



Statistics
Canada

Statistique
Canada

2016 Census of Population of Canada: Integration of immigration administrative data

www.statcan.gc.ca



Telling Canada's
story in numbers

Scott McLeish

Social and Aboriginal Statistics Division

UNECE Work Session on Migration
Statistics

October 30-31, 2017

Canada 



Presentation overview

- Census immigration data
- Administrative data
- 2016 Census developments
 - Record linkages
 - Improved processing
 - New admission category variables
- Summary and next steps



The Census of Population is an important source of information about immigrants in Canada

- The Canadian Census of Population collects several variables related to immigration through the questionnaire:
 - Immigrant status and year of immigration
 - Place of birth
 - Citizenship
 - Generation status (based on place of birth of parents)
- In addition, several socio-economic topics are covered to permit detailed analysis of immigrants in Canada
 - Demographics
 - Income, labour and education
 - Housing, households and families
 - Languages

Census processing

- Consistency edits and donor imputation are performed to address missing, invalid or inconsistent values
- Missing, invalid or inconsistent values could be the result of different factors:
 - Respondents may not understand the question (e.g. year of immigration vs. year of arrival)
 - Respondents may not remember their exact year of immigration
 - Proxy responses (the person filling out the questionnaire may not know the true values for everyone in the household)
 - Ambiguous responses which could refer to multiple values (e.g. 'Ireland')
 - Changing boundaries and place names

Donor imputation methods are used for missing values

- Goal: to find the nearest neighbour (donor)
- How do we define the “nearest neighbour”?
 - Determine a list of variables that are correlated with the missing variable (matching variables)
 - Languages
 - Age
 - Completed immigration questions
 - Other ethnocultural variables
 - Geography
 - Determine the relative importance (weight) of the matching variables
 - Calculate the dissimilarity between the record to be imputed and the potential donor records using the matching variables
- For most census immigration questions, less than 2% of records require imputation



Administrative data provides detailed information on the flow of immigrants to Canada

- Immigration, Refugees, and Citizenship Canada (IRCC) collects data on all immigrants when they are admitted to Canada
 - Detailed information on admissions since 1980
 - Includes variables also collected on the census questionnaire:
 - Immigrant status and year of immigration
 - Place of birth
 - Includes additional details on conditions at time of admission
 - Immigrant admission category (e.g. economic immigrant, refugee, etc.)
 - Intended occupation, knowledge of official languages, previous experience in Canada
 - No updates are made and data does not account for deaths or emigration
- This data is linked with annual tax files to create the Longitudinal Immigration Database (IMDB)

