SPSD/M Microsimulation for EI

Regan Meloche - Jan 21, 2021

This post is an overview of the Social Policy Simulation Database/Model tool created by Stats Canada. Official Documentation and description can be found **here**.

Introduction

The SPSD/M is a tool designed to analyze the financial interactions of governments and individuals in Canada. It allows the assessment of cost implications in the cash transfer system. It has two main components: The database and the model. The database consists of a simulated population of Canada imputed from various data sources, and the model is the set of parameters and algorithms that encode the various policies that are modelled. The SPSD/M can be run in the command-line "classic" mode, or the "visual" mode, which includes a graphic user interface. The SPSD/M also has two modes of running. The "blackbox" mode involves simply altering parameters on the pre-built system. This can be used for a wide variety of simulations. If a user needs to add more complex and custom functionality, then they can use "glassbox" mode, which can involve writing some code and compiling your own executable. One feature of the SPSD/M is to model certain simulations related to Employment Insurance (EI). This article will discuss the SPSD/M from the EI perspective.

Data Sources

A microsimulation requires a simulated population, which must be based on reliable data sources. The various data sources can be aggregated and matched together in order to create, or "impute" simulated individuals in a population. The SPSD/M is primarily based on the Canadian Income Survey. This source contains details on income and family, but is lacking info on unemployment. The CIS is further supplemented by Personal Income Tax

Return data, EI claim histories, and the survey of household spending.

Together, these data sources can be used to impute a population that closely models the population of Canada. It is an imprecise model, but as the quote goes: “All models are wrong, but some are useful”. Considerations must be made in order to balance out different goals of the database. The database must be publicly accessible, therefore it cannot contain any “simulated” individuals that could be traced back to an actual individual. When constructing the database, there is a tradeoff between confidentiality and correctness of the data. Another important consideration when imputing data such as this, is to make sure that unrealistic individuals aren’t created, for example an elderly childless couple that is assigned child care expenses.

Setting up an Example

The “Model” component of the SPSD/M contains the coded algorithms (in C++) for calculating the various taxes and transfers, including EI. When a user runs a simulation, all of the simulated individuals from the imputed database are run through these algorithms to calculate a huge list of variables that relate to the individuals and their households.

We will do a brief overview of a simple scenario that involves changing a parameter related to the Maternity Benefits portion of EI. For starters, a quick intro to maternity benefits. This program uses the same calculation as the regular EI benefits, taking a percentage (e.g. 55%) of your average income over the course of the previous year. The average income is calculated by adding together your weekly income over a certain number of “best weeks”. This number depends on the unemployment rate in your region. In a place that has higher unemployment, you will use a lower number for the best weeks. Then you take that sum and average it out using that same number of best weeks. If your income is consistent throughout the year, then you don’t really need to worry about this part. Finally, there is a maximum amount that you can get per

week, as well as a maximum number of weeks for which a mother can claim benefits. More info can be found **here**.

So we've identified a few different "parameters" for this calculation. At the very least, we have the percentage of the average income, the number of best weeks based on unemployment rates, a maximum weekly value, and a maximum number of weeks. These are all components of the policy itself. This is in contrast to "variables", which can be thought of as attributes belonging to an individual or household. Related to maternity benefits, this may include the employment and income history, the date the claim was submitted, where they live, etc.

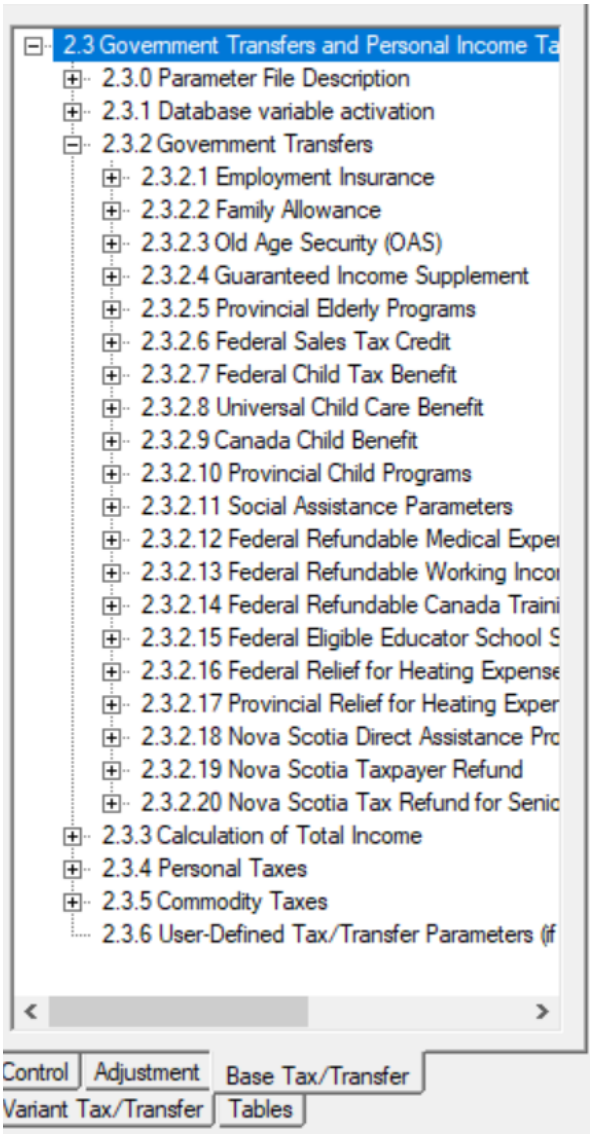
So we have a simulated population imputed from the data sources. Some of these variables are already filled in for our simulated individuals (e.g. province of residence, age, etc.) and some are still unknown. We are going to alter one of the parameters for the Maternity Benefits calculation, and run the simulation twice. The first run will be for the case where there is no change to the policy, and the second will be with the change applied. This will allow us to calculate the difference for our entire population and then draw insights from the aggregate values

