**``Preliminary Form to Add Social Determinants to CSDUL**

**Request date (2024-11-26):**

|  |  |
| --- | --- |
| **Researcher (name and affiliation):** | Tianjiao Yang, Dalhousie University |
| **Node Lead (name and affiliation):** | Daniel Dutton, Dalhousie University |
| **Indicator or Model Name:** | CCOFOG and Macroeconomic Indicators |

**Purpose of the document**

This document includes several questions that must be answered by the researcher interested in incorporating dataset information, indicators or models into CSDUL. These questions pretend to briefly explain the mathematical and theoretical framework of the indicator or model being incorporated. The researcher must be able to fill out every question clearly and concisely, supporting their explanation with respectable academic sources.

The document will be added to the model or indicator documentation in CSDUL-OUT and CSDUL-RDC. It must serve as a quick and straightforward introduction to the indicator or model for anyone interested and give relevant references to guide the learning process to other researchers.

**To be completed by the responsible analyst.**

**If there are questions that cannot be answered because of the nature of the indicator/model, write N/A.**

**For the second part, you can support your completion using the example document located in this link:** [**Documents - Add inputs to CSDUL - 02 - Example.docx - Google Docs**](https://docs.google.com/document/d/1t4_Bh5pRtHzd8GQ3ifJWY2zjjjch8DFf/edit)