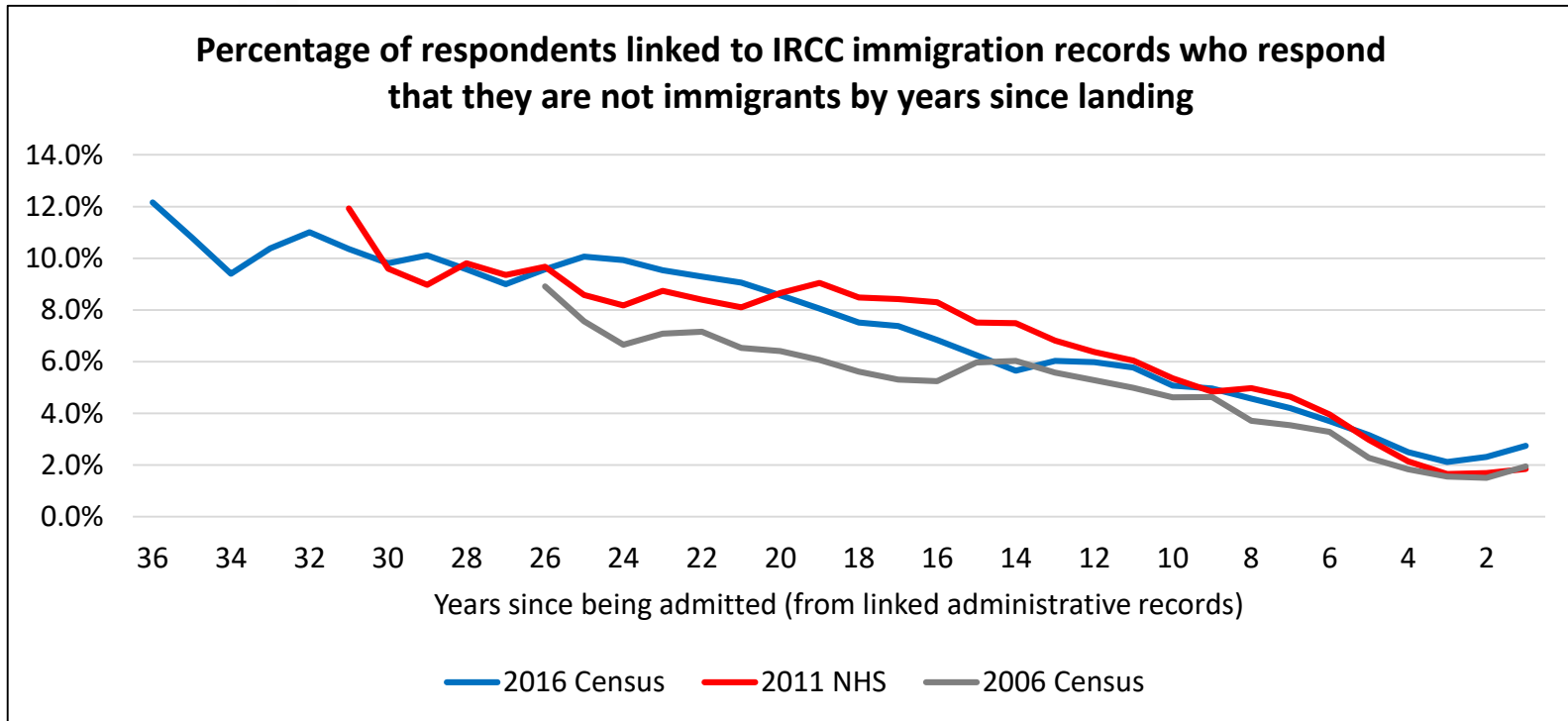


2016 Census: Record linkage to combine both data sources

- Similar to linkages done in 2006 and 2011, the 2016 Census was linked with the administrative data:
 - Immigrants from 1980 to 2016
 - Immigrants from 1952 to 1979
 - Less detailed records
 - Temporary residents
 - Temporary foreign workers
 - International students
 - Asylum claimants
- About 90% of immigrants who landed from 1980 to 2016 were linked
 - Estimated 0.3% false positives



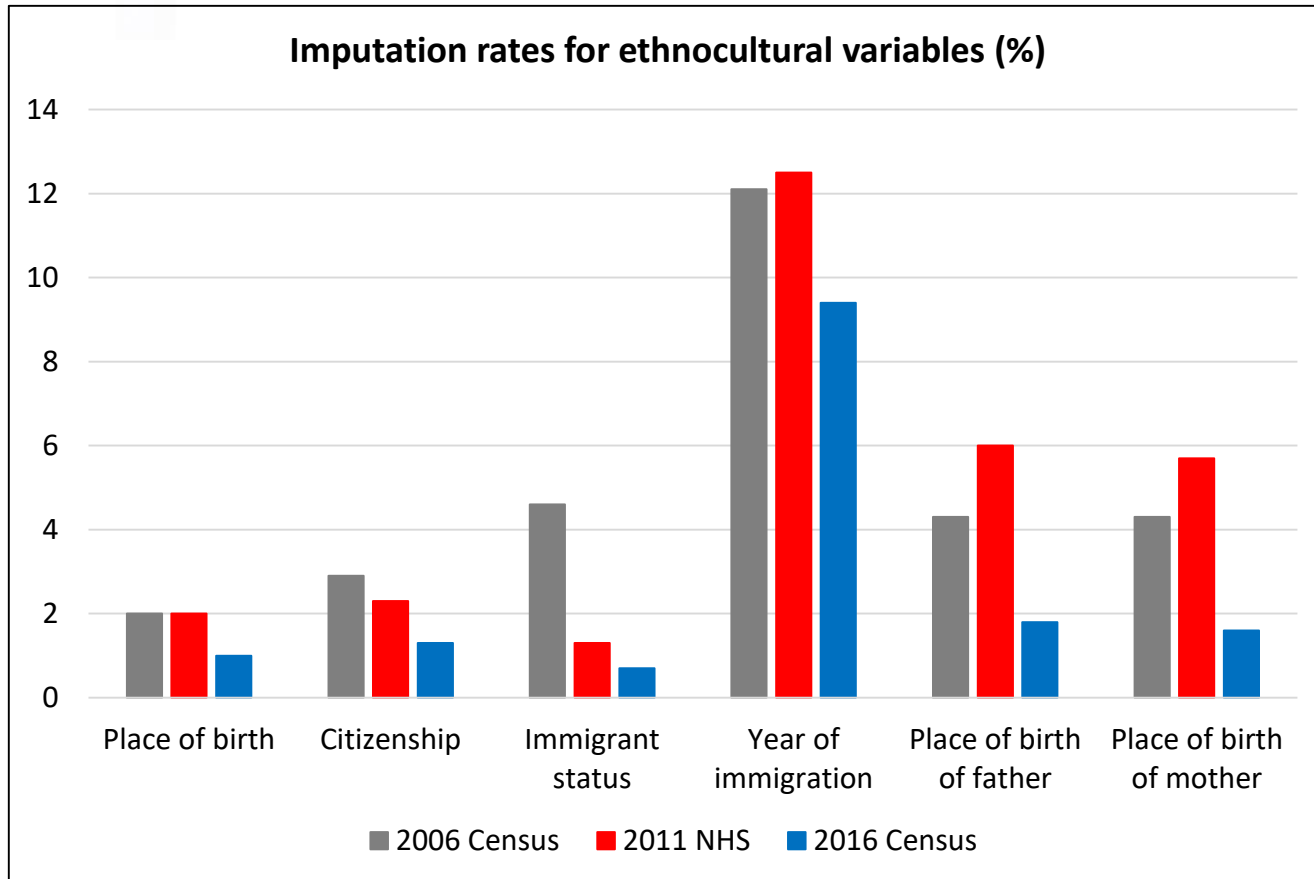
Administrative data illustrated effects of inconsistent responses and corresponding consistency edits



