Catalogue no. 11-633-X — No. 018
ISSN 2371-3429
ISBN 978-0-660-27484-3

Analytical Studies: Methods and References

Data Sources and Linking Methods for Long-run Provincial and Territorial Data: An Update for Gross Domestic Product, Urbanization, Unemployment and Depreciation

by Ryan Macdonald

Release date: November 2, 2018

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" >"Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Data Sources and Linking Methods for Long-run Provincial and Territorial Data: An Update for Gross Domestic Product, Urbanization, Unemployment and Depreciation

by

Ryan Macdonald

Economic Analysis Division
Statistics Canada

11-633-X No. 018
ISSN 2371-3429
ISBN 978-0-660-27484-3

October 2018

Analytical Studies: Methods and References

Papers in this series provide background discussions of the methods used to develop data for economic, health, and social analytical studies at Statistics Canada. They are intended to provide readers with information on the statistical methods, standards and definitions used to develop databases for research purposes. All papers in this series have undergone peer and institutional review to ensure that they conform to Statistics Canada's mandate and adhere to generally accepted standards of good professional practice.

The papers can be downloaded free at www.statcan.gc.ca.

Acknowledgements

The author would like to thank John Baldwin, Danny Leung and Wulong Gu for their support and input in creating the extended estimates. He would also like to thank Mark Brown for his feedback and Jim Tebrake and Catherine Van Rompaey for their comments, as well as participants at seminars where preliminary results were presented.

Table of contents

Abstract

This paper describes the data sources and methods used to backcast provincial and territorial income-based gross domestic product (GDP), expenditure-based GDP, real gross domestic income, unemployment rates, depreciation rates and urbanization rates. Instrumental variable techniques are used to estimate the historical movements of these economic variables back to 1950.

1 Introduction

This paper describes the linking and estimation procedures employed to estimate a new set of variables that describe long-run changes in the economies of the provinces and territories. Historical data sources do not always accord exactly with modern definitions, and not all variables can be perfectly recreated over the time spans described below. Nevertheless, estimates can be produced that are very close and which are useful for understanding the evolution of the provincial and territorial economies.

The first vintage for a number of the series described here was created for Brown and Macdonald (2015) and was used for analysis in that paper. To build on the original series, a second version was created for release in Table 36-10-0229-01 (Statistics Canada n.d.*j*). The second version has an increased scope that includes the territories. In addition, newly incorporated data sources and updated methodologies were used for some estimates. The result is a balanced panel dataset for the provinces, as well as a set of territorial estimates that expands through time as data become increasingly available.

The variables being added in this expansion are long-run estimates for provincial and territorial gross domestic product (GDP) and its constituent elements. The estimates are based on the income and final expenditure approaches to measuring GDP (see below). Price indexes for most components of expenditure-based GDP are also present. The exception is export and import prices, which were not recorded before 1981. Real estimates for final demand components and real gross domestic income (GDI) are reported in 2007 dollars. Finally, a number of analytically important ratios are provided.

Measures added to the long-run data

Current dollar measures (nominal)

- Gross domestic product (GDP)
    - Gross operating surplus
    - Gross mixed income
    - Taxes less subsidies
    - Effect of the statistical discrepancy
    - Gross fixed capital formation (GFCF)
        - Residential GFCF
        - Non-residential GFCF
    - Inventories
    - Net exports plus statistical discrepancy
    - Final domestic demand
    - Capital consumption allowances
    - Net domestic product

Price indexes

- GFCF
    - Residential GFCF
    - Non-residential GFCF
- Final domestic demand

Real measures

- Real gross domestic income (GDI)
    - Real total consumption
    - Real GFCF
        - Residential GFCF
        - Non-residential GFCF
    - Real final domestic demand

Ratios

- Depreciation rates — total capital stock
- Employment rates
- Saving rates
- Urbanization rates

Many of the new series do not have the same time span as the original series. Consequently, the expansion contains data with different starting years. One group of variables related to household income, population, consumer prices and urbanization rates starts in 1926, while a second set of variables related to GDP and analytically important ratios begins in 1950. Although the series come from different data programs and cover different aspects of economic activity, all the series were placed into a single table. The goal is to reduce search costs for analysts and publicize knowledge about the methods and quality of the series provided. The methods used to construct historical estimates can differ significantly from those used for modern estimates.

