

Education and Labour Market Longitudinal Platform 2.0: Exponential Potential

Key Findings

- New administrative data on apprentice, college and university student loans and grants are now available in the Education and Labour Market Longitudinal Platform (ELMLP). These data complement existing information on earnings, education/training and apprenticeships.
- In 2020, the 2018 National Graduate Survey will be available in the platform, and the 2016 Census will be linked sometime thereafter. Both datasets will offer a wealth of new details on apprentice, college and university graduates' sociodemographic characteristics and occupational information.
- Additional tax records from the Longitudinal Administrative Databank will be linked to the ELMLP, which will help identify workers who did not attend post-secondary education in Canada and could therefore serve as a comparison group.

Introduction

In December 2018, Statistics Canada released **the Education and Labour Market Longitudinal Platform (ELMLP)** — a revolutionary data environment that enables researchers and policy-makers to study the earnings trajectories of people who attended publicly funded colleges and universities or participated in registered apprenticeship programs in Canada. Researchers from across the country have already started to dig into this expansive trove of administrative data. Notably, **Statistics Canada** has released initial findings on the earnings of college and university graduates. For our part, LMIC has partnered with the **Education and Policy Research Initiative** to further explore the platform and investigate the earnings trajectories of

graduates across credentials, sociodemographic characteristics and fields of study.

LMI Insights no. 4 noted that while the ELMLP environment offers many exciting possibilities for research and policy analysis, there are still some noteworthy limitations. Three of the noted limitations—information on financial assistance programs, occupations and a potential earnings comparison group—are in the process of being resolved, thanks to new datasets that will be linked with the core elements of the ELMLP (namely, tax-linked records of apprentices, college and university students). This issue of *Insight* provides an overview of the newly available and forthcoming data linkages in the new and improved ELMLP 2.0.

Additions to the Education and Labour Market Longitudinal Platform

New datasets related to student and apprentice loans and grants are now available for linkage with the core ELMLP datasets (see [Table 1](#)). These new data directly address the lack of observations on student assistance programs available to study the role of such programs. Two other previously noted limitations—the lack of an earnings comparison group and occupational information—will start

to be resolved in 2020 with linkages to the Longitudinal Administrative Databank, the 2016 Census and the 2018 National Graduate Survey. These additions to the ELMLP data environment are discussed in greater detail below, along with how they can help address the data limitations associated with the core ELMLP datasets.

Table 1. Datasets available through versions 1.0 and 2.0 of the Education and Labour Market Longitudinal Platform

Dataset	Time Coverage	Description
Core datasets		
Postsecondary Student Information System (PSIS)	2009-10 to 2015-16	Student records, including field of study, program type and graduation year
Registered Apprenticeship Information System (RAIS)	2008 to 2016	Apprentice records from registered programs, including type of program, current status and location
T1 Family File of annual income	2004 to 2016	Tax files of people identified in PSIS and RAIS
New datasets now available upon special request		
Apprenticeship Grants	2007 to 2017	Records of taxable cash grant for individuals in a designated Red Seal apprenticeship ¹
Canadian Apprenticeship Loans (CAL)	2015 to 2017	Information only on approved CAL disbursements. The data include one disbursement per period of technical training for each person, up to 5 training periods.
Canadian Student Loan Program	2005-2006 to 2015-2016	Longitudinal information on loan recipients by loan year (August 1 to July 31) through all phases of the loan or grant lifecycle
Datasets to be added in 2020		
National Graduate Survey	2018	Provides cross-sectional records of all graduates from the 2015 calendar year. The survey contains information about graduates' job satisfaction and the type of employment obtained.
Potential datasets to be added in 2020 and beyond		
Canadian Education Saving Program	2005 to 2016	Records of all recipients of the Canada Education Savings Grant and the Canada Learning Bond through all phases of pre- and post-secondary education.
Long-form Census	2016	Detailed sociodemographic information on 25% of Canadian households
Longitudinal Administrative Databank	2004 to 2016	Random sample of 20% of all T1 Family Files

Note: Full information for each dataset can be found at [Research Data Centres website of Statistics Canada](#).

¹ The Red Seal program sets common standards to assess the skills of tradespeople across Canada. Tradespersons who have successfully passed the Red Seal examination receive a Red Seal endorsement on their provincial/territorial trade certificate.

Analyzing Financial Support for Students and Apprentices

New and important questions can be explored by linking the Apprenticeship Grants (AG), Canadian Apprenticeship Loans (CAL) and Canadian Student Loan Program (CSLP) to the core ELMLP data environment. A fourth related dataset, the Canadian Education Saving Program (CESP), will also be available soon for linkage. Three examples discussed below help illustrate how these new datasets can be used to answer important policy questions.