7.6% of respondents who were linked to administrative immigration records responded “No” to the immigrant status question. This proportion increases the longer they have been in Canada. These cases are largely corrected during processing based on other responses (e.g. reporting going through the immigration process to become Canadian citizens).

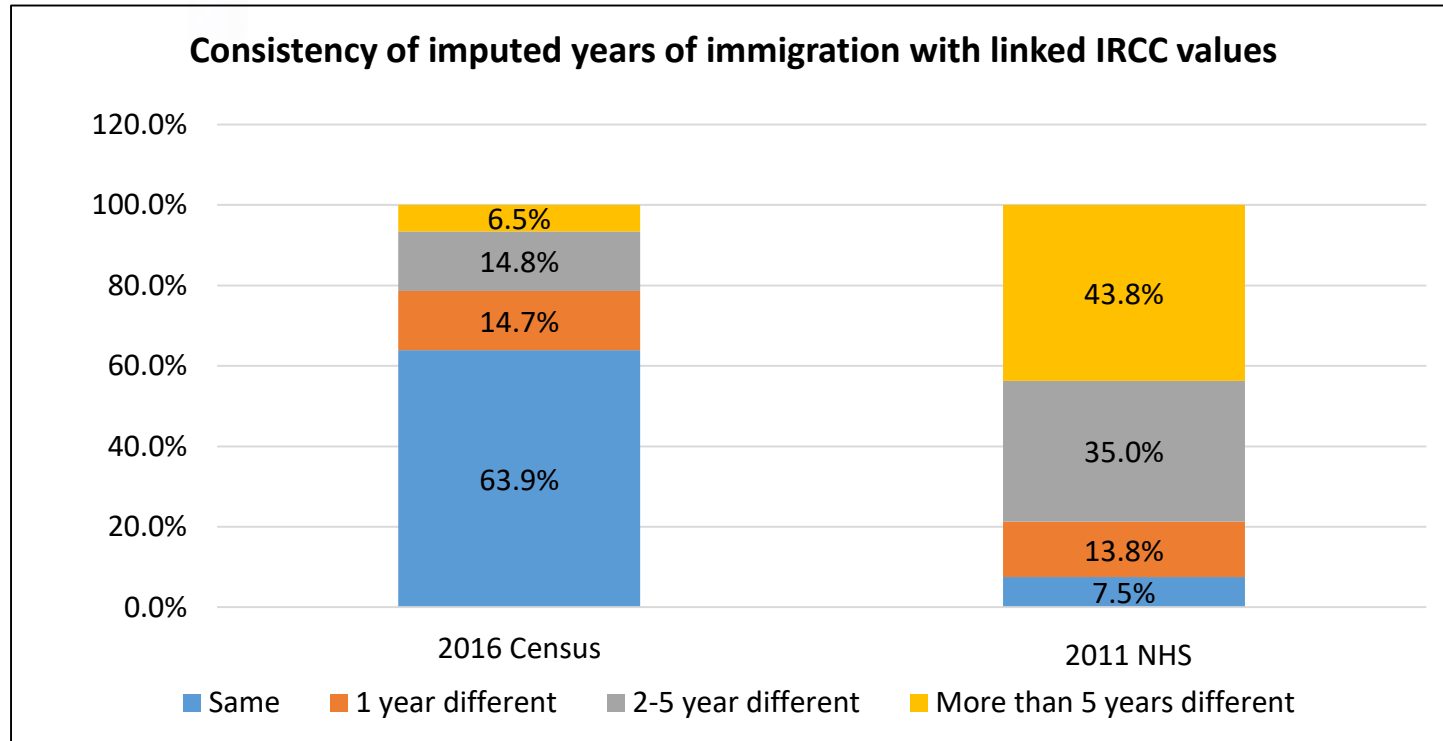


Year of immigration requires more imputation due to consistency edit for immigrant status question





Administrative data improved donor selection by introducing strongly correlated matching variables