The long-run data are constructed from previous vintages of official publications and instrumental variable techniques. This approach enables the construction of an extended time series, but it cannot be viewed as being equivalent to the methods and practices employed in current statistical programs. The data are marked with quality indicators that differ by data source. Where modern data are employed, the data quality is indicated by an “A.” Where historical data from Statistics Canada or the Dominion Bureau of Statistics that are closely related to modern series are employed, the data quality is indicated by a “B.” Where data from external sources or instrumental variables are employed, the data quality is indicated with a “C.”

Because of the different data sources, data vintages and linking methodologies employed, the long-run data lend themselves to instrumental variable estimation techniques. In particular, the structure of the panel dataset lends itself to the use of internal instrumental variable estimators such as generalized method of moments (GMM) estimators that account for the measurement error inherent in the construction of long-run series. Previous knowledge of these types of instrumental variable estimators is one reason why considerable effort was made to create balanced panel datasets for the provinces, as their use allows for more accurate statistical inference.

The remainder of this document is structured as follows. The next section briefly discusses the methodological approach employed. Section 3 then discusses how the long-run time series for each variable was constructed. Section 4 concludes.

2 Linking methodology

The linking methodology ranks data sources from newest to oldest and by source (e.g., Statistics Canada, Dominion Bureau of Statistics, instrumental variables). The data are then linked using growth rates (preferred), ratios (second best), or nearest neighbour or regression techniques (third best). Macdonald (2015) discusses the philosophy used to construct the long-run data.

When statistical programs such as the Canadian System of Macroeconomic Accounts (CSMA) undergo major revisions, the most recent data are often the most heavily revised, and historical estimates are often the least affected. This is especially the case for growth rates. The relatively larger revision occurs as statistical systems adjust to structural changes in the economies upon which they are reporting. In practical terms, this means that more recent estimates are more heavily affected by the introduction of new vintages, and that while levels may change, growth rates (especially those for periods further in the past) are largely maintained.

In the majority of cases, historical estimates are linked to modern estimates using their growth rates. That is, the modern estimates are projected back through time using the growth rates of historical estimates. This is equivalent to treating historical estimates as though they are indexes with a base value equal to the earliest data point from the modern data vintage. The assumption applied for using growth rates is that each historical vintage captures the tendencies, variation and cyclical behaviour of the relevant variables over the period for which it is reported.

However, in some instances, growth rates cannot be used to create historical projections. For example, this occurs when historical estimates were not created but source data are present or when historical estimates are not present. In these instances, one of three alternative methods is employed.

The first method creates a projector for the modern series based on constituent elements, and the growth rate of the projector is then used to backcast the modern data.

The second method is to use ratios. In this case, the ratio of an alternative data source is employed to predict the historical ratio of the desired variable. When ratios are used, an adjustment must be made so that the ratio of the more recent data source and the ratio of the historical source data are comparable. For the series presented in the provincial dataset, the average difference for the overlapping periods between the two series is used.

When using growth rates or ratios is not feasible, regression-based estimates are employed to link data vintages. This is a less preferred method because regression estimates can expand or contract the variance of historical estimates in relation to modern estimates based on what is typically a limited sample of overlapping observations. Moreover, the overlapping observations may not be particularly representative of previous periods, and this can affect the reported magnitude of historical events.

In cases where data are more limited, a nearest neighbour match, a regression-based estimate from data that do not correspond as closely to the original data, or an alternative instrumental variable estimator is employed as a last resort option. The goal is for the dataset to provide the best possible estimates for a balanced panel dataset over an extended period. Data constraints require the use of instrumental variable construction techniques to accomplish this, and data quality indicators are included so that data users have the appropriate information to make decisions regarding data selection or estimation methodologies.

Once links are created and the data are backcast, provincial and territorial variables are benchmarked to the Canada levels. Benchmarking is performed for nominal dollar series. This ensures the additivity of GDP aggregates and household income estimates.

For each variable, the final data quality is assessed in a number of dimensions. First, the degree of concordance across vintages or instrumental variables is examined. The data are then examined to determine whether the backcasting affects trends, cycles, variations or ratios. This requires the use of judgment because no single rule is generally admissible. In most cases, the historical linking procedures work well. However, in a limited number of specific cases, an approach is widely applicable, but one jurisdiction presents less consistent results. In these situations, judgment is used to resolve the discrepancy. Finally, the data are examined for their information content. This is done to understand how and to what extent major historical events are present in the data. This type of analytical examination (see, for example, Brown and Macdonald 2015) enables data users to determine the extent to which the data correspond with what is currently known about the long-run evolution of the provincial and territorial economies and what new avenues of investigation the long-run data facilitate.