- 1. Core datasets of *ELMLP + AG + CAL*:**
The CAL dataset contains records of loans disbursed to apprentices registered in a designated Red Seal trade apprenticeship beginning in 2015, whereas the AG dataset contains records of cash grants given to individuals who have completed either the first or second year of the Red Seal program since 2007. To date, little is known about how, or if, such loans and grants affect apprentices' completion rates and labour market outcomes.
- 2. Core datasets of *ELMLP + CSLP*²**
The Canadian Student Loan Program was launched in 1964 and is the longest-running student loan program in Canada. The CSLP administrative dataset contains detailed records of financial assistance provided to students who demonstrate financial need. By linking the CSLP to the core ELMLP datasets, researchers can examine how the student loan acceptance rate has changed in the face of rising tuition fees, identify the factors associated with student loan defaults, and compare the labour market earnings of those who receive student loans versus those who do not.
- 3. Core datasets of *ELMLP + CESP***
The Canadian Education Saving Program provides tax incentives to encourage parents

to save for a child's post-secondary education, including apprenticeship programs. The program provides administrative data records that allow researchers to study its uptake rate. Linkage to the ELMLP will provide, for the first time, empirical evidence of the labour market outcomes of children who benefit from the CESP.

Generating a Comparison Group

Contextualizing the level of earnings of apprentices and college and university students is a challenging but necessary undertaking. An important way to do this is to have a group that allows for "apple to apple" comparisons. In [Insight no. 4](#), we noted that a key limitation is the lack of a baseline earnings group, and suggested the ideal group would be the set of comparable individuals who have never participated in an apprenticeship or a college or university program. Such a comparison will eventually be possible when the Longitudinal Administrative Databank (LAD) is made available within the ELMLP data environment.

Once integrated into the ELMLP, a comparison group can be created through a reverse linkage. As Ted MacDonald described in an [LMIC blog post](#), a reverse linkage identifies the individuals in one dataset (i.e., the LAD) who are not present in another (i.e., the core ELMLP data).

While the integration of LAD represents a big and important step forward, the reverse-linked comparison group will come with many important caveats. First, since the LAD is a random 20% sample of all tax filers in Canada, this linkage will not provide observations on all individuals not recorded in the ELMLP. Secondly, this approach is limited to those who have studied at a college or university in Canada since 2009–2010 or have participated in a registered apprenticeship program since 2008. In other words, the set of

² Records include information on full- and part-time students at designated colleges, universities and institutions across Canada, except for Quebec, Nunavut and the Northwest Territories.

individuals identified through a reverse linkage of the LAD with the ELMLP data will only include individuals who studied abroad or completed their education before 2009. Despite these caveats, LAD data linked to the ELMLP will facilitate longitudinal observations of labour market earnings of a comparison group that reasonably approximates those who did not participate in college, university or apprenticeship programs.

Identifying Occupational Information

Finally, the forthcoming linkage (planned for early 2020) between the 2018 National Graduate Survey and core ELMLP datasets will provide, among other details, occupational information about graduates from the 2015 calendar year. This data linkage will provide much-needed information about graduates' types of employment, job satisfaction and the qualifications required in their current positions. However, the survey includes approximately 60,000 respondents (approximately 13% of graduates), which will limit the number of ELMLP student records that can be enhanced with details from this survey.

The planned linkage between the core ELMLP datasets and the 2016 long-form Census in late 2020 will offer even more granular details related to sociodemographic information and the type of employment held. Moreover, given the size of the 2016 Census, roughly 25% of individuals recorded in the ELMLP will be associated with this granular

information.³ The 2016 Census also provides information on a wide range of time-invariant information—such as country of birth, ethnicity and mother tongue—that can be used to analyze the role of sociodemographic factors in education and labour market outcomes.

Both the National Graduate Survey and Census are conducted every five years. As a result, the granular information will be available at one point in time due to the cross-sectional nature of the dataset.

The Way Forward

The ELMLP has continued to evolve since its release in December 2018. The onus is now on policy-makers and researchers, including the team at LMIC, to take full advantage of the exponential potential of the ELMLP data environment. Our collective aim is to provide all stakeholders—students, parents, institutions, policy makers, researchers and the broader Canadian public—with reliable insights so they can make the most informed decisions possible.

Overall, we hope that analyses of the ELMLP data will help bring clarity to Canadians navigating the world of work. To that end, LMIC will soon launch a series of in-depth analyses in partnership with the **Education Policy Research Initiative** and other stakeholders to paint a more complete picture of the role that education plays in the working lives of Canadians.

Acknowledgments

This issue of LMI Insight was prepared by Young Jung of LMIC. We would like to thank Sylvie Gauthier (Tourism and Centre for Education Statistics Division) and Vince Dale (Centre of Labour Market Information) at Statistics Canada for their comments and suggestions. For more information about this edition of *LMI insights* or other LMIC activities, please contact Young Jung, Economist, at young.jung@lmic-cimt.ca or Tony Bonen, Director of Research, Data and Analytics, at tony.bonen@lmic-cimt.ca.

³ The 2016 long-form Census was filled out by 25% of Canadian households. Therefore, roughly 25% of individuals in the core ELMLP files could be linked to detailed Census data.