63.9% of imputed 2016 Census year of immigration values matched their linked IRCC values. This compares with only 7.5% from the 2011 NHS.

Note: The use of IRCC data in donor selection necessarily ensures that the imputed values will be more similar to IRCC values. However, these were not imputed deterministically - the linked IRCC values were only used to select a donor.

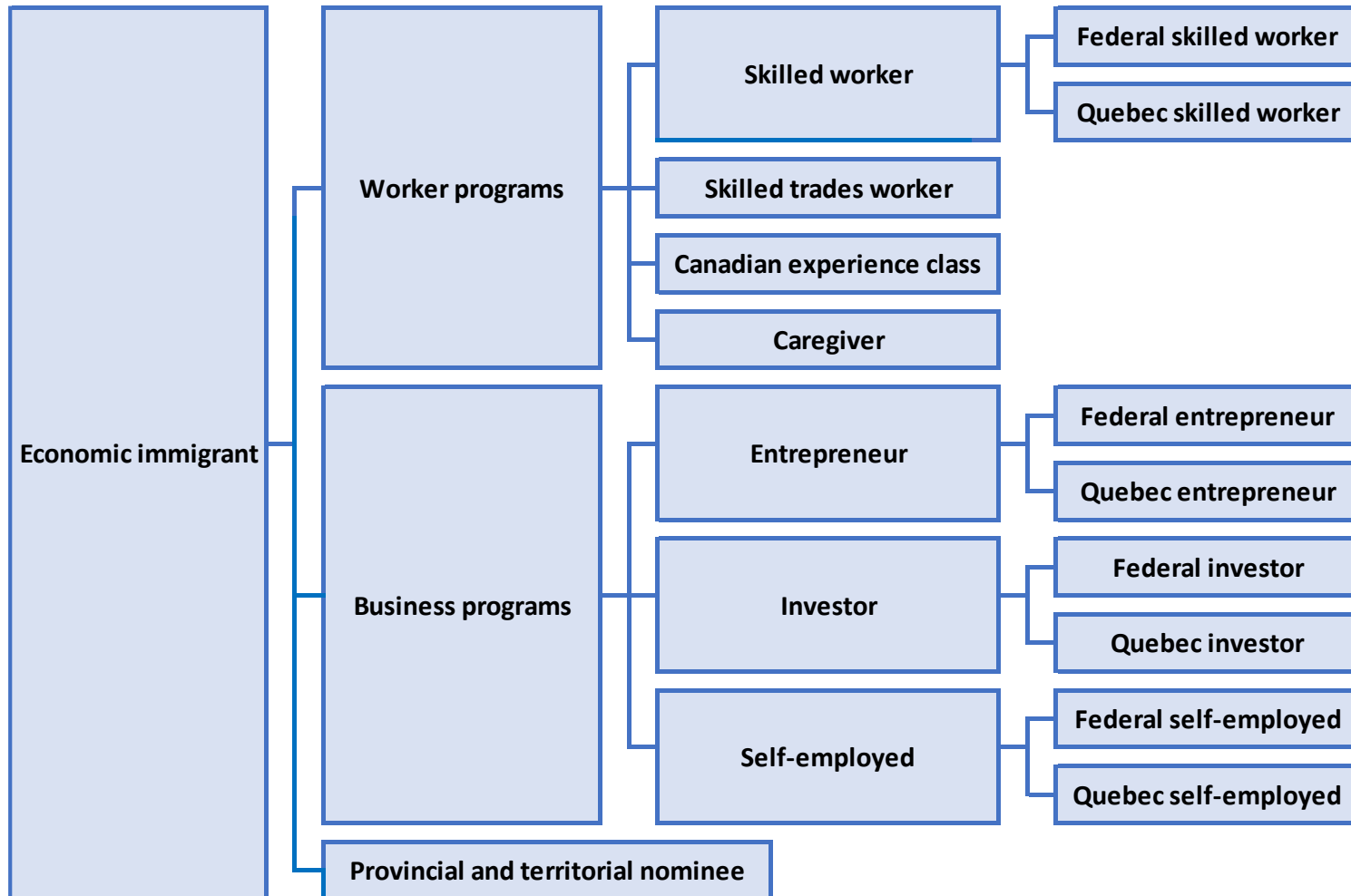


Administrative data provided new analytical variables on immigrant admission category

