



Statistics
Canada Statistique
Canada

Canada



Statistics Canada
www.statcan.gc.ca

Using tax data to estimate the number of families and households in Canada

2012 Applied Demography Conference
Session 2C

Julien Bérard-Chagnon
Demography Division
Statistics Canada

Monday, January 9, 2012



Context

- Households (HHs) estimates are used in the calibration of many of Statistics Canada's surveys
- Census families (CFs) estimates are released yearly
- Limitations in previous methodologies fuelled the need to develop a new, integrated methodology



Plan of the presentation

- Concepts
- Previous methodologies
- T1 Family File (T1FF)
- New methodology
- Comparisons
- Concluding remarks



Concepts

- **Census families**
 - Married or common-law couple
 - Lone-parent family
- **Households**
 - Person or a group of persons who occupy the same dwelling



Previous methodologies

- **Census families**
 - Number : Component-cohort approach
 - Characteristics : extrapolation based on previous censuses
- **Households**
 - Headship ratios based on previous censuses
- **Significant limitations**



T1 Family File (T1FF)

- T1 income tax data from the Canada Revenue Agency (CRA)
- Aims at recreating Canadian population and families with tax data
- Produced yearly since 1982
- Very good coverage
- Contains basic demographic information



New method

- T1FF heads
 1. Coverage adjustment
 2. Bias adjustment
 3. Period adjustment
 4. Census coverage adjustment

$${}_i EST_{t+n}(X) = \underset{(1)}{CF} \underset{(2)}{CH_{t+n}^{T1FF}}(Y) \times \underset{(3)}{\varepsilon_t(X)} \times \frac{P_{t+n(\text{July})}^{Dem-CNU}(p, a, s)}{P_{t+n(\text{CD})}^{Dem-CNU}(p, a, s)} + \underset{(4)}{{}_i CNU_t(X)}$$



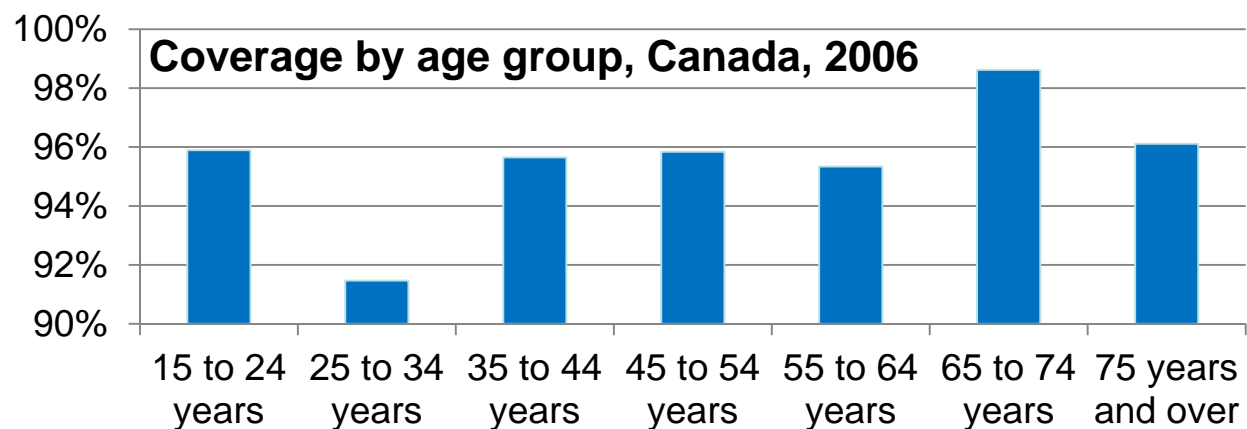
1. Coverage adjustment

$${}_i EST_{t+n}(X) = {}_{CF} CH_{t+n}^{T1FF}(Y)$$

- Adjust for coverage with population estimates

$$CH_t^{T1FF}(Y) = H_t^{T1FF}(Y) \times \frac{P_t^{Dem.}(p, a, s)}{P_t^{T1FF}(p, a, s)}$$

- Done every year





2. Bias adjustment

$${}_i EST_{t+n}(X) = {}_{CF} CH_{t+n}^{T1FF}(Y) \times \varepsilon_t(X)$$

- Comparison of T1FF and last Census

$$\varepsilon(X) = \frac{H^{Cens.}(X)}{CH^{T1FF}(Y)}$$

- Assumed to be constant until next Census
- From CFs to HHs : Census CF-HH relationship



3. Period adjustment

$${}_i EST_{t+n}(X) = {}_{CF} CH_{t+n}^{T1FF}(Y) \times \varepsilon_t(X) \times \frac{P_{t+n(\text{July})}^{Dem-CNU}(p, a, s)}{P_{t+n(\text{CD})}^{Dem-CNU}(p, a, s)}$$