3 Linking procedures by variable

3.1 Unemployment rate

Unemployment rates are calculated from a combination of official data sources. Following the Second World War, Statistics Canada began collecting and publishing labour force information from the newly created Labour Force Survey (LFS). In 1976, a new version of the LFS was implemented. At that time, changes were made to methodologies, such as to the definition for the youngest participants in the labour market,[1] to the sampling methodologies and to the questions asked on questionnaires. As a consequence, the two sets of information do not exactly correspond, and their levels can differ. However, they do capture highly correlated features of the provincial unemployment rates through time, and the similarity in their progress through time is used to link them.

As the starting point, estimates for the period from 1976 to the present are drawn from Statistics Canada's current LFS in Table 14-10-0287-01 (Statistics Canada n.d.*f*), and these estimates form the base for the backcasting procedure. For the years 1966 to 1990, a set of estimates that was created to bridge the gap between the modern LFS and its predecessor is employed. These data are not currently available on the Statistics Canada web site, but they are available in paper format. This dataset is used to backcast the modern level to 1966. From 1950 to 1966, a set of provincial unemployment rates from Series D491-497 of the *Historical Statistics of Canada* (Denton, n.d.) is used to complete the backcasting.

For British Columbia, Ontario and Quebec, separate series exist in each vintage. For the remaining provinces, the original LFS and the estimates in the *Historical Statistics of Canada* (Denton, n.d.) contain an aggregate for the Prairie Provinces (Alberta, Saskatchewan and Manitoba) and the Atlantic Provinces (New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland and Labrador). For the period from 1950 to 1966, the aggregate is used to backcast each of the constituent provinces except Newfoundland and Labrador. For Newfoundland and Labrador, *Historical Statistics of Canada* estimates for the unemployment rate are used back to 1957, and the Atlantic aggregate is used as an instrument thereafter. As a result, the relative difference between each province in a given aggregate is fixed, but the relative difference between a province and provinces not in its aggregate continues to evolve separately (Table 1). Because of their high quality, the source data for the unemployment rates are taken as is, and their growth rates are used to backcast the level of unemployment in each province.

1. The introduction of the modern LFS in 1976 led to changes in the definitions of labour force variables and the questionnaire used for the survey. For a discussion of the changes and their impact, see Statistics Canada (n.d.*a*).

Table 1
Unemployment rate data sources by province and for Canada

	Present to 1976 (modern Labour Force Survey)	1975 to 1966 (original Labour Force Survey)	1965 to 1950 (Historical Statistics of Canada)
Newfoundland and Labrador	Provincial estimate	Provincial estimate	Provincial estimate 1965 to 1957; Estimate based on aggregate of Atlantic provinces 1950 to 1956
Prince Edward Island	Provincial estimate	Provincial estimate	Estimate based on aggregate of Atlantic provinces
Nova Scotia	Provincial estimate	Provincial estimate	Estimate based on aggregate of Atlantic provinces
New Brunswick	Provincial estimate	Provincial estimate	Estimate based on aggregate of Atlantic provinces
Quebec	Provincial estimate	Provincial estimate	Provincial estimate
Ontario	Provincial estimate	Provincial estimate	Provincial estimate
Manitoba	Provincial estimate	Provincial estimate	Estimate based on aggregate of Prairie provinces
Saskatchewan	Provincial estimate	Provincial estimate	Estimate based on aggregate of Prairie provinces
Alberta	Provincial estimate	Provincial estimate	Estimate based on aggregate of Prairie provinces
British Columbia	Provincial estimate	Provincial estimate	Provincial estimate
Yukon	Territorial estimate, 1992 to present only	…	…
North West Territories	Territorial estimate, 2001 to present only	…	…
Nunavut	Territorial estimate, 2004 to present only	…	…
Canada	National estimate	National estimate	National estimate

… not applicable

Source: Statistics Canada.

Data for the territories are taken from Table 14-10-0292-01 (Statistics Canada n.d.*g*), which reports estimates of the unemployment rate based on three-month moving averages. The data for the territories are shorter than those for the provinces and are reported on a different basis. For this reason, the data are reported separately from the provinces when LFS estimates are reported (Statistics Canada 2017).

3.2 Gross domestic product, income based

Estimates of long-run GDP are derived using the income approach to estimating GDP. To create the long-run GDP series, a bottom-up approach is used. Each of the major components of income is backcast individually, and their values are benchmarked to the Canada-level data. The sum of the components forms the projector for income-based GDP in each jurisdiction. For the Canada-level data, previous vintages of GDP are employed rather than a projector. This enables more consistent backcasting at the Canada level, which, when used as a benchmark, helps limit the extent of measurement error from the regional projectors.

