INCOME DISTRIBUTION DATA REVIEW – CANADA

1. Available data sources used for reporting on income inequality and poverty

1.1. OECD reporting:

OECD Income Distribution Data for Canada are computed by Statistics Canada and based on the *Survey of Labour and Income Dynamics* (SLID). SLID is an annual cross sectional and longitudinal survey conducted since 1996. Data from 1976 to 1996 are from the Survey of Consumer Finances (SCF), an annual cross sectional survey. In 2009, Statistics Canada has revised the SCF data for the OECD in order to match SCF and SLID. Canada provides the full questionnaire response on an annual basis starting with the year 1976.

1.2. National reporting and reporting in other international agencies:

1.2.1 National reporting:

Income distribution and poverty indicators for Canada are also available from the Census which is conducted every five years.

An additional source is the administrative source from The Canada Customs and Revenue Agency (CCRA), T1FF which is a census of all individuals who file taxes, or whose social insurance number (SIN) appears on another family members' tax file.

As Census does not include information on taxes, a Census after-tax (Census-AT) dataset has been created by Statistics Canada imputing taxes for families using a reduced form approach with estimation of regressions of taxes paid on family characteristics using tax data and then using the estimated coefficients from those regressions to predict taxes paid in the Censuses (see http://publications.gc.ca/Collection/Statcan/11F0019MIE/11F0019MIE2006274.pdf).

1.2.2 International reporting:

LIS uses, such as OECD, SLID as a basis for data reporting (see http://www.lisdatacenter.org/our-data/lis-database/by-country/canada-2/).

OECD earnings indicators for Canada refer to gross weekly earnings distribution for full-time workers from Labour Force Survey, Statistics Canada, V0902_06 (see Annex III).

Table 1 presents the main characteristics of the different sources:

