold Including Value of Lot. These tables also show the base year average quantities ($q's$) used in the calculation of the indexes.

A resistant regression procedure is used that diminishes the impact of unusual observations by down-weighting the data for those observations.

Since the regression does not include all of the characteristics that explain price variability and because the characteristics are interdependent, the estimated regression coefficients should not be regarded as estimates of the true price factors. The estimated regression coefficients are not shown.

Stratum Weights

The base year for the indexes is currently 1996. The weights (levels of activity in 1996) used to combine the five stratum indexes to form the Price Indexes of New One-Family Houses Sold Including Value of Lot are shown in Table A-3.