Running the Simulation

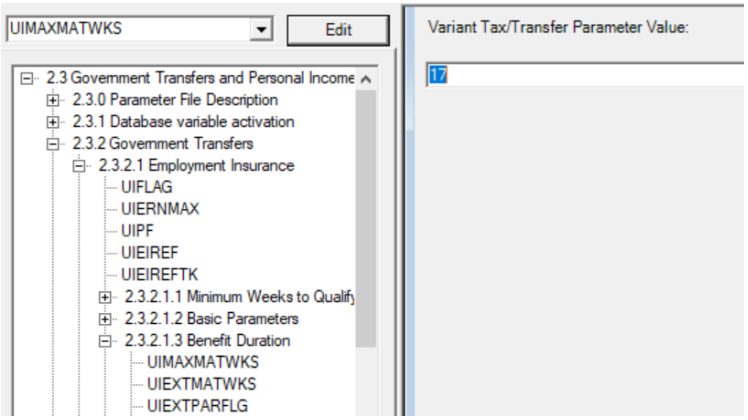
We will now open up the visual SPSD/M and demo the simulation. For this demo, we will simulate the effect of changing the maximum weeks for maternity benefits from 15 to 17. As one might imagine, this would cause people to collect more money from the program. The SPSD/M will allow us to confirm that and will give an estimate of exactly how much more, and we can look at fine-grained amounts across different income groups and provinces.

Inside the program, we can see all of the different parameters that we can change. They are grouped by program, so we will be looking into the

“Government Transfers” section, and then more specifically, the “Employment Insurance” section.



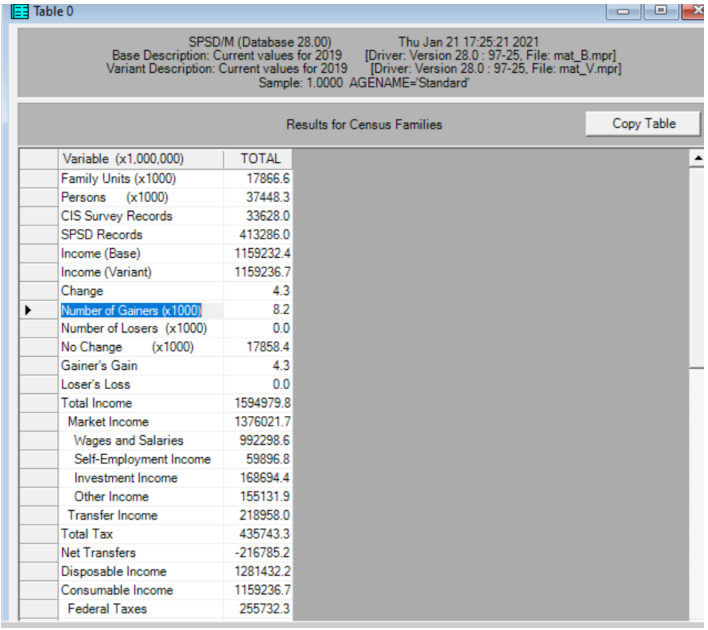
Expanding the EI section, we see the different groups of parameters involved in the calculation. The documentation is also very detailed and searchable, and can be used to find parameters that you want to change. The parameter we are interested in is called ‘UIMAXMATWKS’, and it refers to the maximum weeks for Maternity benefits. The base value for 2019 is 15. We want to see the effect of changing that value to 17, so we will update the variant parameter value from 15 to 17.