Table 1. Characteristics of datasets used for income reporting, Canada

	OECD income distribution database	National survey SCF Survey of Consumer Finances (SCF) pre 1996 (Income)	National survey Survey of Labour and Income Dynamics (SLID) 1996-2011 (Income)	Census (1980,1985,1990,1995,2000,2006, 2011)	Administrative source: The Canada Customs and Revenue Agency (CCRA) T1FF (the T1 Family File)
Name	see SLID	Survey of Consumer Finances (SCF)	Survey of Labour and Income Dynamics (SLID)	Census	
Name of the responsible agency		Statistics Canada	Statistics Canada	Statistics Canada	
Year (survey and income/wage)		1980 available up to 1996	1996-2011	Every five years	
Period over which income is assessed			Each year, a panel is interviewed in January (mainly to collect labour information) and in May (to collect income information).	Shortly after the tax season (sometime in May in recent years).	
Covered population		The SCF is an annual crosssectional survey that targets all households in Canada, with the exception of those living in the territories, institutions, or on native reserves. The exceptions account for less than 3% of the Canadian population.	The survey covers all individuals in Canada, excluding residents of the Yukon, the Northwest Territories and Nunavut, residents of institutions and persons living on Indian reserves. Overall, these exclusions amount to less than 3 percent of the population. Due to the survey coverage error, SLID covered 84 percent of its target population.	The Census is a survey aimed at collecting information on the entire population, and is conducted every five years shortly after the tax season (sometime in May in recent years).	T1FF is a census of all individuals who file taxes, or whose social insurance number (SIN) appears on another family members' tax file.Population coverage rate has been over 95% since 1992 in T1FF.
Sample size		Small: The sample of roughly 35,000 households is selected as a supplement to the April Labour Force Survey (LFS).	Small: The sample contains 64,783 persons, 27,843 economic families and 26,745 households.	Large 20% of the population	Large
Sample procedure			Cross-sectional estimates, it is also designed for longitudinal analysis. The samples for SLID are selected from the monthly Labour Force Survey (LFS) and thus share the latter's sample design (stratified, multi-stage design that uses probability sampling).Panels are interviewed for up to 6 years, with new (and overlapping) panels introduced every 3 years. For the income interview, respondents have the option of allowing Statistics Canada to link to their T1 tax files (if possible) in order to collect their income information, thus eliminating the need for an income questionnaire. More than 80% of respondents provide Statistics Canada with the permission to attempt this match, and the income of about 70% of all respondents is obtained from the tax files in this way.		
Response rate		Response rates generally over around 80%	Slightly higher (80% to 85%)		
Imputation of missing values					
Unit for data collection					
Break in series			New survey after 2011.Survey of Longitudinal and International Study of Adults (LISA)		The T1FF has been particularly well-suited for estimates of income at the lower end of the distribution since 1992 given the large number of incentives for lower income families to file taxes.2 On the other hand, the creation of those incentives implies that tax data are unlikely to be based on a consistent sample of observations before versus after 1992. Indeed, given changes from the Child Tax Credit to the Child Tax Benefit in 1993, the 1992 data may also not be comparable to subsequent years.3 This, plus the fact that transfer income was not reported at all before 1989 and not reported consistently before 1992 implies that tax data before 1992 are not comparable. 1993 should be considered as a starting point in establishing trends.
Definition of reference person		Families are derived with respect to the "head" of the family, which gives priority to the husband	In SLID, it is with respect to the "major income recipients"		
Definition of household			A household is defined as being composed of a person or group of persons who co reside in,or occupy the same dwelling and do not have a usual place of residence else where in Canada or abroad.Household can refer to the economic family defined as a group of two or more persons who live in the same dwelling and are related to each other by blood,marriage,a common law union or adoption.Household can also contain two or more economic families composed of individuals who are unrelated by blood or law but live together in the same dwellin,such as roommates or alodger (non-familyhousehold). An ndividual living alone is considered as one person household. The household members who are temporarily absent (e.g. temporary residents elsewhere) are considered part of their usual household.		
Web source:		http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3502&lang=en&db=imdb&adm=8&dis=2	http://www.statcan.gc.ca/pub/75f0011x/75f0011x2012001-eng.htm	http://www5.statcan.gc.ca/subject-sujet/result-resultat.action?pid=3868&id=3874&lang=eng&type=CENSUSTBL&pageNum=1&more=0	http://publications.gc.ca/Collection/Statcan/11F0019MIE/11F0019MIE2004219.pdf

