



Statistics  
Canada

Statistique  
Canada

Canada



Statistics Canada  
[www.statcan.gc.ca](http://www.statcan.gc.ca)



# *Examples of Recent Data Comparability Issues in Canada*

**Brian Murphy**

**Statistics Canada**

Prepared for the 2<sup>nd</sup> Meeting of Providers of OECD Income Distribution Data  
Paris, France  
February 18-19, 2016



# Presentation Outline

- Purpose
  - To help motivate a discussion on various data issues confronted by data-experts that affect comparability of estimates across data sources and countries
  - Discuss a few broad categories of issues and provide a few specific Canadian examples



## Multiple Income sources

- Statistics Canada has multiple sources of household income data from survey and administrative sources across which regular comparisons are made
  - IID income sources
    - Survey of Consumer Finances (SCF) 1976-1997
    - Survey of Labour and Income Dynamics (SLID) 1993-2011
    - Canadian Income Survey(CIS) 2012-2013
  - Functional Urban Area administrative data
    - T1 Family File (T1FF) 1982-2013
  - Census long form, New Census income linkage

## Breaks in Series

- Recently we have changed the methodology for two key income series
  - Survey of Labour and Income Dynamics (SLID) → Canadian Income Survey (CIS) in 2012
  - Census long form → National Household Survey(2010) → Full Census-T1 linkage (2015)
- Resulted in many comparisons of income estimates
  - [Aggregate comparisons](#)
  - [NHS reference guide](#)
  - [Canadian Income Survey re-weighting paper](#)
  - [Outreach presentation for working poor research](#)



# Data Issues affecting Comparability

- Coverage
- Survey/register non-response
- Item non-response
- Coherence/alignment



## Canadian Examples: Coverage

- Coverage/universe
  - Age of respondent for income surveys:
    - 15 on census – 16 on IID income surveys
    - Children of any age who file taxes are on T1FF
    - T1FF children (CTB based construction v HH docket)
  - Household income surveys exclude certain populations (Hutterites, on-reserve, institutions, ...)
  - FUA geographical boundaries: Postal code mapping
- Collection mode – CAPI(Paper interview), CATI(telephone interview), EQ(internet), Admin (link register)



# Canadian Examples: Survey Non-Response

- Non-response
  - Income surveys
    - SLID/CIS: T4 raking margins (payroll slips)
    - No dual frame oversampling of high income strata
      - Survey of Financial Security and SPSSD/M adjust for the high income population
  - Administrative
    - 5% late filers on T1FF/LAD not accounted for
    - 20-40 top filers missing on tax files
- Collection mode – CAPI(Paper interview), CATI(telephone interview), EQ(internet), Admin (link register)



# Canadian Examples: Survey Non-Response

- SLID vs CIS
  - SLID sample
    - Initial response bias for 6 year panel
    - 67% response rate
    - Attrition
  - CIS
    - LFS respondents, short questionnaire
    - 73% response rate





## Canadian Examples: Item Non-Response

- Item non response
  - Surveys
    - SLID/CIS employ nearest neighbor techniques
    - NHS imputation of employment income
    - SPSD/M Social Assistance imputation
  - Administrative
    - T4 earnings added to non-filers on T1FF
    - Variables not on master file available on netfile for some filers



## Canadian Examples: Alignment

- Coherence/alignment
  - IID surveys
    - CIS and SLID do not force alignment with SNA totals
      - Old Age Security and Canada Pension Plan aggregates include lump-sum payments
      - SNA based social assistance control totals can include lump sum and administrative costs
  - Respondents report fewer negative incomes to interviewers than they do to administrative registers



**Thank you**