- In collaboration with IRCC, new variables were added to the 2016 Census regarding admission category of immigrants
 - Immigrants are admitted into Canada for different reasons:
 - Economic immigrants
 - Sponsored family members
 - Refugees
 - Analysis can connect the conditions under which immigrants were admitted with socio-economic outcomes from the census

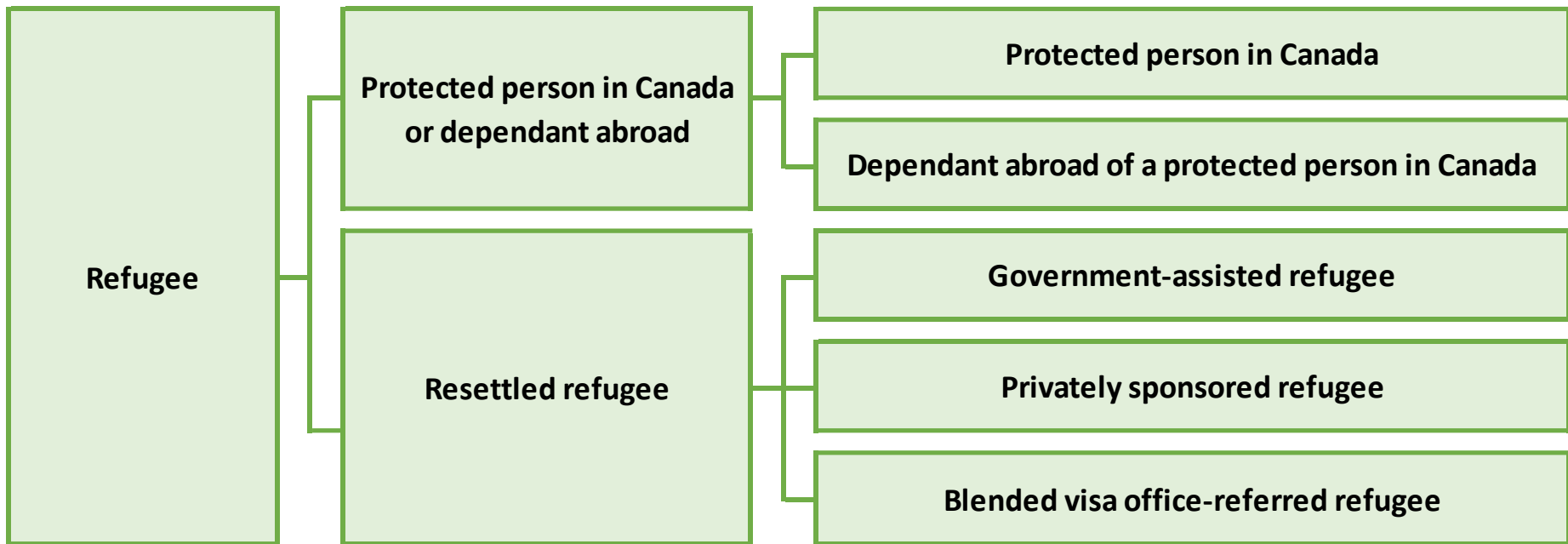


Economic immigrant categories





Refugee categories





Similar to other census variables, consistency edits and donor imputation methods were developed