2. Comparison of main results derived from sources used for OECD indicators with alternative sources

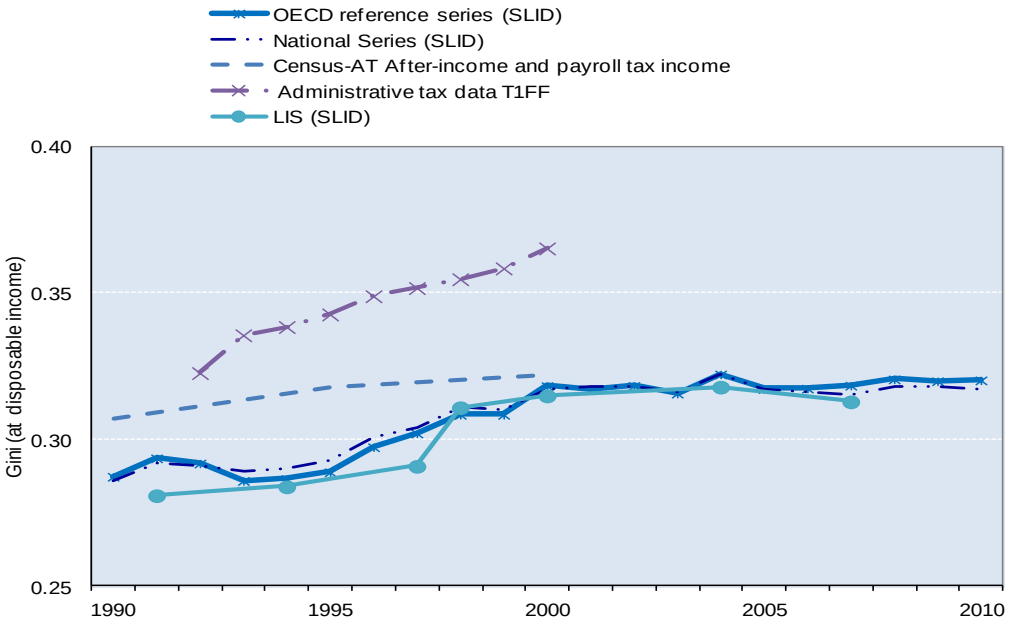
2.1 Income

2.1.1 Time series of Gini coefficients and other inequality indicators

Income indicators and Gini coefficients based on surveys are lower than those recorded under census or tax data. As response rates are lower in survey data, and non response is more likely in lower-income families, this results to a relative undercoverage at the very bottom of the income distribution and lower-income families are under-represented. The weights in the survey data are adjusted to make the sample more representative of various province-age-sex groups, as well as household and family sizes. However, no specific adjustments are made to correct the non-representativeness of the sample by income level that may be introduced by non-response.

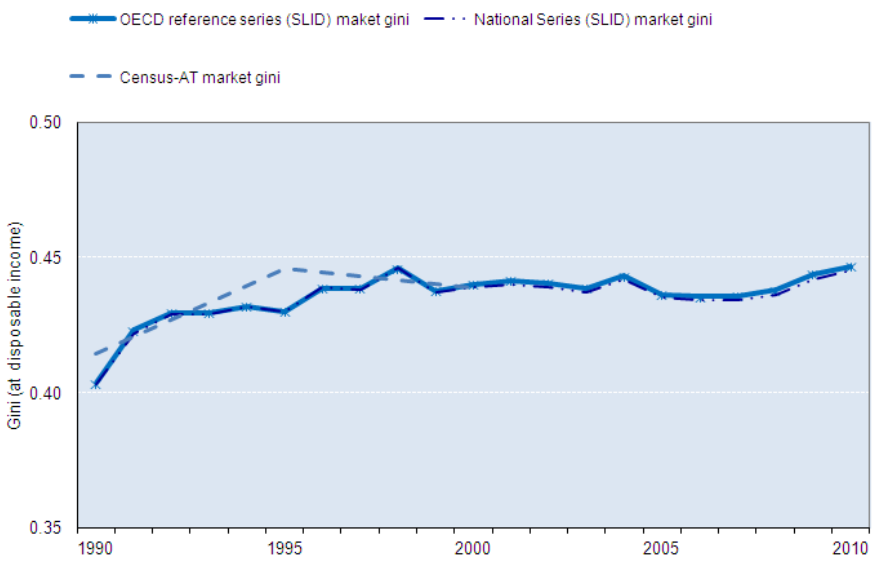
The OECD reference series for Gini coefficients of disposable and market income are quasi identical to the national series based on SLID and very similar to the LIS series which are equally based on SLID (Figures 1.1 and 1.2)

Figure 1.1 Trends in Gini coefficient (disposable income)



Source: Census-AT After-income and payroll tax income: Analytical Studies Branch Research Paper Series: Revisiting Recent Trends in Canadian After-Tax Income by Marc Frenette, David Green and Kevin Milligan Family and Labour Studies Division; Administrative tax data T1FF: Research Paper, Inequality Using Census Data: Analytical Studies Branch research paper series, Rising Income Inequality in the 1990s: An Exploration of Three Data Sources By Marc Frenette, David Green and Garnett Picot. Business and Labour Market Analysis Division; National Series (SLID): Survey of labour and Income Dynamics; LIS: Cross national data center in Luxembourg http://www.lisdatacenter.org/.

