Catalogue no. 36-28-0001
ISSN 2563-8955

Economic and Social Reports

Leveraging Statistics Canada data integration opportunities for program evaluation

by Marc Frenette, Winnie Chan and Tomasz Handler

Release date: March 26, 2025

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 infostats@statcan.gc.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

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

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2025

Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Leveraging Statistics Canada data integration opportunities for program evaluation

by Marc Frenette, Winnie Chan and Tomasz Handler

DOI: https://doi.org/10.25318/36280001202500300002-eng

How can Statistics Canada benefit program evaluation?

Government programs are designed to help individuals, families, communities and businesses through grants, loans, subsidies, job training, informational or counselling services, etc. The costs of setting up and running these programs often run well into the millions and, in some cases, billions of dollars. Establishing what sorts of impacts a program has on beneficiaries is crucial for understanding its value (i.e., do the benefits outweigh the costs?), which can then be used to justify the development, continuation, or cessation of the program.

Proposed programs are sometimes pilot tested on a small scale in an experimental framework (e.g. Ford et. 2012). Selected individuals who agree to be part of the pilot test are randomly assigned to a program group (that is offered the program) or a comparison group (that is not offered the program). The outcomes of both groups are then followed over time. The difference in these outcomes provides researchers with an estimate of the program impact—the difference the program made in the lives of those who participated in it.[1] Random assignment more or less ensures that the difference in outcomes can be attributed to program participation rather than pre-existing differences in the characteristics of individuals participating or not in the program.

Although this approach is appealing from a scientific perspective, pilot tests can be very expensive and may take several years or even decades to complete. In many instances, policymakers may wish to address today's problems by implementing the full program as soon as possible. What this means is that in most cases, programs are often evaluated after implementation (ex post) rather than before implementation (ex ante).

Evaluating programs after implementation can be very challenging. Entities responsible for a program often collect information on beneficiaries, such as demographic and socioeconomic background information, details on the services or financial assistance they received, and outcomes following program participation. All of this information is critical as it corresponds to the usual intended objectives of government programs—programs are often intended to be taken up by specific population groups or businesses, which are expected to benefit from it. Unfortunately, data on program participants alone are

1. Since participants in the program group were offered the program (or treatment, in experimental language), as opposed to being forced to receive it, the estimated impact is termed an "intention-to-treat effect". When comparing individuals who received the program or not, the estimated impact is termed a "treatment effect on the treated". Both treatment effects are useful as the former (the intention-to-treat effect) indicates to the policymaker the impact of offering the program to participants (i.e. the action that a policymaker can take), while the latter (the treatment effect on the treated) informs potential program participants of the impact of taking up the program.

not sufficient to conduct a rigorous program evaluation because it is usually not possible to make comparisons with people or businesses that did not participate in the program. How can policymakers know whether a program improved the lives of its participants when the outcomes of (otherwise similar) non-participants are not known? Likewise, determining whether the program was well targeted may be challenging with no information on non-participants. Another issue arises with participant follow-up—many programs may benefit participants several years later, but measuring the outcomes of participants for several years often requires follow-up surveys that may suffer from low response rates and high attrition rates. Both of these factors limit the credibility of the survey instrument and the program evaluation as a whole. Finally, it is also beneficial to track outcomes of participants and non-participants prior to program implementation, as it provides a useful reference point for comparing to post-program trends.

Through a user-centric and comprehensive suite of data and analytical services, Statistics Canada can assist government departments and others with program evaluation. Data integration involves combining multiple administrative and survey data sources together through microdata record linkage (e.g., person-to-person or business-to-business matching across datasets). This approach yields a more comprehensive set of information than program data alone and thus enables more rigorous program evaluation. The integrated data may include individuals who participated in a program (through the program data file) and those who did not (from the broader Statistics Canada data files, which include participants and non-participants). Moreover, participants and non-participants can be followed for several years with very little attrition, since many Statistics Canada data files, particularly longitudinal administrative data files, are available annually with comprehensive coverage rates. As a result, it may be possible to compare the outcomes of program participants with those of non-participants for many years before or after their participation.

Microdata linkage is conducted in accordance with Statistics Canada's Directive on Microdata Linkage, which has been in place since 1986.[2] As part of its governance over microdata linkages, Statistics Canada has pre-approved specific types of linkages. The linkages involved are those where the privacy risks and situations of potential conflict of interest are low and where procedures to mitigate risk to confidentiality and privacy are in place. All other microdata linkages must undergo a prescribed review and approval process, which involves the submission of documented proposals to senior management. After receiving approval to linkage, source data will be processed further to remove all personal, business and other identifiable information. The resulting linked data cannot be used to make decisions on any specific individual.[3]

Statistics Canada data and analysis in action

One recent example of this approach concerns the Canada Summer Jobs (CSJ) program (OAG 2024).[4,5,6] The CSJ is a program under the Youth Employment and Skills Strategy and is geared towards youth aged 15 to 30. The CSJ program aims to provide high-quality work experiences for youth, respond

2. Please visit https://www.statcan.gc.ca/en/record/policy4-1 for Statistics Canada's Directive on Microdata Linkage.
3. For more information, please visit https://www.statcan.gc.ca/en/record/gen.
4. For more details on the CSJ program, please consult the guide by Employment and Social Development Canada (ESDC 2024).
5. The program evaluation was initiated and conducted by the Office of the Auditor General as part of an audit, with Statistics Canada assisting in the process. However, program areas may contact Statistics Canada proactively if they require assistance in evaluating their programs.
6. This project was funded by the Office of the Auditor General (OAG) and was a collaboration between the OAG, Employment and Social Development Canada, and Statistics Canada (the Social Analysis and Modelling Division and the Centre for Income and Socioeconomic Well-being Statistics).

to national and local priorities to improve access to the labour market for youth who face unique barriers, and provide opportunities for youth to develop and improve their skills.

