Demographic Forecasting Model 2024

A Demographic Forecasting Model for Canadian Provinces and Sub-Provinces.

Prepared by Kojo Laryea

Overview

The Demographic model is an econometric forecasting model for the geographical regions of Canadian provinces and their sub-provincial levels. It is designed to be a full cohort-level population model, providing provincial and sub-provincial-level forecasts by single-age cohorts and a breakdown of the sources of projected population growth: births, deaths, domestic net migration and international net migration.

The input data for the model and the forecasts for exogenous variables and other assumptions are prepared and stored in a collection of Excel files. The model is coded in Python and contained in one large Python script. When successfully run, the model produces an Excel output file of the generated forecasts.

Running The Model

The first step in running the model is to update the four Excel files in the "Population Model 2024" folder. The files are titled Population_Cohort," "Population_Data," "Population_Assumptions," and "Population Exogenous." The cohort and data files contain the historical data used in the regressions, while the last two include the assumptions used in generating the forecasts.

Once these four files have been updated, the next step is to ensure Python is appropriately installed on your computer. For the Python script to be run, you must have Python installed (version 3.8 or later) and several additional Python packages. The non-built-in packages that need to be installed are numpy, pandas, xlsxwriter, openpyxl, statsmodels, and linearmodels. These packages can be installed using pip (e.g. "pip install numpy") in a command prompt window. When the Python script is run successfully, it generates an output Excel file containing the forecast results in a file that is suffixed with the current date and time.

The third step is to define your forecast window. This will be done in the Python script in lines 24 and 24. The forecast window is currently set to (2024,2029) but can be adjusted depending on how many years of forecasts are needed.

The final step is to run the Python script. The model results will be stored in the "Population Model 2024" folder.

Model Structure

The model forecasts net migration first. The net migration figure for a given forecast year represents the migration that occurred between July 1 of the previous year and July 1 of the forecast year (except for the city of Calgary population, which primarily relies on civic census numbers dated between April of the previous year and April of the forecast year).

Net migration at the provincial level is split into international and net migration.

1. Provincial International Net Migration= f(Lagged Province/Canada Population Share, Lagged Change In Province Total Employment, Federal Immigration Policy Dummy)

2. Provincial Domestic Net Migration= f(Lagged Provincial Domestic Net Migration, Province/Canada Unemployment Rate Ratio, Province Residual Deviation, Lagged Change In Province Total Employment)

3. Economic Region (ER) International Net Migration= f(Province International Net Migration, ER/Province unemployment rate ratio

4. Economic Region (ER) Domestic Net Migration= f(Province Domestic Net Migration, ER/Province Unemployment Rate Ratio)

5. CMA International Net Migration= f(Economic Region International Net Migration)

6. CMA Domestic Net Migration= f(Economic Region Domestic Net Migration)

7. City Total Net migration= f(Economic Region/Canada Unemployment Rate Ratio, Lagged Province/Canada Population Share, Lagged Change In Economic Region Total Employment, City House Price, CMA Inflation Rate, Federal Immigration Policy Dummy)

International net migration is typically less volatile and primarily related to federal policy, though there is some correlation with economic conditions. The current nature of federal immigration policy is having such a strong effect on net migration that a dummy variable has been included to account for this at the provincial and city levels. For the ER, the effects of immigration are already accounted for since it relies on provincial international net migration. The same applies at the CMA level.

Domestic net migration is very volatile and sensitive to relative economic conditions across Canada, as measured by the relative unemployment rate. Actual job growth in Alberta also reflects newcomers arriving for specific jobs. The lagged variables account for this phenomenon: domestic migration has a certain "momentum," so its lagged value has a positive effect. The residual deviation is a variable from the Statistics Canada components of population growth data added when a new census reveals errors in the total population growth between censuses. Since the individual components of population growth are not revised in this case, the discrepancy is solved by adding a "residual deviation." Much of the errors likely come from domestic migration. The residual deviation is always zero for future years since errors in the population growth estimates are expected to have a mean of zero. The labour market assumptions for forecasting can be retrieved from the overall macroeconomic model or a consensus survey.

The rest of the model for all geographical regions takes the following steps:

-Take the previous year's age/sex cohorts and shift them all one year older.

-Calculate total births by applying fertility rate assumptions[1] to age cohorts

[1]. The cohort-level fertility rate assumptions represent the total births per woman expected in an age cohort during a given year. The process for generating these assumptions is based on death probabilities provided by Statistics Canada. These can be provided on an annual basis separately.

-Distribute births by sex according to the assumed (historical) birth-sex ratio and add them to the total population. These births become the new 0-year-old cohorts.

-Calculate deaths by multiplying the cohorts (including the new 0-year-old cohorts) by the mortality rate[2] assumptions. Sum up the total deaths and subtract associated deaths from each population cohort.