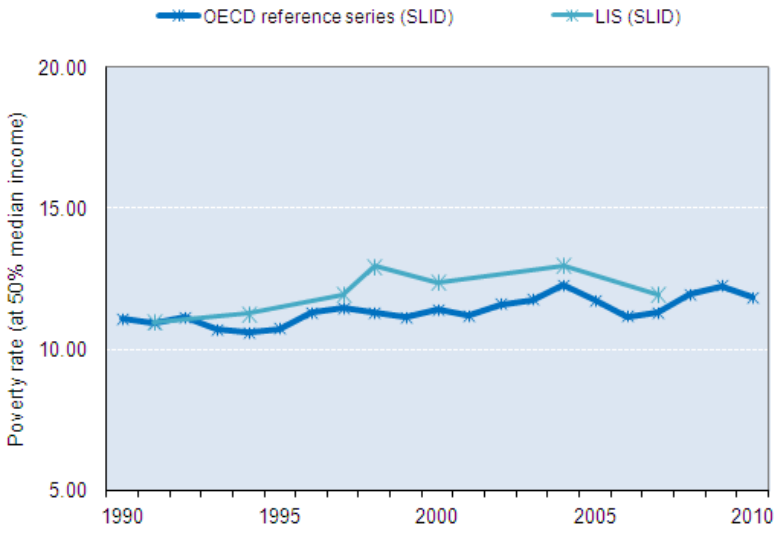
Figure 1.2 Trends in Gini coefficient (market income)



Source: Census-AT market income: Analytical Studies Branch Research Paper Series: Revisiting Recent Trends in Canadian After-Tax Income by Marc Frenette, David Green and Kevin Milligan Family and Labour Studies Division; National Series (SLID): Survey of labour and Income Dynamics; LIS: Cross national data center in Luxembourg http://www.lisdatacenter.org/.

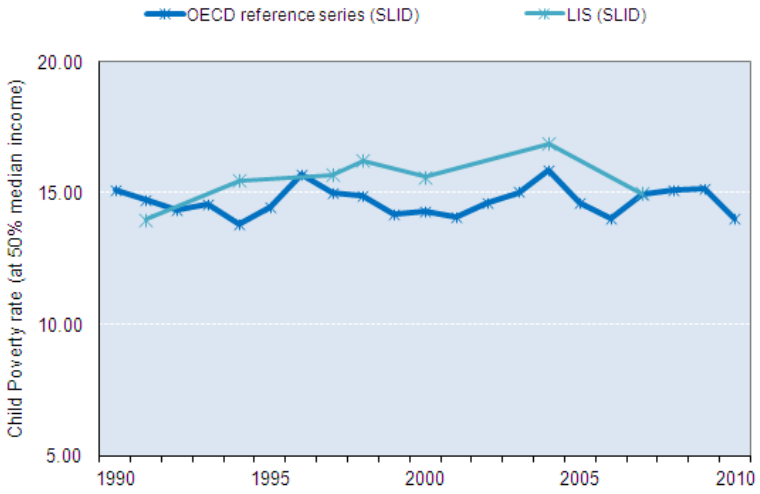
2.1.2 Time series of poverty rates

Figure 2.1 Trends in poverty rates



Source: Survey of labour and Income Dynamics (SLID); LIS: Cross national data center in Luxembourg http://www.lisdatacenter.org/.

Figure 2.2 Trends in Child poverty rates



Source: Survey of labour and Income Dynamics (SLID); LIS: Cross national data center in Luxembourg http://www.lisdatacenter.org/.

Despite the fact that data are both from SLID, poverty rates based on the OECD income distribution database and LIS differ slightly. LIS series are somewhat higher, in particular during the period 1997 through 2004. This is probably due to the fact that Statistics Canada has revised the data for the OECD in order to match SCF and SLID.

2.2 Wages

See Part II of the present Quality Review.

3. Consistency of income components shares with alternative data sources

3.1. Comparison of main aggregates: earnings, self-employment income, capital income, transfers and direct taxes

Table 2 shows shares of income components for the latest available year, according to the OECD benchmark series. Unfortunately, such information is not available for the other data sources described in table 1.