- Desired date of reference : July 1st
- ${}_{CF} CH_{t+n}^{T1FF}(Y) \times \varepsilon_t(X)$ date of reference : Census Day (CD)
- Headship ratios
 - Exclusion of Census net undercoverage (CNU) from population estimates



4. Census coverage adjustment

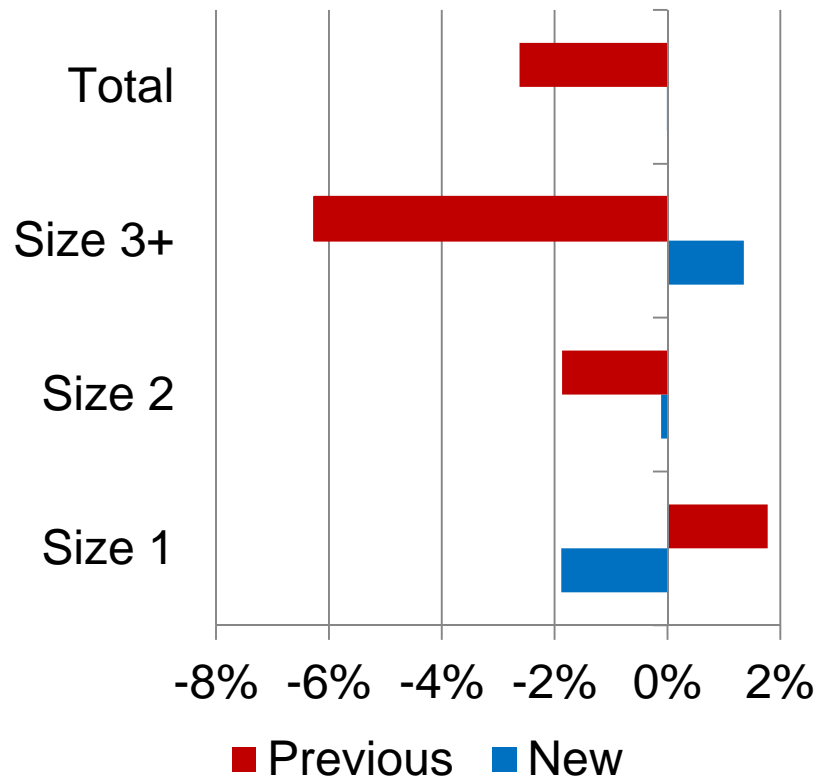
$${}_i EST_{t+n}(X) = {}_{CF} CH_{t+n}^{T1FF}(Y) \times \varepsilon_t(X) \times \frac{P_{t+n(\text{July})}^{Dem-CNU}(p, a, s)}{P_{t+n(\text{CD})}^{Dem-CNU}(p, a, s)} + {}_i CNU_t(X)$$

- Method uses Census coverage studies to estimate CF net undercoverage
 - HH net undercoverage is obtained with Census ratios
- Assumed to be constant until next Census
 - 2006 : around 150,000 CFs and 335,000 HHs

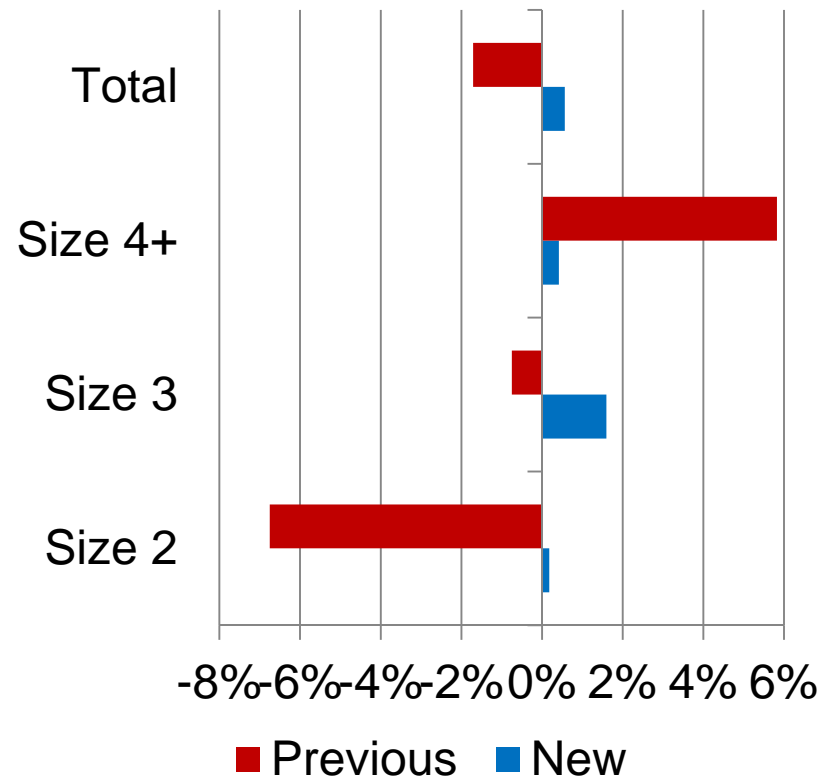


Comparison by size (with the 2006 Census)