Because of differences in historical data granularity, the components of income-based GDP are grouped into compensation of employees, gross operating surplus, gross mixed income and taxes less subsidies. Estimates for compensation of employees (i.e., wages, salaries and supplementary labour income) were taken from previously reported long-run estimates published in Table 36-10-0229-01 (Statistics Canada n.d.*j*). For the remaining components, compiling long-run estimates of provincial income-based GDP requires the use of projectors for the 1950s, as well as past vintages of official CSMA estimates. The earliest set of provincial GDP estimates, which spans from 1961 to 1980, is in Table 36-10-0325-01 (Statistics Canada n.d.*k*), while the

modern set of provincial GDP estimates, which spans from 1981 to the present, is in Table 36-10-0221-01 (Statistics Canada n.d.*m*). An internal version of the data bridges the two published sets of information and is used to connect the two.

To enforce the additivity of results, the difference between the benchmarked GDP level of each province and the projected value is referred to as the effect of the statistical discrepancy. This reflects the inclusion of the statistical discrepancy in the modern data that cannot be backcast.

3.2.1 Capital consumption allowance

For most years, estimates of capital consumption allowance (CCA) come directly from CSMA sources. Modern estimates of CCA for 1981 to the latest data point are taken from current income-based GDP estimates included in Table 36-10-0221-01 (Statistics Canada n.d.*m*). Estimates for 1961 to 1980 are taken from a terminated set of provincial income-based GDP estimates in Table 36-10-0325-01 (Statistics Canada n.d.*k*).

For 1950 to 1960, the CCA of the provinces must be estimated indirectly. Because the CCA in GDP estimates is derived from across the economy, the CCA from residential investment must be incorporated with the CCA from non-residential investment.

For residential investment, estimates of CCA for persons and unincorporated businesses for the provinces are derived by allocating CCA estimates for Canada from Statistics Canada (1988) among the provinces according to annual weights derived from the flows and stocks of residential capital published in Table 34-10-0122-01 (Statistics Canada n.d.*c*).

For non-residential investment, CCA estimates from Statistics Canada (1988) are divided among the provinces using weights calculated from the flows and stocks of non-residential capital in Table 36-10-0236-01 (Statistics Canada n.d.*d*). Because the non-residential depreciation estimates begin in 1955, annual weights are calculated for the years 1955 to 1960, while the average of the years 1955 to 1960 is used to allocate the national values for the years 1950 to 1954.

3.2.2 Gross operating surplus

Gross operating surplus for the years 1981 to the present is taken directly from modern estimates published in Table 36-10-0221-01 (Statistics Canada n.d.*m*). To construct the historical estimates, two projectors were created.

The first uses an internal dataset that contains estimates for income-based GDP from 1961 to 1995. The estimates from this dataset, except for some revisions to the Yukon and the Northwest Territories from 1976 to 1980, are identical to the data published in the earliest set of provincial data found in Table 36-10-0255-01 (Statistics Canada n.d.*h*). However, the internal dataset includes an overlapping period with the modern data. This makes it more suitable for linking.

Estimates for corporate profits before taxes and the inventory valuation adjustment from 1961 to 1995 are combined with the above-described measure of non-residential CCA to produce a projector for each province and territory. This projector is then used to backcast the level of gross operating surplus from 1980 to 1961 using growth rates.

For the period from 1950 to 1960, a projector that combines estimates of corporate taxable income by province included in Table 33-10-0049-01 (Statistics Canada n.d.*e*) and non-residential CCA is employed. Corporate taxable income is used here as an approximation of net operating surplus. This projector is then used to backcast the level of gross operating surplus from 1960 to 1950.

3.2.3 Gross mixed income

Modern estimates of gross mixed income are taken directly from Table 36-10-0221-01 (Statistics Canada n.d.*m*). Historical estimates are based on measures of net farm income, net unincorporated business income (including imputed rental income for owner-occupied dwellings) and interest, dividends and miscellaneous investment income of persons taken from the 1968 System of National Accounts published by Statistics Canada (1988). This is the same data source used for estimates of household income and compensation of employees.

The estimates of net farm income, net unincorporated business income (including imputed rental income for owner-occupied dwellings) and interest, dividends and miscellaneous investment income of persons are summed with the long-run estimate of residential CCA to produce an estimate of gross mixed income from 1950 to 1986. This estimate is used to backcast the level of modern gross mixed income from 1980 to 1950.