1. **Data Description:** We are compiling theCanadian Classification of the Functions of Government **(**CCOFOG) and Macroeconomic Indicators datasets which share the same characteristics in terms of start and end dates, frequency, and number of observations. Both are pre-existing indicators identified and accessible through Statistics Canada. CCOFOG describes government expenditures based on their final use which is structured into three levels: division, group, and class, each with a unique numerical code to indicate its position. The first three digits represent the division, which is the highest level in CCOFOG. The inclusion of a fourth digit indicates the group within a division. With five digits, the code represents a class within a group. Macroeconomic indicators describe the overall state of the economy and include measures such as Gini coefficient, unemployment rate, Gross Domestic Product, median after-tax income and consumer price indexes.
2. **Beginning Date:** 2008
3. **End Date:** 2023
4. **Sampling Frequency of Data:** Annual
5. **Available Periods:** 2008 - 2023
6. **Number of observations:** 160province-year (10 provinces, each observed over a consecutive period of 16 years from 2008 to 2023)
7. **Key variables**
   * 1. **Dataset 1 - Canadian Classification of the Functions of Government (CCOFOG) - Values are in millions of dollars (Nominal/Current dollars[[1]](#footnote-2)):** 
        1. **Health [707]:** government expenditures relating to the administration, development, and management of health consist of the following groups
           1. **Medical products, appliances, and equipment [7071]:** government expenditures on medicaments, prostheses, medical appliances, and equipment, and other health-related products obtained by individuals or households, either with or without a prescription, usually from dispensing chemists, pharmacists, or medical equipment suppliers for consumption or use outside a health facility or institution
           2. **Outpatient services [7072]:** government expenditures on medical, dental, and paramedical services delivered to outpatients by medical, dental, and paramedical practitioners and auxiliaries. The services may be delivered at home, in individual or group consulting facilities, dispensaries, or the outpatient clinics of hospitals and the like
           3. **Hospital services [7073]:** government expenditures on the services of general and specialist hospitals, the services of medical centers, maternity centers, nursing homes, and convalescent homes that chiefly provide in-patient services; the services of military base hospitals; the services of institutions serving elderly people in which medical monitoring is an essential component; and the services of rehabilitation centers providing in-patient health care and rehabilitative therapy where the objective is to treat the patient rather than to provide long-term support
           4. **Public health services [7074]:** government expenditures on the provision of public health services; administration, inspection, operation, or support of public health services, such as blood bank operation, disease detection, prevention, monitoring, epidemiological data collection, family planning services, and so forth; and preparation and dissemination of information on public health matters
           5. **Health not elsewhere classified [7075, 7076]:** government expenditures on administration and operation of government agencies engaged in applied research and experimental development related to health, as well as grants, loans, and subsidies to support applied research and experimental development related to health undertaken by non-government bodies, including research institutes and universities. This also includes expenditures on collective health prevention programs and all other health-related expenditures not elsewhere classified
        2. **Social protection [710]: g**overnment expenditures on the administration, development, and management of social protection are categorized into the following groups. These expenditures include spending on services and transfers provided to individuals and households, as well as spending on services provided on a collective base
           1. **Sickness and disability [7101, 7103]:** government expenditures on the provision of social protection for sickness or disability and for survivors of a deceased person. This includes expenditures on cash or in-kind benefits for loss of earnings due to temporary inability to work caused by sickness or injury, as well as support for persons who are survivors of a deceased person. This also includes expenditures on administration, operation, or support of these social protection schemes
           2. **Old age [7102]:** government expenditures on the provision of social protection for old age. This includes expenditures on cash or in-kind benefits for loss of income, inadequate income, or lack of independence in carrying out daily tasks, etc. due to aging, as well as expenditures on administration, operation, or support of such social protection schemes
           3. **Family and children [7104]:** government expenditures on the provision of social protection for households with dependent children. This includes expenditures on cash or in-kind benefits such as maternity allowances, birth grants, parental level benefits, shelter and board to preschool children, etc., as well as expenditures on administration, operation, or support of such social protection schemes
           4. **Unemployment [7105]:** government expenditures on the provision of social protection for unemployment. This includes expenditures on cash or in-kind benefits (e.g. unemployment benefits, early retirement benefits, mobility and resettlement payments, and vocational training, etc.) for persons who are capable of work and available for work but are unable to find suitable employment, as well as expenditures on administration, operation, or support of such social protection schemes
           5. **Housing [7106]:** government expenditures on the provision of social protection for households’ cost of housing. This includes expenditures on in-kind benefits such as the provision of low-cost or social housing, payments to alleviate the current housing costs of oner-occupiers, etc., as well as expenditures on administration, operation, or support of such social protection schemes
           6. **Social exclusion not elsewhere classified [7107]:** government expenditures on the provision of social protection for victims’ services and social exclusion not elsewhere classified. Victim’s services refer to the provision provided for persons who have suffered physical or emotional harm, property damage, or economic loss as a result of crime or for their surviving dependents (i.e. spouse, child, or other relatives who were, in whole or in part, dependent upon the victim for support at the time of their death). Social exclusion not elsewhere classified refers to the provision of social protection to persons who are socially marginalized or at risk of social marginalization (e.g. lower income earners, refugees, etc.)
           7. **Social protection not elsewhere classified [7108, 7109]:** government expenditures on social protection-related applied research and experimental development and social protection not elsewhere classified. This includes expenditures on the administration and operation of government agencies engaged in applied research and experimental development related to social protection, as well as expenditures supporting social protection-related research and development undertaken by non-government bodies. Social protection not elsewhere classified includes expenditures on administration, operation, or support of activities such as formulation, administration, coordination, and monitoring of overall social protection policies, plans, programs, and budgets; preparation and enforcement of legislation and standards for the provision of social protection; and production and dissemination of general information, technical documentation, and statistics on social protection
     2. **Dataset 2 - Macroeconomic Indicators:**
        1. **Gini coefficient - adjusted for after-tax income:** a number between 0 and 1 that measures the degree of inequality in the distribution of income. A value of **0** represents perfect equality, where everyone has the same adjusted household income, while a value of **1** indicates maximum inequality, where one person receives all the income and everyone else receives none. Although a single Gini coefficient does not have a simple standalone interpretation, it is useful for comparing income inequality across different populations or over time: **the higher the coefficient, the greater the inequality.**
        2. **Unemployment rate:** the number of unemployed individuals expressed as a percentage of the total **labour force**. For a specific group (such as by age, gender, or marital status), the unemployment rate is calculated as the number of unemployed persons within that group divided by the labour force of that same group, expressed as a percentage.
        3. **Gross domestic product (GDP) - Values are in millions of dollars (Nominal/Current dollars):** income-based GDP that measures the total income generated by the production of all final goods and services within a region and time frame at market prices.
        4. **Median after-tax income:** median after-tax income (total income less income tax) for all families and unattached individuals in 2023 constant dollars[[2]](#footnote-3).
        5. **Consumer price index (CPI) - all-items:** an index that measures the cost of a fixed market basket of goods and services purchased by households in 2002 constant dollars.
        6. **Consumer price index - general government final consumption expenditure:** an index that measures the average changes in cost for all general government final consumption expenditure in 2017 constant dollars.
8. **Support Documentation**
   1. [**PowerPoint slides for an overview of CCOFOG and Macroeconomic Indicators**](https://dalu-my.sharepoint.com/:p:/r/personal/tn339828_dal_ca/Documents/CSDUL/Node1_CCOFOG_and_Macroeconomic_Indicator.pptx?d=wc283b7c3b29e4f66bd9d303d0b636292&csf=1&web=1&e=ODoEio)
   2. [**An Excel spreadsheet file with variables and datasets**](https://dalu-my.sharepoint.com/:x:/r/personal/tn339828_dal_ca/Documents/CSDUL/Data/CCOFOG_and_Macroeconomic_Indicator.xlsx?d=wa862b2b24dc74e2f93cbc23630a5535e&csf=1&web=1&e=RqNaC0)
   3. [**A Word doc with all links direct to data/variable Statistics Canada website**](https://dalu-my.sharepoint.com/:w:/r/personal/tn339828_dal_ca/Documents/Node1_Links_to_StatsCan_Datasource.docx?d=w07fc25efa84f4aa78ff499ba236f3189&csf=1&web=1&e=y3itqq)
   4. [**A link to Statistics Canada’s website with further detailed information on other variables of government expenditures by function in the CCOFOG**](https://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=201435&CVD=201436&CPV=7&CST=01011960&CLV=1&MLV=4)
9. **Indicator Description:**
10. **Will you share the inputs through CSDUL-RDC, CSDUL-OUT, or both?**

CSDUL-Out

1. **Explanation of the indicator/model.**

Macroeconomic indicators describe the overall state of the economy and include measures such as Gini coefficient, unemployment rate, Gross Domestic Product, median after-tax income and consumer price indexes.