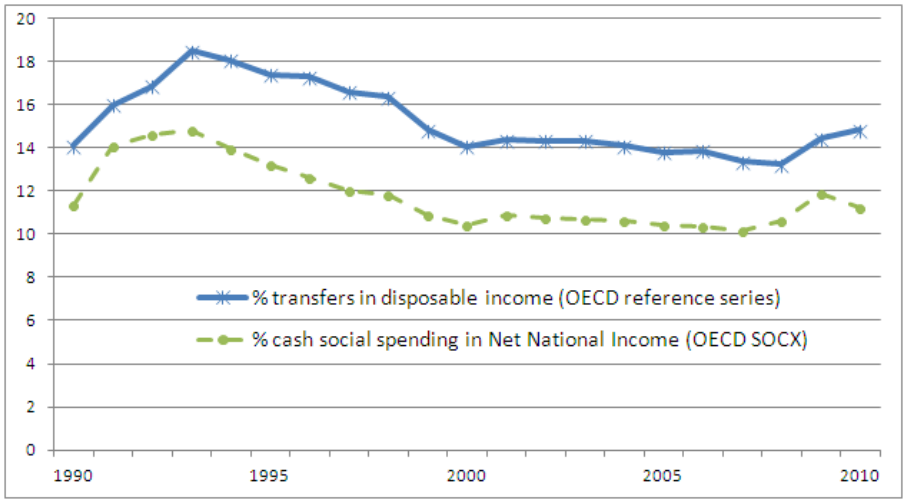
Table 2. Shares of income components in total disposable income, OECD reference series

			Average income			Average	K	SE	TR	TA		HDI
Survey	Year	Unit	EH	ES	EO	Wages	Capital	Self Employment	Transfers	Taxes		Disposable income (HDI)
OECD reference suvery	2010	natcur	25 558	6 761	3 508	35 827	6 185	3 374	6 146	-10 085		41 447
		% av HDI	61.7%			86%	15%	8%	15%	-24%		

Figure 3 compares the trend in shares of public cash transfers in equivalised disposable income from the OECD reference series with the share of total cash social spending in net national income, reported from the OECD Social Expenditure database (OECD SOCX). OECD SOCX series include pensions,

incapacity, family, unemployment, social assistance. Both series show similar trends throughout the period, except for the latest year.

Figure 3 Trends in shares of public social transfers



4. Metadata of data sources which could explain differences and inconsistencies

Definitions, methodology, data treatment

Differences between survey and tax data:

The differences lie more in coverage than in differential reporting between survey and tax data and whether survey data (both SCF and SLID) cover a different population than tax data. Lower-income families are under-represented in survey data. The weights in the survey data are adjusted to make the sample more representative of various province-age-sex groups, as well as household and family sizes. However, no specific adjustments are made to correct the non-representativeness of the sample by income level that may be introduced by non-response.

Tax data is less likely to suffer from income-related response bias since the coverage rate is above 95% (since 1992). And finally, given the larger sample sizes in tax data, inequality estimates are less susceptible to high levels of sampling variability.

In addition there is a break in series in the tax data as data improved in 1993 with the introduction of the Child Tax Benefit and a sudden increase in lower-income filers. Survey data also have a break in the series in the mid to late 1990s in the transition from SCF to SLID.

Differences between survey and Census:

Census has a greater tendancy to accurately represent families in the bottom of the income distribution as this is mandatory and the coverage is much larger than in surveys. However census data do not collect information on taxes.

The response rate in survey data is generally around 80% to 85%, and no adjustments are made to make the samples representative of the Canadian population in various income groups.

5. Summary evaluation

When comparing the pre-tax income distribution from survey and tax data with that of Census data, the bottom end of the income distribution in Census data more closely resembles the tax data than the survey data, both in terms of levels and trends.

The Census collects income data in much the same way as the SCF did, yet the coverage rate is much higher (as in the tax data). However as survey data are annual data and include complete information on income composition including taxes, they were chosen as the most appropriate for analysis in trends in household income distribution and poverty in Canada.

Bibliography:

Analytical Studies Branch Research Paper Series: Revisiting Recent Trends in Canadian After-Tax Income by Marc Frenette, David Green and Kevin Milligan Family and Labour Studies Division.

Research Paper, Inequality Using Census Data: Analytical Studies Branch research paper series, Rising Income Inequality in the 1990s: An Exploration of Three Data Sources By Marc Frenette, David Green and Garnett Picot. Business and Labour Market Analysis Division