3.2.4 Taxes less subsidies

Estimates of taxes less subsidies from 1981 to the present combine taxes less subsidies on production and taxes less subsidies on products. Estimates of total taxes less subsidies from the internal income-based GDP vintage spanning from 1961 to 1995 are then used to backcast the modern level.

The period from 1950 to 1960 is more challenging, because historical sources are more difficult to find. The average share of taxes less subsidies from each jurisdiction in the Canada-level total from 1961 to 1970 is used to split the level of taxes less subsidies for Canada. The Canada-level data are taken from the 1968 System of National Accounts published by Statistics Canada (1988).

3.3 Gross domestic product, expenditure based

Expenditure-based estimates of GDP are assumed to equal the income-based estimates. This enables the disaggregation of the components of expenditure-based GDP into total consumption, investment, inventories and a residual measure of net exports plus statistical discrepancy. The measurement of each component is described below.

3.3.1 Consumption

Two types of consumption are recorded in the final expenditure estimates of GDP—collective consumption and individual consumption. Here, the sum of the two is estimated historically. Households, non-profit institutions serving households (NPISHs) and governments do not have separate values.

Estimates of consumption by province are linked according to the share of consumption in each province from three data sources. The modern data are drawn from expenditure-based GDP estimates in Table 36-10-0222-01 (Statistics Canada n.d.*n*). An internal dataset with estimates spanning from 1961 to 1995 is used to estimate the share of consumption for the years 1961 to 1980. This dataset is consistent with historical estimates of expenditure-based GDP by province published in Table 36-10-0255-01 (Statistics Canada n.d.*h*), but it contains an overlapping period with modern estimates that facilitates linking. Lastly, estimates of retail trade by province from Series V89-99 from Moyer (n.d.) are used to estimate consumption for the years 1950 to 1960.

For the period from 1961 to 1980, the share of consumption in GDP is calculated for the internal estimates and the modern data. The average difference between the modern share of consumption and the share of consumption in the historical dataset is calculated. The difference is used to adjust the historical share of consumption up or down to be consistent with the modern data.

For the period from 1950 to 1960, the share of retail sales by province is used to estimate the consumption share of GDP for each province and for the territories. The data from the *Historical Statistics of Canada* (Moyer n.d.) contain estimates that span from the 1920s to 1976, but, during the early period, estimates for the Atlantic Provinces are provided annually as an aggregate and separately only every decade. Therefore, the aggregate must be split to the provincial level before estimating consumption shares. To do so, a linear interpolation for the share of retail sales in each province is created, and this interpolated share is used to split the aggregate retail sales for the Atlantic Provinces into the constituent parts.

Next, the share of retail sales in each province is calculated, and, according to the overlapping period from 1961 to 1976, an average adjustment factor is calculated. The share of retail sales plus the adjustment factor (which may be positive or negative) is then used to estimate the share of consumption in GDP by province.

Finally, the level of consumption in current dollars for each province and for the territories is estimated by multiplying the share of consumption by GDP.

3.3.2 Gross capital formation (or saving equals investment)

No direct method exists to calculate the aggregate saving rate for each of the provinces over the period from 1950 to 2016. As a result, the aggregate saving estimate for the provinces is based on the saving-equals-investment identity embedded in the System of National Accounts (United Nations. Statistical Office 1968). To use this identity, the total value of investment across all assets and sectors is aggregated. Consequently, the investment series includes gross fixed capital formation and inventory investment from corporate and unincorporated business, NPISHs, households, and governments.

To construct the provincial saving or investment rate estimates, separate links are made for inventories, residential investment and non-housing investment. The inventory values are described in the next section. The provincial residential series links residential investment from the current expenditure-based GDP estimates published in Table 36-10-0222-01 (Statistics Canada n.d.*n*) with historical estimates based on the flows and stocks of residential capital published in Table 34-10-0049-01 (Statistics Canada n.d.*c*). The historical estimates extend back to 1941 and therefore enable estimation of investment in housing by province back to 1950 with a single data source. The historical estimates are used to backcast the level of the modern series from 1980 to 1950.

At the Canada level, residential investment is linked with estimates of residential investment from Statistics Canada (1988). These estimates produce a nearly identical result to estimates from the flows and stocks, but enable a longer period to be used.

To construct provincial non-housing investment, estimates of public and private investment excluding residential investment are calculated from current expenditure-based estimates of GDP. Historical estimates for the period from 1955 to 1980 are taken from the flows and stocks of fixed non-residential capital in Table 36-10-0236-01 (Statistics Canada n.d.*d*). These estimates are then used to backcast the modern level of investment to 1955.