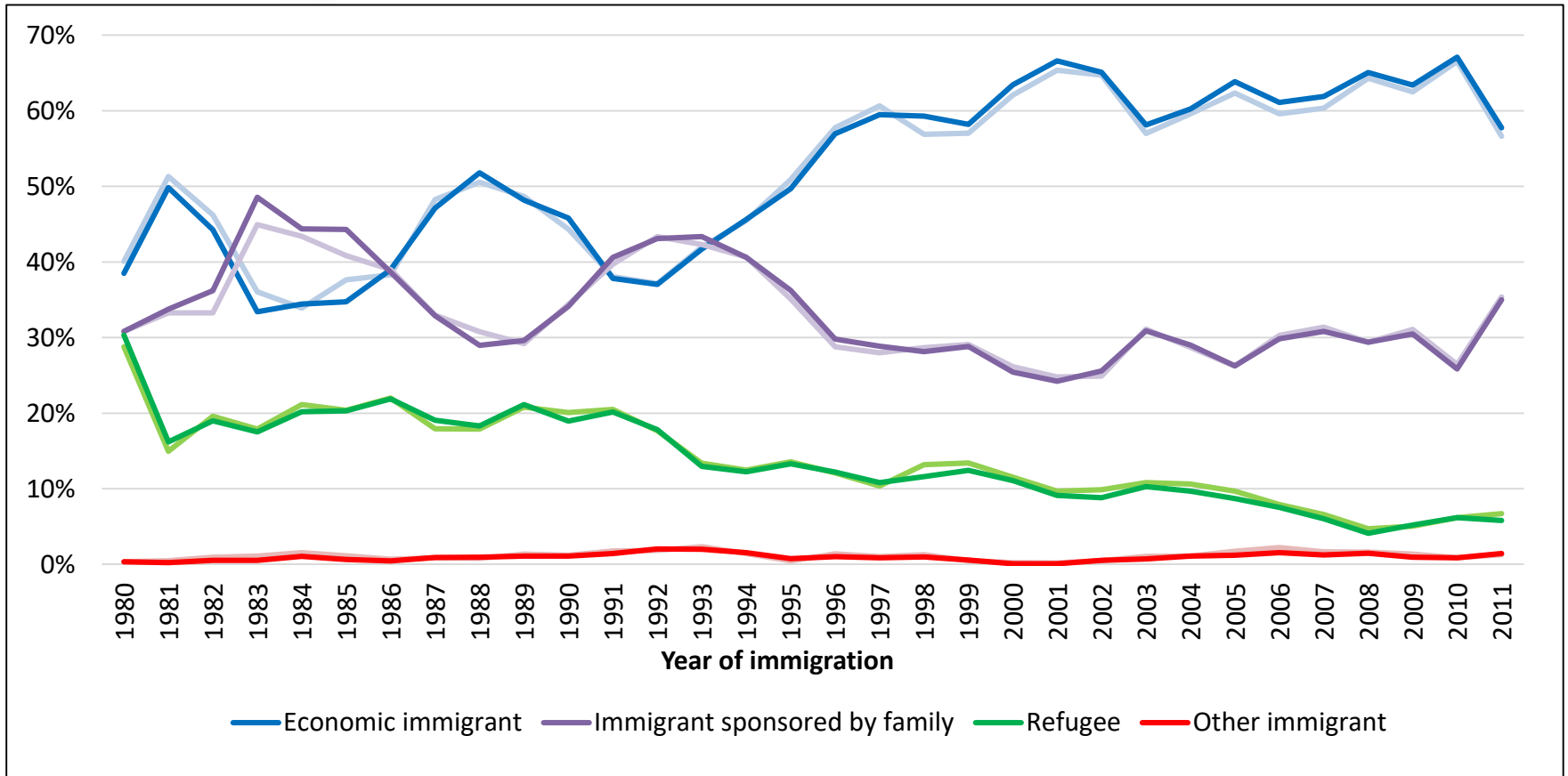
- Imputation was stratified based on family characteristics
 - Couples (especially with matching years of immigration)
 - Children and parents
- ReliefF algorithm was used to determine the relative weight of matching variables when selecting donors
 - Non-parametric method using similar assumptions as nearest-neighbour imputation: records with similar matching variables should come from the same class
 - More predictive variables are given more importance when selecting a donor
 - Matching variables which are different between classes have weights increased
 - Matching variables which are different within classes have weights decreased

Matching variables by predictive value (descending)

Adults not in a couple	Adults in a couple	
	Spouse has same year of immigration	Spouse has a different year of immigration
Place of birth Age at immigration Year of immigration Marital status Highest level education Knowledge off. languages Province of residence Occupation Sex Subsidized housing	Spouse's admission cat. Spouse's applicant type Place of birth Year of immigration Age at immigration Highest level education Occupation Province of residence Knowledge off. languages Sex	Place of birth Year of immigration Spouse's admission cat. Age at immigration Spouse's year immigration Highest level education Spouse's applicant type Sex Knowledge off. languages Province of residence



