**Request for Proposals**

**Research Data Centre Program**

**Statistics Canada**

**Longitudinal Immigration Database (IMDB)**

**Fall 2017**

The Longitudinal Immigration Database (IMDB) combines linked administrative immigration and tax data files. It is a comprehensive source of data on the socio-economic outcomes of the immigrant tax filer population in Canada.

The database is managed by Statistics Canada on behalf of a federal-provincial consortium led by Immigration, Refugees and Citizenship Canada (IRCC). The IMDB provides detailed and reliable information on the labour market behaviour of different categories of immigrants over a period that is long enough to assess the impact of characteristics at admission, such as education and knowledge of French or English. The database also provides information on pre-admission work or study experience in Canada, provincial mobility and family composition. The benefits of analysis using the IMDB included:

* Support for evidence based policy making;
* Increase the use of administrative data sources for research, thereby reducing the cost of data collection and the burden on respondents for surveys;
* Allow inferential statistical analysis on the confidential microdata and analytical work complex in nature or not suitable for other forms of data access.

The decision to conduct a pilot project at the Research Data Center (RDC) aims to address the feasibility of supporting the extended use of the IMDB. This will allow further testing of the confidentiality vetting rules and testing of the IT resources necessary to transfer and analyze these large data files.

**The Data**

[Longitudinal Immigration Database (IMDB)](http://www23.statcan.gc.ca:81/imdb/p2SV.pl?Function=getSurvey&SDDS=5057&lang=en&db=imdb&adm=8&dis=2&fpv=30000)

A person is included in the Longitudinal Immigration Database if he or she obtained permanent resident status since 1980 and filed at least one tax return since 1982. This survey is a census with a longitudinal design where data are collected for all units of the target population, therefore no sampling is done.

The data are combined from administrative files through exact matching record linkage techniques. The IMDB brings together immigration from IRCC, taxation data from the Canada Revenue Agency ([Annual Income Estimates for Census Families and Individuals –T1 Family files](http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4105&lang=en&db=imdb&adm=8&dis=2)), and the date of death from Statistics Canada’s Amalgamated Mortality Database (AMDB). Each year the IMDB is updated with new immigrant cohorts, their non-permanent resident information, and new taxation data. Individuals admitted in previous years may be added later on if they subsequently linked to a tax record.

Researchers unfamiliar with administrative data are cautioned that the IMDB requires the manipulation of multiple large data files via linkages. Consequently, many researchers have found it takes some time to become familiar with the IMDB and to be able to operationalize it in their research.

**Additional record linkages**

The linking of records from additional data sources can be a useful and cost-efficient technique in the design, production, analysis and evaluation of statistical data. Please note that in accordance with the Statistics Canada Directive on Microdata Linkage, the IMDB has been linked to the following dataset available to researchers in the RDC:

* [Longitudinal Survey of Immigrants to Canada](http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4422) (LSIC)

Researchers interested in the LSIC-IMDB should apply to the pilot project using the linked data.

**Submissions**

While a limited number of proposals for both cross sectional and longitudinal analyses will be considered, research that includes the following types of analyses or results is of particular interest:

* Proposals making use of the longitudinal aspects of the database
* Proposal making use of smaller geographical areas
* A mix of SAS users and STATA users.

Researchers are invited to submit proposals for consideration by **December 22, 2017**. The submitted research proposals will be assessed based on the research areas of interest as well as the proposed types of analysis, including the viability of the proposed research. The proposals that are better aligned with the above described research areas of interest, types of analysis and results will receive more favorable consideration.

All researchers will be notified by **early 2018**. The researchers whose proposals have been accepted *should* be able to access the data by early 2018.

Researchers with approved projects are expected to:

* Attend user group meetings as required
* Provide feedback on the data and documentation
* Be prepared that vetting rules are being tested and vetting is being done by committee, and therefore, the release of output may take longer than non-pilot projects.

We will be testing the robustness of the vetting rules, over the next year, with these proposals. Therefore, vetting could be delayed. Researchers should bear this in mind when considering the appropriateness of applying for access to IMDB, at this time, because of the potential impact on the timely completion of their research.

Please refer to the [IMDB technical report](http://www5.statcan.gc.ca/olc-cel/olc.action?objId=11-633-X2017007&objType=46&lang=en&limit=0) .

Please contact Mustafa Ornek (rdc2@mcmaster) for questions related to proposal development and submission. Please contact Dong Shen (qrdc@queensu.ca) for data related questions.

**Proposals should be submitted by December 22, 2017 to:**

**Lisa Oliver, Regional Manager**

**Research Data Centre Program**

[**lisa.oliver@canada.ca**](file:///%5C%5Cstchsfsb%5Colivlis%24%5CRegional%20Manager%5CData_unit%5CIMDB_pilot%5Ctosend%5Clisa.oliver%40canada.ca)

**Proposals received after this date will be considered as long as space permits.**