Before 1955, estimates of investment by province are not available. Therefore, for 1950 to 1954, investment in non-housing capital is estimated using the average industry shares of investment from each province over the period from 1955 to 1959. To do so, the share of each province's investment in a particular three-digit industry of the North American Industry Classification System is averaged over the 1955-to-1959 period. Then, this weight is used to attribute a share of investment from that industry to the province for the years 1950 to 1954. The industry-based series are subsequently aggregated, and the implied investment series is used to backcast the provincial estimate for investment in non-housing fixed assets.

At the Canada level, the estimates from Statistics Canada (1988) for gross fixed capital formation excluding residential investments are used to backcast the modern level.

3.3.3 Inventories

Inventories are linked based on the share of inventories in GDP from the modern data found in Table 36-10-0222-01 (Statistics Canada n.d.*n*), the internal dataset that is consistent with the original version of the expenditure-based GDP estimates published in Table 36-10-0325-01 (Statistics Canada n.d.*k*) and a projector of inventory investment by industry based on inventory shares. The modern data span 1981 to the present while the internal dataset consistent with the original provincial GDP estimates is employed for the period from 1961 to 1980.

For the period from 1950 to 1960, the share of GDP from inventory investment is inferred from a projector. The projector is created by regressing inventory investment estimates for agriculture; forestry; mining, quarries and oil wells; manufacturing; construction; transportation, storage and communications; electric power, gas and water utilities; trade, grain in commercial channels; wholesale trade; retail trade; finance, community business and personal services; and total government found in Table 36-10-0175-01 (Statistics Canada n.d.*i*) on the share of inventory investment in each province and the aggregate of the three provinces for the period from 1961 to 1995. The projector is then used to estimate the share of inventory investment in each province between 1950 and 1960.

Despite the fact that the projector provides an indication of inventory investment, a word of caution is warranted here. Since the industry breakdown used to create the projector is not overly detailed, stronger assumptions about the structure and dispersion of inventory investment are imposed on the inventory projector than on any other instrument created for this dataset to date. For some provinces, such as Ontario, Quebec and Saskatchewan, the projector nearly duplicates the official data for the overlapping period from 1961 to 1995. However, for other provinces, such as Nova Scotia, the difference between historical estimates, the projector and modern estimates is greater. Similarly, the estimates for Newfoundland and Labrador use assumptions about a period when Newfoundland and Labrador was more fully integrated into Canada to predict what happened during the period of its transition after the Second World War. As a result, while the data are deemed useable overall, year- and province-specific changes may be overstated or understated in some instances when this projection methodology is used.

3.3.4 Net exports plus statistical discrepancy

The estimates of GDP, final consumption, gross fixed capital formation and inventories can be used to infer the value of net exports plus statistical discrepancy. This was the practice in the original version of the expenditure-based GDP estimates published in Table 36-10-0325-01 (Statistics Canada n.d.*k*), because trade flows by province were not available. The same practice is followed here. Net exports plus statistical discrepancy are estimated as GDP less consumption, less gross fixed capital formation, less inventories.

For net exports, a note of caution is warranted. Because they are measured residually, net exports potentially contain errors from all projectors employed. As a result, they are subject to greater uncertainty than other estimates.

3.3.5 Final domestic demand and final domestic expenditure

Estimates of final domestic demand are calculated as the sum of total consumption and gross fixed capital formation. Estimates of final domestic expenditure are calculated as the sum of total consumption, gross fixed capital formation and inventories.

3.4 The gross fixed capital formation deflator and the real gross domestic income deflator

Real GDI is a measure of the purchasing power of domestic production. It combines changes in real GDP with the effects of relative price movements such as changes in the terms of trade (Macdonald 2010). Two ways exist to estimate real GDI. The first is to begin with real GDP and add the effects of relative price changes. The second is to deflate nominal GDP with a broad measure of the movement of domestic prices. Here, a final domestic demand deflator is employed. This approach is applied because it is less data-intensive than beginning with real GDP.

To form the real GDI deflator, the index of final domestic demand for the period from 1981 to the present is taken from estimates of expenditure-based GDP from Table 36-10-0221-01 (Statistics Canada n.d.*m*). A historical estimate for the final domestic demand deflator based on data from Table 36-10-0331-01 (Statistics Canada n.d.*l*) is then used to backcast the modern index to 1971.