Households



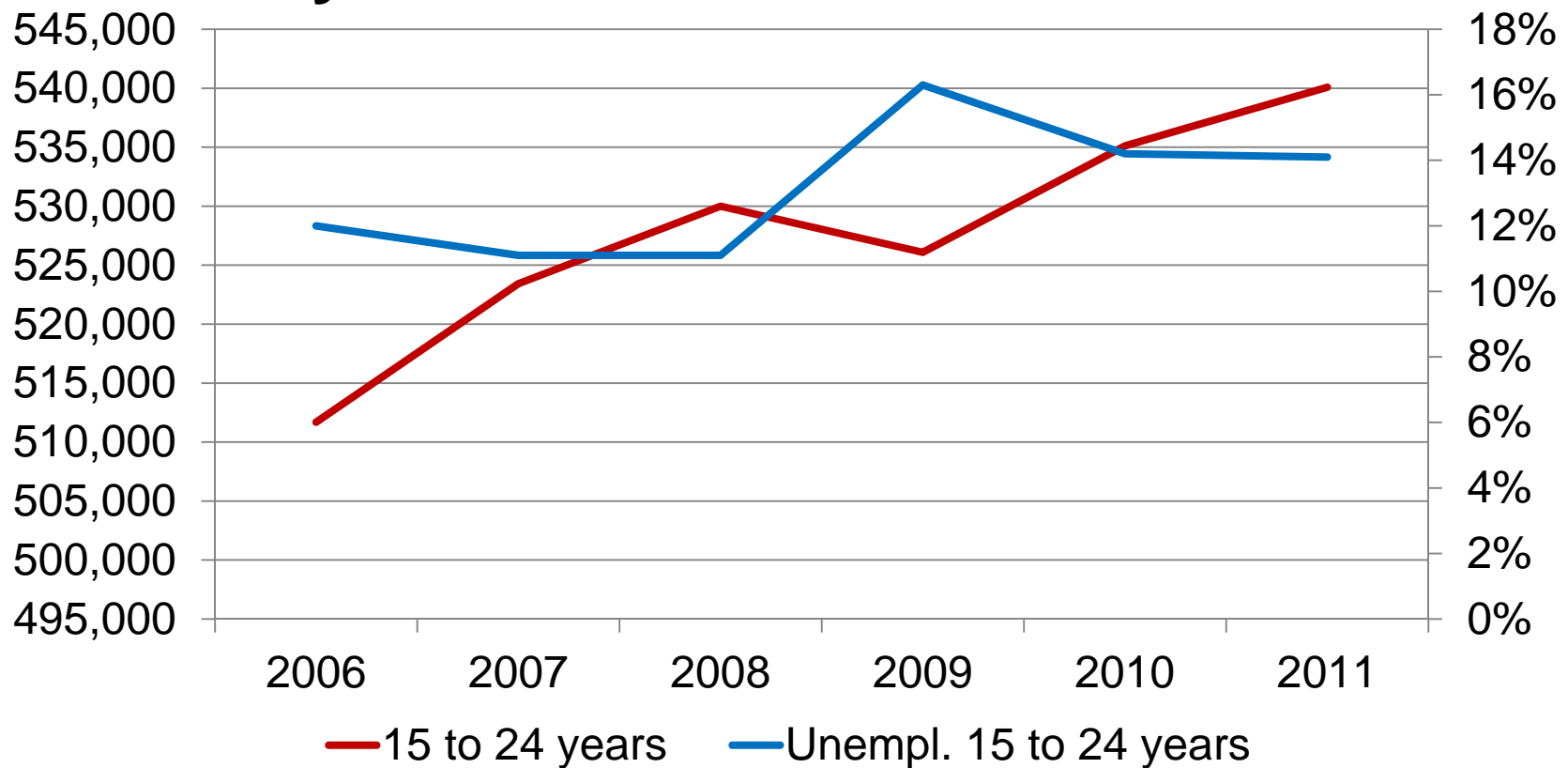
Census families





Dynamism of the estimates

HHs (left) and Unemployment rate (right), 15 to 24 years





Concluding remarks

- New method based on tax data
- Internal consistency between CFs and HHs
- Dynamic estimates
- Better results



Questions or comments ?

- You can consult our population and family estimation methods:
 - Internet : www.statcan.gc.ca
- Contact us :
 - Client Services:
1-866-767-5611
 - Email :
demography@statcan.gc.ca
- Thank you very much !
- Merci beaucoup !