Program data from the CSJ were integrated with various Statistics Canada data files, including the 2016 Census of Population, several years of the Postsecondary Student Information System and the Longitudinal Worker File. Adding these files to the program data was critical to the success of the evaluation, since they enabled the creation of a comparison group of observationally similar individuals who did not participate in the program. It was also possible to measure various program outcomes and follow participants and non-participants for many years (with very little sample attrition) using longitudinal administrative data housed at Statistics Canada.[7]

The analysis performed on these files revealed two key findings. First, CSJ participants generally registered more favourable labour market outcomes than non-participants. Indeed, the median earnings of program participants were greater than those of the non-participant comparison group, nine years following participation. However, the second key finding was that the CSJ program did not always reach the most vulnerable youth. For example, compared with non-participants, CSJ participants were less likely to report having a disability and more likely to report being White. Moreover, program participants were equally likely as non-participants to report an Indigenous identity.

Key factors for success

The data integration between program files and Statistics Canada data files was crucial to the success of the CSJ program evaluation, but what factors enable successful program evaluation by leveraging data integration? The following sequential steps may be helpful[8]:

- Determine what questions need to be answered
    - This will guide data collection and data linkage needs.

- Determine the data and methodological requirements
    - The ultimate goal should be to identify data (both from the program, as well additional data from Statistics Canada) and methods to enable a robust program evaluation on key target groups and sub-groups of individuals. Understanding who is taking up the program is also important.
    - Subject matter expertise at Statistics Canada will be able to provide more information related to the subject area and data requirements for the program team.
    - It is particularly important to speak with Statistics Canada prior to collecting program data since data linkage is only possible only with high-quality linkage keys (individual identifiers).

- Develop an evaluation data plan
    - This step may require a significant amount of time for discussion and approvals. Statistics Canada operates on a cost-recovery basis. As a result, programs will also need to budget for linkage and related activities, which may include costing, contracting, data acquisition, record linkage approval, the creation of anonymized linkage keys, record linkage implementation and validation of linked data, derivation of vetting rules by methodology

7. Comparing outcomes of the CSJ participants with the non-participants yielded estimates of treatment effects on the treated. It was not possible to estimate intention-to-treat effects since no information was available on program offers.
8. The list draws heavily from TBS (2025), a guide written specifically for Federal government departments, and is adapted for the purposes of this article.

services, and data access arrangements (if analysts in the program area would like to use the microdata directly) or analysis (if the program area would prefer that Statistics Canada generate custom tabulations, perform more complex analyses or even write a report). To get started, contact Statistics Canada early in the process.

- Identify and address challenges
    - Programs should take stock of obstacles that the program faces in the implementation of a linking-based data initiative, including partnership agreements or information technology limitations. Plans should be developed to address identified obstacles.
    - The challenges might impact what the evaluation plan is going to be and a discussion with Statistics Canada can shed light on alternative strategies or options that the program teams can put in place to address the issues or limitations.

- Collect and share the data with Statistics Canada
    - Coordinate with Statistics Canada to deliver any data files. These files should contain only the relevant information required for the linkage.
    - As noted earlier, there is a data acquisition process at Statistics Canada that should be factored into timelines. The length of this process can depend on the number of files and level of data sensitivity.

- Work with Statistics Canada to produce the analysis
    - Statistics Canada has considerable experience in working with linked micro-data files, and conducting complex analysis to support program evaluation.
    - Programs that wish to access the microdata to conduct the analysis can leverage various data access arrangements with Statistics Canada. These arrangements may already be in place with the sponsoring department. Members of the program team will need to be deemed as having Statistics Canada employee status in order to work directly with the data.

Authors

Marc Frenette, Winnie Chan and Tomasz Handler are with the Social Analysis and Modelling Division, Analytical Studies and Modelling Branch (ASMB), at Statistics Canada. The ASMB is responsible for supporting a range of data linkage environments (that is, the Canadian Employer-Employee Dynamics Database) and the use of existing linked data files such as the Longitudinal Worker File, the Intergenerational Income Database (IID) and the IID-Census to support analytical activities for program development and evaluation. ASMB provides expert advice to policy departments on the development of data and analytical plans to support program development and evaluation as well as training in quantitative evaluation methods. The Branch is also responsible for providing access to microdata for those departments wishing to conduct analysis.

References

ESDC [Employment and Social Development Canada]. 2024. *Canada Summer Jobs 2024: Providing Youth with Quality Work Experiences (Applicant Guide).* Employment and Social Development Canada. https://publications.gc.ca/collections/collection_2024/servcan/SG2-10-2024-eng.pdf.

Ford, R., Frenette, M., Nicholson, C., Kwakye, I., Hui, T.S., Hutchison, J., Dobrer, S., Smith Fowler, H., Hébert, S. 2012. *Future to Discover (FTD) - Post-Secondary Impacts Report.* Social Research and Demonstration Corporation. https://www.srdc.org/project/Future-to-Discover-FTD--Post-secondary-Impacts-Report/.

OAG [Office of the Auditor General]. 2024. *Canada Summer Jobs—Employment and Social Development Canada.* 2024 Reports 8 to 12 of the Auditor General of Canada to the Parliament of Canada. https://www.oag-bvg.gc.ca/internet/English/parl_oag_202412_12_e_44595.html.

TBS [Treasury Board of Canada Secretariat]. 2025. *Data Linking for Program Monitoring, Evaluation and Reporting.* Treasury Board of Canada Secretariat. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32808.