Finally, for the period from 1950 to 1970, an estimate of the final domestic demand deflator is created using a linked consumption price index based on data from Table 36-10-0229-01 (Statistics Canada n.d.*j*) and an estimate of the price of gross fixed capital formation based on the flows and stocks of fixed non-residential capital published in Table 36-10-0236-01 (Statistics Canada n.d.*d*) and the flows and stocks of residential capital published in Table 34-10-0122-01 (Statistics Canada n.d.*c*). In the case of the former, the price index for non-residential capital is available only to 1955. As a result, an estimate for the progress of non-residential gross fixed capital formation for Canada from Statistics Canada (1988) is employed for all jurisdictions for the period from 1950 to 1954. This gross fixed capital deflator is linked to the modern gross fixed capital price deflator to create its long-run estimate.

3.5 Real variables

Real GDI is estimated by deflating the historically linked estimates for income-based GDP by the historically linked estimates for the final domestic demand deflator. The same deflator is used to estimate real final domestic demand. The long-run linked CPI deflator is used to estimate real consumption, while the long-run gross fixed capital formation deflator is used for that variable.

3.6 Depreciation rate

Estimates of the depreciation rates are taken from Statistics Canada's gross fixed capital formation survey results shown in Table 33-10-0049-01 (Statistics Canada n.d.*e*) and correspond to the depreciation rates for all assets except housing. For the years 1955 to the present, the depreciation rate is the ratio of geometric depreciation to capital stock, constructed using a geometric depreciation rate.

For the years 1950 to 1954, an estimate of the depreciation rate is constructed by regressing time and time squared on the logarithm of depreciation and the logarithm of the capital stock. The fitted values from the regressions are then employed to estimate the depreciation rate according to their levels after taking the anti-log of the predicted values.

3.7 Urbanization rate

Urbanization rates are taken from calculations made by the Demography Division of Statistics Canada.[2] The data are available by census period, every 10 years from 1881 to 1951 and every 5 years thereafter. A linear interpolation between decadal or bi-decadal reference points is used to create an annual time series. For those years after the latest Census, a projection based on the last 15 years is used to estimate the urbanization rate for each province and territory.

2. See Statistics Canada (n.d.*b,o*) for more details. Internal calculations were done for the years before 1971.

4 Conclusion

The long-run data are linked using instrumental variables techniques that combine modern and historical data sources. The estimation process employs current and historical vintages of the Canadian System of Macroeconomic Accounts (CSMA), as well as data from surveys and censes that the CSMA employ as source data.

The estimation methods allow for examinations of economic phenomenon such as the size and evolution of provincial and territorial income shares, comparisons across cycles or for comparisons among provinces and territories. They are also constructed with the aim of employing panel data estimators to derive values for parameters in economic models.

Despite their strengths, the linked data are not necessarily suitable for all uses, and assessments should be made about the coherence of the linked data with other data sources prior to use. The linking process creates historical level estimates for CSMA variables that are consistent within the data set. However, these estimates will not necessarily accord with other data sources in the same way that modern data does. Studies examining government finances, for example, would need to exercise caution as the linked data for historical periods will not be as consistent with government finance statistics as modern data sources are.

Nevertheless, the linked data present a rich dataset for understanding the evolution of provincial and territorial economies. They illustrate how complementary data sources may be combined to produce estimates of aggregate economic variables over extended periods of time, and permit greater insight into a range of Canadian economic phenomena.

References

Brown, W.M., and R. Macdonald. 2015. *Provincial Convergence and Divergence in Canada, 1926 to 2011*. Economic Analysis Research Paper Series, no. 96. Statistics Canada Catalogue no. 11F0027M. Ottawa: Statistics Canada.

Denton, F.D. n.d. "Section D: The Labour Force." *Historical Statistics of Canada.* Statistics Canada Catalogue no. 11-516-X. Last updated July 2, 2014. Available at: http://www.statcan.gc.ca/pub/11-516-X/sectiond/4057750-eng.htm (accessed April 11, 2018). Archived.

Macdonald, R. 2010. "Real gross domestic income, relative prices and economic performance across the OECD." *Review of Income and Wealth* 56 (3): 498–518.

Macdonald, R. 2015. *Constructing Provincial Time Series: A Discussion of Data Sources and Methods*. Income and Expenditure Accounts Technical Series, no. 77. Statistics Canada Catalogue no. 13-604-M. Ottawa: Statistics Canada.

Moyer, M.S. n.d. "Section V: Internal trade." *Historical Statistics of Canada.* Statistics Canada Catalogue no. 11-516-X. Last updated July 2, 2014. Available at: http://www.statcan.gc.ca/pub/11-516-x/sectionv/4057758-eng.htm (accessed December 7, 2017). Archived.