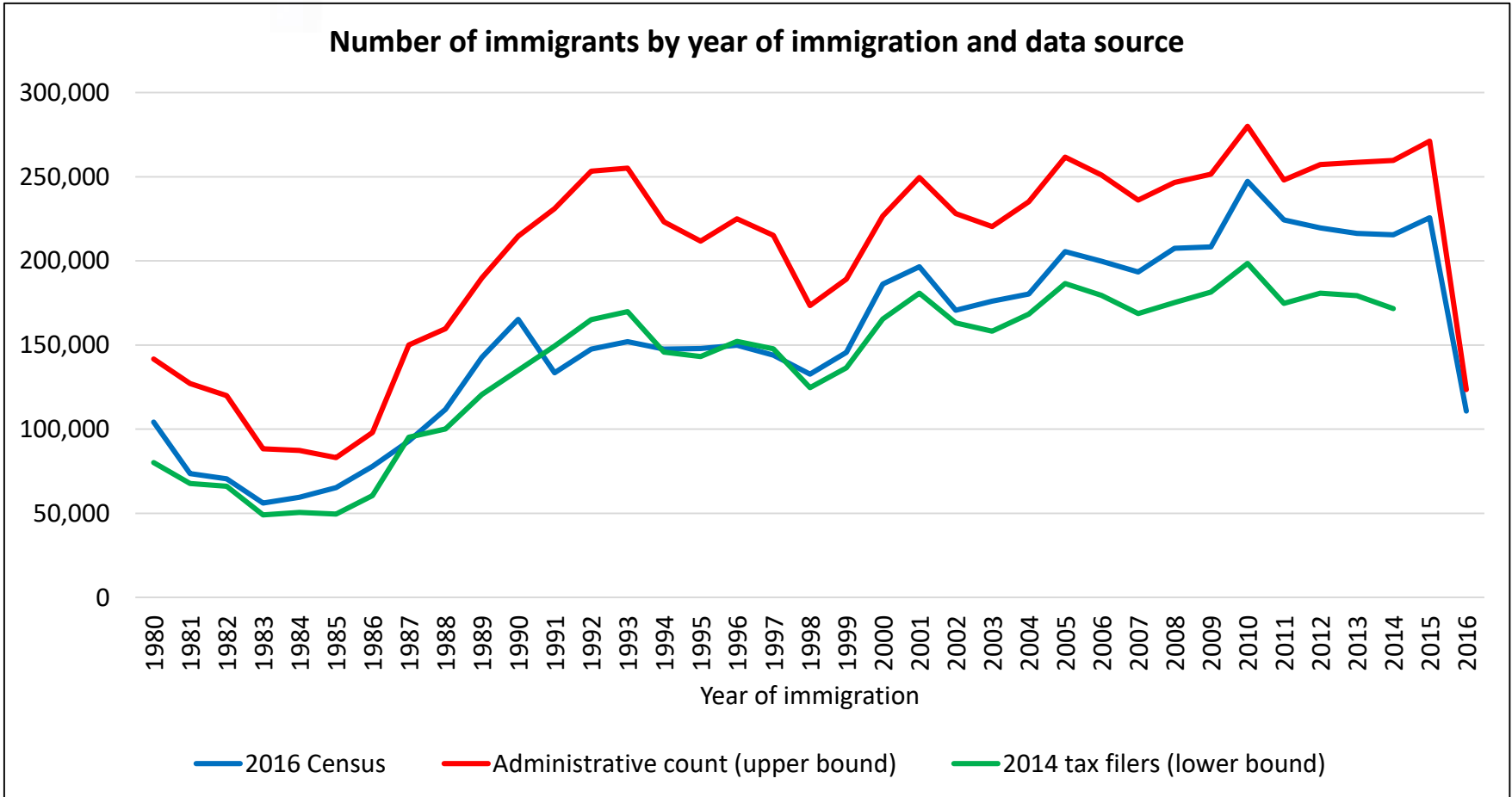
During testing, imputation results were in line with original results even for small subpopulations