There are a few other control parameters and triggers that we need to define. These can be found in the documentation and illustrated by the tutorials that come with the documentation. We will save and run our simulation. The way we've configured it will actually trigger both the base and the variant simulation to be run, so we only need to run it once.

Results

If the simulation runs successfully, the program will give us a message letting us know the simulation is complete. We can then analyze the tables that were generated from the simulation. Two tables with general information are generated by default, and there are facilities for creating your own tables.



Looking at the first default table, we can see the number of gainers/losers that result from the change. As expected, there are only gainers, since we've increased the number of weeks for which new mothers can receive benefits. We can also see the change in cost of the program (4.3 million).

Variable (x1,000,000)	NFLD	PEI	NS	NB	QUE	ONT	MAN	SASK	ALTA	BC	CANADA
Family Units (x1000)	253.0	75.8	483.5	367.5	4225.8	6700.3	597.8	531.5	2019.4	2612.0	17866.6
Persons (x1000)	518.0	157.7	945.1	748.0	8558.8	14547.8	1297.5	1129.0	4443.4	5103.1	37448.3
CIS Survey Records	1433.0	1044.0	1996.0	1867.0	6037.0	7982.0	3346.0	2774.0	3449.0	3700.0	33628.0
SPSD Records	11709.0	4719.0	5971.0	10885.0	34283.0	141009.0	14436.0	15829.0	96664.0	77781.0	413286.0
Income (Base)	14168.2	4090.0	25735.7	19664.8	237335.4	456664.7	37204.6	35442.0	159749.7	169177.1	1159232.4
Income (Variant)	14168.2	4090.1	25735.8	19664.9	237335.4	456667.5	37205.1	35442.1	159749.8	169177.8	1159236.7
Change	0.0	0.0	0.1	0.1	0.0	2.8	0.5	0.0	0.1	0.7	4.3
Number of Gainers (x1000)	0.0	0.1	0.2	0.1	0.0	5.7	0.8	0.1	0.1	1.2	8.2
Number of Losers (x1000)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No Change (x1000)	253.0	75.7	483.4	367.3	4225.8	6694.6	597.0	531.4	2019.3	2610.8	17858.4
Gainer's Gain	0.0	0.0	0.1	0.1	0.0	2.8	0.5	0.0	0.1	0.7	4.3
Loser's Loss	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total Income	20691.7	5777.2	36717.4	27545.9	335106.7	630654.0	51687.5	48401.0	212119.7	226278.6	1594979.8
Market Income	16681.2	4720.6	30642.8	22382.3	278713.9	549541.3	44449.0	42387.9	190572.1	195930.6	1376021.7
Wages and Salaries	13208.7	3382.5	21682.7	16954.2	200597.3	395317.2	32226.6	29710.9	146543.1	132675.3	992298.6
Self-Employment Income	677.6	283.0	1214.5	621.8	13057.2	26782.2	1920.0	2211.6	4814.7	8314.0	59896.8
Investment Income	812.6	412.2	3096.2	1597.3	27270.4	65243.9	5126.2	6086.9	24514.5	34534.1	168694.4
Other Income	1982.2	642.9	4649.4	3208.9	37788.9	62197.9	5176.2	4378.4	14699.8	20407.1	155131.9
Transfer Income	4010.4	1056.6	6074.6	5163.6	56392.9	81112.6	7238.5	6013.1	21547.6	30348.0	218958.0
Total Tax	6523.4	1687.2	10981.6	7881.0	97771.4	173986.6	14482.4	12958.9	52369.9	57100.8	435743.3
Net Transfers	-2513.0	-630.5	-4907.0	-2717.4	-41378.5	-92873.9	-7243.9	-6945.8	-30822.3	-26752.8	-216785.2
Disposable Income	16292.0	4673.0	29264.2	22470.1	266485.9	507168.4	41166.2	39188.5	169384.1	185339.9	1281432.2
Consumable Income	14168.2	4090.1	25735.8	19664.9	237335.4	456667.5	37205.1	35442.1	159749.8	169177.8	1159236.7
Federal Taxes	3447.4	889.6	5812.8	4250.8	46132.3	104478.4	8239.5	7950.0	37696.3	36835.2	255732.3

The second table breaks down the results by province, with the totals for the country on the right-most column. Note that there is no change in Quebec, since that province has a separate system for maternity benefits, which is also captured in the SPSD/M. If we wanted to measure that change, then we could adjust the parameters related to that calculation and run another simulation.

Conclusion

This is just a very simple example of what can be done with the SPSD/M with regards to EI. We will continue to explore the use cases of the SPSD/M as well as search for policy experts and analysts working on EI who may be able to make use of such a tool. If we can identify value from it, then we may integrate it as a key component in a larger Policy Difference Engine system.