-Add net migration to the population, distributing international and domestic migration according to their respective assumed (historical) shares.

The model is set up to account for year-to-date (YTD) monthly data that may exist for the current year. While the model is based on annual data, monthly information published before the year is complete provides valuable information, mainly when producing forecasts near the end of a year. Thus, for the current year, where a forecast needs to be produced, but some YTD monthly data is available, the model takes a weighted average of the base model forecast and a YTD-derived value to produce the final forecast.

This section illustrates the population model for Alberta, Calgary Economic Region (CER), Calgary CMA, and the city of Calgary. The Calgary model contains variables at the CER, CMA, and City (COC) levels, corresponding to what data happens to be available for each variable. The Calgary model follows the Alberta model in the model sequence for each forecast year. Since Calgary is contained within Alberta, and Alberta-level forecasts are produced before the Calgary forecast, the Calgary forecast is not independent of the Alberta forecast. Rather, most of the Calgary forecasts are driven by the Alberta forecasts, with an eye on factors that may cause trends in Calgary to differ from those in Alberta. The Alberta-level variable is often used as the primary explanatory variable, or the Calgary/Alberta share is forecasted rather than the Calgary variable itself.

This model produces Population forecasts for the CER, CMA, and COC levels. However, while the CER and CMA population forecasts are based on Statistics Canada data and have a reference date of July 1, the COC population forecast is based on the former Civic Census (available up to 2019), with a reference date of Apr 1.

For the City of Calgary, the net migration forecasts are not broken down into international and interprovincial components, but the assumptions about the distribution of net migration into cohorts, which come from the CMA, have this breakdown. Thus, for this distribution assumption, the share of CMA net migration for international and domestic is applied to the COC net migration forecast, and then the distribution assumptions can be applied.

Model Results-Alberta-Calgary Illustration

[2] Cohort-level mortality rates represent the probability that a cohort member will die each year. Fertility rates are assumed to change each year, though they converge to constant values after 15 years. The process for generating these assumptions is based on birth rates provided by Statistics Canada and Health Services. These are provided on an annual basis separately.

Alberta International Net Migration			
	Coefficients	P-Value	T-Statistic
Constant Term	-9.1e+04	0.000	-6.933
Lagged Alberta-Canada Population Ratio	1.1e+06	0.000	8.226
Differenced Lag in Total Alberta Employment	0.1107	0.000	3.773
Federal Immigration Policy Dummy	4.1e+04	0.000	5.899
R-squared/ Adjusted R-Squared	0.8270/ 0.8150		
Alberta Domestic Net Migration			
Constant Term	2.6e+04	0.052	2.001
Lagged Alberta Domestic Net Migration	0.3708	0.005	2.973
Alberta-Canada Unemployment Rate Ratio	-2.9e+04	0.043	-2.091
Alberta Residual Deviation	0.2821	0.480	0.713
Differenced Lag in Total Alberta Employment	0.1707	0.001	3.660
R-squared/ Adjusted R-Squared	0.686/0.654		
Calgary Economic Region (CER) International Net Migration			
Constant Term	-1.5e+04	0.056	2.046
Alberta International Net Migration	0.5277	0.000	29.595
CER-Alberta Unemployment Rate Ratio	-1.6e+04	0.037	-2.255
R-squared/ Adjusted R-Squared	0.980/0.978		
Calgary Economic Region (CER) Domestic Net Migration			
Constant Term	1.3e+04	0.193	1.329
Alberta Domestic Net Migration	0.3397	0.000	12.579
CER-Alberta Unemployment Rate Ratio	-9040.1	0.337	-0.974
R-squared/ Adjusted R-Squared	0.861/0.852		
Calgary CMA International Net Migration			
Constant Term	-150.8	0.123	-1.611
CER International Net Migration	-0.9935	0.000	207.153
R-squared/ Adjusted R-Squared	1.000/1.000		
Calgary CMA Domestic Net Migration			
Constant Term	-643.5	0.014	-2.704
CER Domestic Net Migration	0.9545	0.000	43.633
R-squared/ Adjusted R-Squared	0.990/0.989		
Calgary Municipality (COC) Net Migration			
Constant Term	-7515.205	0.885	-0.146
CER-Canada Unemployment Rate	-1.9e+04	0.044	-2.120
Lagged COC-Canada Population Share	9.6e+05	0.667	0.435
Differenced Lag in Total CER Employment	0.1746	0.098	1.718
COC Average House Price	9.9524	0.841	0.202
Calgary CMA Inflation	608.1275	0.612	0.513
Federal Immigration Policy Dummy	2.3e+04	0.016	2.588
R-squared/ Adjusted R-Squared	0.607/0.516		