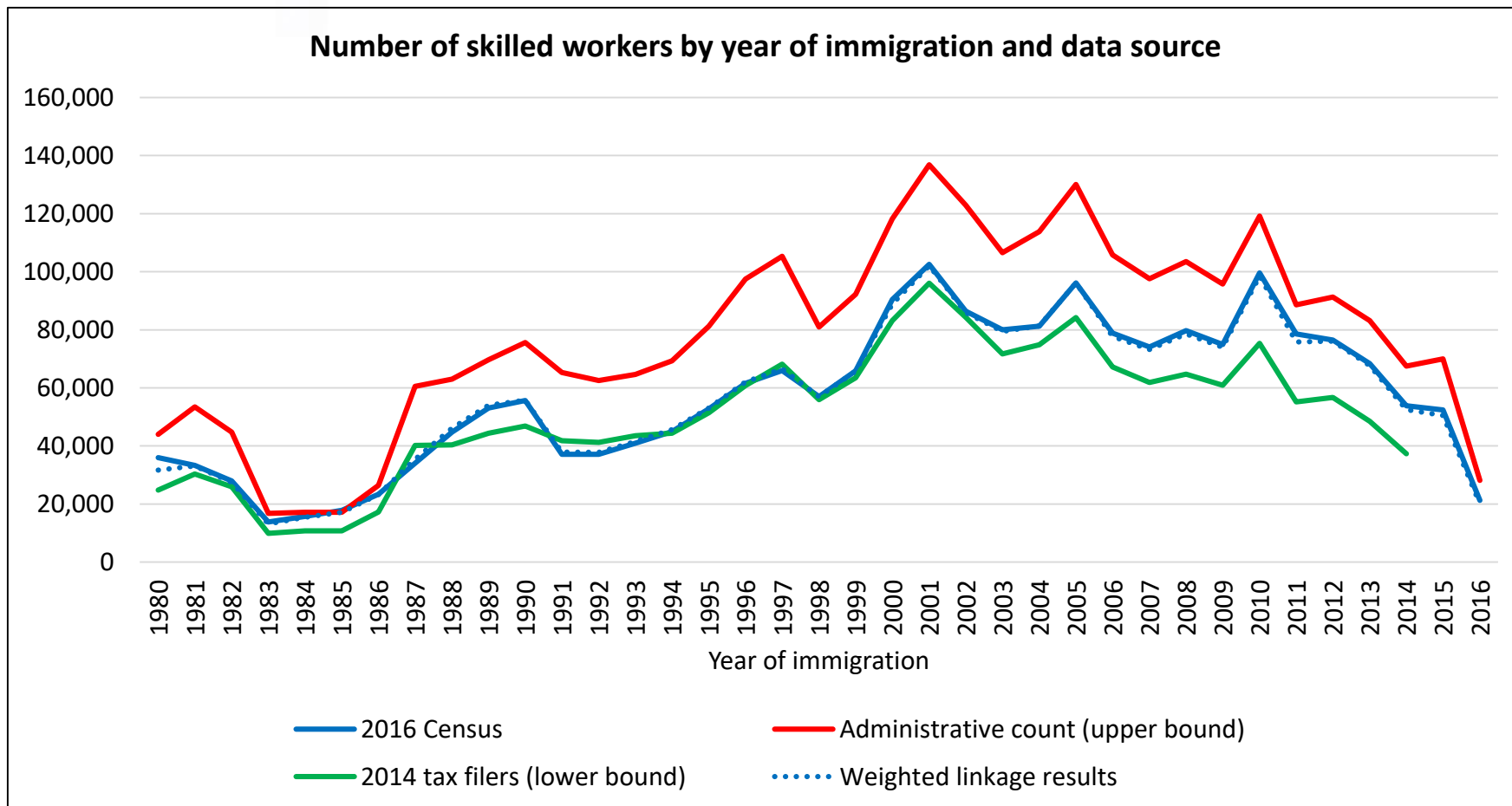
Source: Statistics Canada, Edit and imputation test data, 2011 National Household Survey and Immigrant landing file linkage



Final 2016 Census results were certified against other data sources



Admission category variables aligned with other data sources after imputation





Summary and next steps

- 2016 Census immigration data were enhanced with integrated administrative data
 - Improvements made to processing
 - New variables add analytical value to important data source for immigrants in Canada (collaboration with IRCC)
- Immigration results were released October 25th, 2017
 - Including new admission category variables
 - Reference material released providing context and explanations of variables and comparability with past results



Acknowledgements

- Partners at Immigration, Refugees and Citizenship Canada
 - Lorna Jantzen and colleagues
- Social Survey Methods and Household Survey Methods Divisions
 - Andrew Stelmack, Lyne Guertin, Sean Crowe, Piotr Biernot, Caroline Pelletier, Paul Cascagnette, and Colin Babyak
- Social and Aboriginal Statistics Division
 - Kathryn Spence, Eric Mongrain, Laetitia Martin, Hélène Maheux, Mireille Vézina, Chantal Poirier, Hyunji Lee, and Jarod Dobson



Contact information

Scott McLeish

Chief, Census and Administrative Data

Social and Aboriginal Statistics Division

Statistics Canada

scott.mcleish@canada.ca

Ces diapositives sont aussi disponibles en français.