Statistics Canada. n.d.*a.* "Comparing Current LFS Estimates to Those Prior to 1976," *Labour Force Survey (LFS).* Last updated May 5, 2017. Available at: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3701 (accessed April 6, 2018). Archived – PDF document 3701_D8_T9_V1-eng.pdf.

Statistics Canada. n.d.*b*. *Table 17-10-0118-01*. *Selected population characteristics, Canada, provinces and territories.* Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710011801 (accessed October 4, 2018).

Statistics Canada. n.d.*c*. *Table 34-10-0122-01 Flows and stocks of fixed residential capital, *Archived* annual (dollars x 1,000,000)* (table). Last updated November 8, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3410012201 (accessed July 4, 2018).

Statistics Canada. n.d.*d*. *Table 36-10-0236-01 Flows and stocks of fixed non-residential capital, by North American Industry Classification System (NAICS) and asset, Canada, provinces and territories, *Archived* annual (dollars x 1,000,000)* (table). Last updated July 4, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610023601 (accessed July 4, 2018).

Statistics Canada. n.d.*e*. *Table 33-10-0049-01 Provincial allocation of corporate taxable income, by industry, *Archived* annual (data in thousands)* (table). Last updated July 4, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3310004901 (accessed July 4, 2018).

Statistics Canada. n.d.*f. Table 14-10-0287-01 Labour Force Survey estimates (LFS), by sex and age group, seasonally adjusted and unadjusted monthly (persons unless otherwise noted)* (table). Last updated June 8, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410028701 (accessed July 4, 2018).

Statistics Canada. n.d.*g. Table 14-10-0292-01 Labour Force Survey estimates (LFS), by territories, sex and age group, three-month moving average, seasonally adjusted and unadjusted monthly (persons unless otherwise noted)* (table). Last updated June 8, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410029201&request_locale=en (accessed July 4, 2018).

Statistics Canada. n.d.*h. Table 36-10-0255-01 Gross domestic product (GDP), expenditure-based, *Archived* annual (dollars x 1,000,000)* (table). Last updated July 4, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610025501 (accessed July 4, 2018).

Statistics Canada. n.d.*i. Table 36-10-0175-01 Investment in inventories, by industry, 1968 System of National Accounts (SNA) annual (dollars x 1,000,000)* (table). Last updated July 4, 2018. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610017501 (accessed July 4, 2018).

Statistics Canada. n.d.*j. Table 36-10-0229-01 Data on long-run provincial and territorial economic performance annual (dollars)* (table). Last updated November 4, 2016. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610022901&request_locale=en (accessed July 4, 2018).

Statistics Canada. n.d.*k. Table 36-10-0325-01 Provincial gross domestic product (GDP), expenditure-based, *Archived* annual (dollars x 1,000,000)* (table). Last updated July 4, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610032501 (accessed July 4, 2018).

Statistics Canada. n.d.*l. Table 36-10-0331-01 Implicit price indexes, gross domestic product (GDP), *Archived* annual (index, 1986=100)* (table). Last updated July 4, 2018. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610033101 (accessed July 4, 2018).

Statistics Canada. n.d.*m. Table 36-10-0221-01 Gross domestic product, income-based, provincial and territorial annual (dollars x 1,000,000)* (table). Last updated November 8, 2017. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610022101 (accessed July 4, 2018).

Statistics Canada. n.d.*n. Table 36-10-0222-01 Gross domestic product, expenditure-based, provincial and territorial annual (dollars x 1,000,000)* (table). Last updated November 8, 2017. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610022201 (accessed July 4, 2018).

Statistics Canada. n.d.*o. Population and Dwelling Count Highlight Tables. 2016 Census*. Last updated June 19, 2017. Available at: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/index-eng.cfm (accessed October 4, 2018).

Statistics Canada. 1988. *National Income and Expenditure Accounts: Annual Estimates, 1926–1986*. Statistics Canada Catalogue no. 13-531-X. Ottawa: Statistics Canada.

Statistics Canada. 2017. *Guide to the Labour Force Survey*. Statistics Canada Catalogue no. 71-543-G. Ottawa: Statistics Canada. Available at: http://www.statcan.gc.ca/pub/71-543-g/71-543-g2017001-eng.htm (accessed February 9, 2018).

United Nations. Statistical Office. 1968. *A System of National Accounts: [1968]*. New York: United Nations.