* 1. **In simple words, explain what the indicator/model to be added consists of.**

Macroeconomic indicators describe the overall state of the economy and include measures such as Gini coefficient, unemployment rate, Gross Domestic Product, median after-tax income and consumer price indexes.

* 1. **Are there assumptions associated with the indicator/model? If there are, please briefly describe them.**
  2. **How is the indicator/model derived? Support your explanation with formulas when possible.**
* Macroeconomic indicators:
  + Gini coefficient: The coefficient would register zero (minimum inequality) for a population in which each person received exactly the same adjusted household income and it would register a coefficient of one (maximum inequality) if one person received all the adjusted household income and the rest received none.
    - * : Income for unit i
      * : Survey weight for unit i
      * : Total survey weight
      * : Total weighted income
  + Unemployment rate:
  + [Gross Domestic Product – income based](https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=1902):
    - Sum all of the factor incomes generated by productive activity - incomes representing the returns to the labour and capital employed.
  + [Median after-tax income](https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/app/index-eng.cfm?ID=A2_3):
    - The amount of after-tax income that divides the income distribution of that group into two halves, i.e., the incomes of half of the units in that group are below the median, while those of the other half are above the median.
      * : Lower boundary of the income group in which falls where
        + N: Number of (weighted) units in the specified group for which the distribution is being shown
        + : Weight of each unit in the specified group, for the short-from census questionnaire, the weight is equal to 1 for each unit
      * : Size (range) of the median income group
      * : Number of units in the specified group necessary from the median income group to reach the middle
      * : frequency or (weighted) total number of units in the median income group
  + Consumer Price Index:
    - The cost of a fixed market basket of goods and services purchased by households.
    - General formula:
  1. **What is the unit of analysis of the indicator/model? (e.g. households, persons, cities)**
  + Province-year
  1. **How can the indicator be integrated with other datasets?**

We can merge CCOFOG and Macroeconomic Indicators with other datasets based on year and geographic unit (i.e., the 10 Canadian provinces).

* 1. **What are the boundaries of the indicator/model?**
* Macroeconomic indicators:
  + Gini coefficient: between 0 and 1
  + Unemployment rate: 0% to 100%
  + Gross Domestic Product:
  + Median after-tax income:
  + Consumer Price Index:
  1. **If you want to add a model to CSDUL, is this associated with a hypothesis? If yes, please describe their:**
     1. **Null hypothesis**
     2. **Alternative hypothesis**
     3. **The implications of rejecting the null hypothesis**
  2. **What is the interpretation of the values of the indicator/model?**

Gini coefficient is a number between 0 and 1 that measures the degree of inequality in the distribution of income. A value of **0** represents perfect equality, where everyone has the same adjusted household income, while a value of **1** indicates maximum inequality, where one person receives all the income and everyone else receives none. Although a single Gini coefficient does not have a simple standalone interpretation, it is useful for comparing income inequality across different populations or over time: **the higher the coefficient, the greater the inequality. Unemployment rate is the percentage of unemployed individuals in the labour force. GDP** measures the total income generated by the production of all final goods and services within a region and time frame at market prices. Median after tax income is the amount that divides the income distribution of that group into two halves, i.e., the incomes of half of the units in that group are below the median, while those of the other half are above the median. CPI measures the cost of a fixed market basket of goods and services purchased by households.

* 1. **Based on the literature and your experience working with this indicator/model, is it possible to identify weaknesses in its calculations or assumptions? To facilitate your answer, you can focus on:**

1. **Potential biases**
2. **Overestimation**
3. **Underestimation**
4. **Omitted variables**
5. **Endogeneity**
6. **Does the indicator/model have other mathematical or computational versions (not syntax) to build it? (provide references)**

[Gini coefficient](https://www.datacamp.com/blog/gini-coefficient):

* Lorenz curve method:
  + - : Area under the Lorenz curve
* Pairwise income gap method:
  + - n: Number of observations
    - and : Unit values (income of person i)
    - : Mean
  1. **Why are you building the indicator/model as you propose? Are there advantages compared to other versions?**

1. **Do you see potential improvements for the indicator/model? This could involve using other datasets, refining calculations, or modifying assumptions, among others.**

**What inputs are to be added to CSDUL? Write “X”**

|  |  |
| --- | --- |
| **X** | Raw or intermediate datasets required to create the indicator/model. |
| **X** | Codes that create the indicator/model (be sure that your code is clear enough to be replicated in the future for yourself or any other researcher). |
| **X** | Documentation that explains step by step the entire process that builds the indicator or model. |
| **X** | Results, which consist of the list of variables, indicators, or model results. |
| **X** | Support files. They can be papers, chapter books, codes, etc. |

**References**

1. Current dollars or the term nominal suggests that the values are expressed in the actual prices of the year they were recorded - not adjusted for inflation. [↑](#footnote-ref-2)
2. Constant dollars or the term real suggests that the values are adjusted for inflation to reflect the purchasing power in a base year. [↑](#footnote-ref-3)