A New Hedonic Rent Index Methodology for Canada

Motivation

As part of modernization and quality improvement efforts at Statistics Canada, a new characteristics-based hedonic model has been developed to replace the previous matched model for the housing rent price index. Starting with the January 2019 CPI, the rent index is calculated and published using this new methodology.

Data and Weights

Data come from the monthly Labour Force Survey (LFS) of households from 2008 to 2018. Although this survey collects data on the price and characteristics of rental units, its sampling is done to satisfy labour market characteristics. The LFS strata weights, corresponding to postal code forward sorting areas (FSAs), are therefore used in aggregating regional price estimations to better reflect the Canadian rental market structure. Prior to running the model, certain codes are run to flag any problems in the data, e.g. missing values, so that subject matter experts will have the ability to modify / impute some of the observations. Further outlier detections are done within the model.

Model

Following the review of different approaches, as well as model specification tests, a characteristics-based hedonic model is chosen as follows:

$$y^{*}=\beta_0 + \beta_1\, services + \beta_2\, age + \beta_3\, bedrooms + \beta_4\, dwellingtype + \beta_5\, FSA + \epsilon$$

where y^{*} is the log of observed rent. Variables in the model are categorical dummy variables for different types of services (cable, heating and gas, laundry, etc.), dwelling (apartment, duplex, single detached, etc.), age, and number of bedrooms. FSAs are geographic dummy variables for the forward sorting area, which can capture regional fixed effects, such as median income and population, which are not available on a monthly basis. Standard tests, such as multicollinearity, are done and necessary adjustments are applied. In the first step the model is run, and using a Cook's distance test, outliers are flagged for subject-matter reviewing and modifications. The second step runs the model to produce price estimates.

Index construction / Aggregation

Using estimates from the model, we can construct metropolitan area (CMA), regional, and national rent indices. We first calculate the quantity $x_j^{char,cma}$ of each characteristic by CMA in period j by aggregating over all LFS strata in the CMA:

$$\boldsymbol{x_j^{char,cma}} = \sum_{n=1}^{N} \boldsymbol{w_n^{strat,j}}\, \boldsymbol{x_{n,j}^{char,cma,strat}}$$

where $w_n^{strat,i}$ are stratum weights and $x_{n,i}^{char,cma,strat}$ is the quantity of characteristic $char$ in LFS strata n. We can then use these quantities to calculate the expenditure on each characteristic by CMA using period i prices and period j quantities:

$$\boldsymbol{pq_{i,j}^{cma}} = \boldsymbol{e^{\beta_i x_j^{cma}}}\, \boldsymbol{e^{\frac{\hat{\sigma}_i}{2}}}$$

Where β_i is the vector of price coefficient estimates from period i. An adjustment factor $e^{\frac{\hat{\sigma}_i}{2}}$ is included. These expenditure estimates can be used to produce CMA and CPI strata level indexes using LFS weights. Superlative indexes can also be calculated. The final price index I_p^{CS} for prices going from the previous period (period 0) to current period (period 1) in a given CPI strata CS is:

$$\boldsymbol{I_p^{CS}} = \left(\frac{\boldsymbol{pq_{1,0}^{CS}}}{\boldsymbol{pq_{o,0}^{CS}}}\,\frac{\boldsymbol{pq_{1,1}^{CS}}}{\boldsymbol{pq_{0,1}^{CS}}}\right)^{\frac{1}{2}}$$

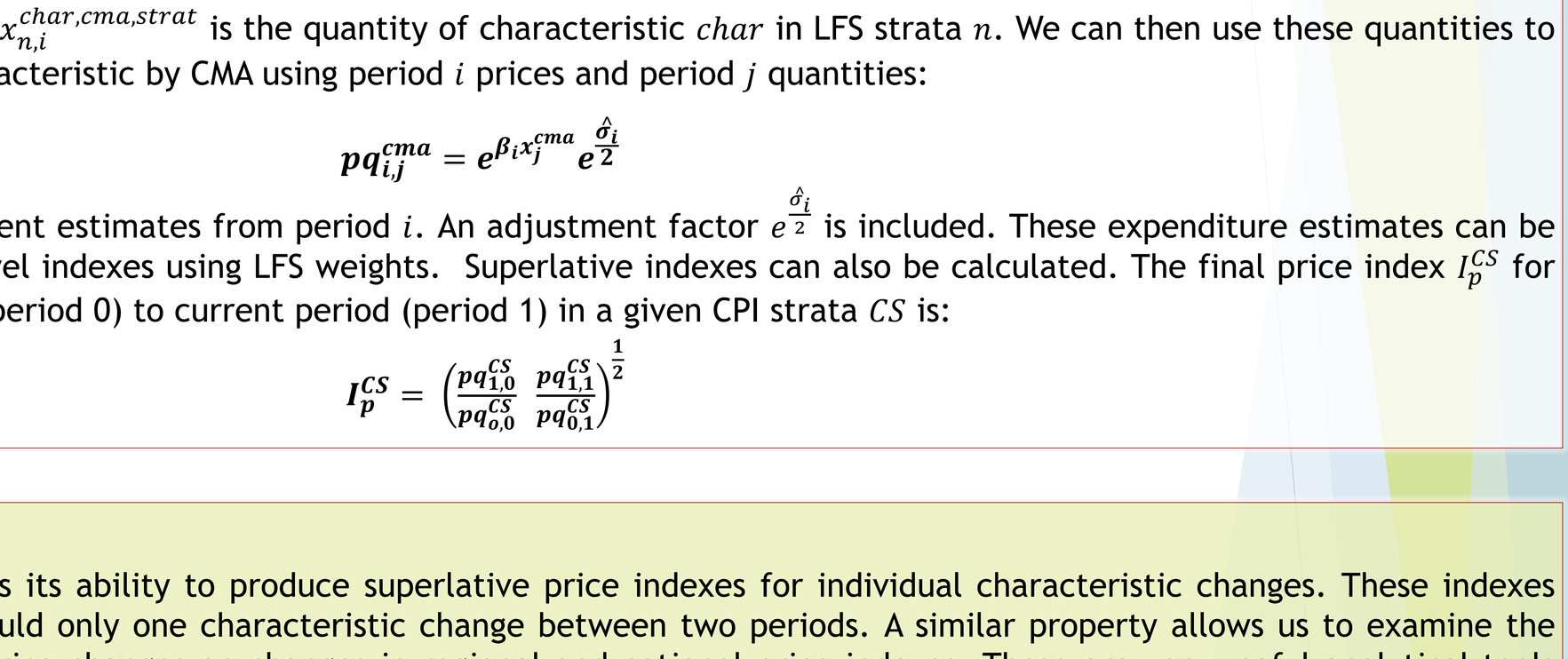
Index / model properties

An interesting property of the model is its ability to produce superlative price indexes for individual characteristic changes. These indexes reflect the change in rental prices should only one characteristic change between two periods. A similar property allows us to examine the effects of specific metropolitan area price changes on changes in regional and national price indexes. These are very useful analytical tools when producing monthly indices.

Model Results

Model estimation results show that the hedonic model is more reflective of the rental market trends over time. The model yields coefficients that are mostly statistically significant at a 5% significance level, and stable over time with the expected signs. The results of location characteristics (FSA) also support the claim that conurbation matters, with denser urban FSAs yielding higher and more statistically significant estimators. The average goodness of fit (adjusted R^2) between periods is 0.65, suggesting that the model provides a satisfactory description of price movements.

Rent Price Index, Sept 2018 – Mar 2019



Source: Statistics Canada; Table 1810000401

Statistics Canada / Statistique Canada