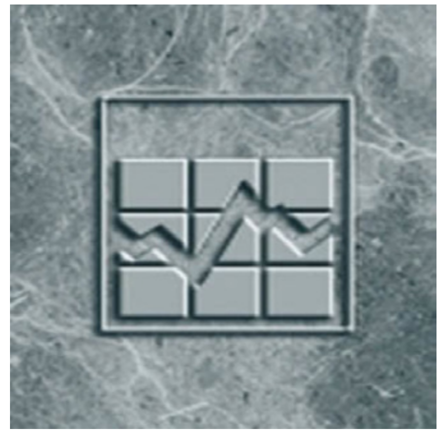
Catalogue no. 89-648-X
ISSN 1927-0100
ISBN 978-0-660-33956-6

Longitudinal and International Study of Adults Research Paper Series

Assessing the Suitability of the Longitudinal and International Study of Adults for the Estimation of Intergenerational Income Mobility



by Gaëlle Simard-Duplain and Xavier St-Denis

Release date: March 17, 2020

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

An HTML version is also available.

Cette publication est aussi disponible en français.

Table of contents

Assessing the Suitability of the Longitudinal and International Study of Adults for the Estimation of Intergenerational Income Mobility

by **Gaëlle Simard-Duplain and Xavier St-Denis**

Abstract

This study investigates the suitability of Canada's Longitudinal and International Study of Adults (LISA) for research on intergenerational income mobility. The LISA combines survey data, collected biennially since 2012, and the personal income tax records of both respondents and their past and present family members. In comparison, existing work on intergenerational mobility in Canada has often used the Intergenerational Income Database (IID), a purely administrative dataset based on the universe of tax filers. The IID's size has allowed researchers to describe the experience of mobility of narrowly defined geographic units and cohorts. However, its potential to investigate the mechanisms underlying these patterns is limited, given the small set of variables it informs. As such, the LISA is a promising candidate to further our understanding of the drivers of mobility. This study reproduces the analysis from four key papers that have documented the intergenerational transmission of income in Canada using the IID. Despite having a much smaller sample size and a different approach to the establishment of parent-child links, it finds that the LISA produces results that are consistent with the existing literature. This study also explores the sensitivity of rank-rank estimates to the choice of different specification and present results that will guide the methodological choices to be made by users of the LISA intergenerational family files in combination with LISA variables from the survey data.

Introduction

Inequalities between Canadians are partly rooted in differences observed early in their life course. For example, several important studies have shown that there is a significant relationship between the income of parents and the income of their children once they reach adulthood, meaning that to a certain extent, the opportunities available to Canadians are determined by constraints emerging shortly after their birth. In Canada, these studies have generally relied on the Intergenerational Income Database (IID), a purely administrative dataset based on data from the T1 Family Files (T1FF), which links all tax filers in the 1963-1985 birth cohort to their parents. The IID has allowed researchers to study intergenerational processes, more specifically the relationship between parental and child income (Corak and Heisz, 1999; Chen et al., 2017; Corak, 2017; Connolly et al., 2019).

More recently, Statistics Canada has extended a record linkage between the Longitudinal and International Study of Adults (LISA) and data from the T1 Family Files (T1FF). The T1FF uses information available in tax records to construct families based on the Census family concept.[1] Therefore, in addition to being linked to their own tax data, LISA respondents were matched to the T1FF records of any current or past family member identified by the Census family concept, following the approach adopted with the IID.

The LISA is a longitudinal survey first conducted in 2012, with a new wave collected biennially, and the T1FF is an administrative dataset of the universe of T1 tax filers. It offers several advantages for the study of intergenerational mobility. Most significantly, its rich survey data allows researchers not only to measure intergenerational income correlation, but also to study the drivers of economic mobility across generations. This sets it apart from purely administrative datasets, which do not document variables such as education, job characteristics, parental

1. Statistics Canada (2018a, p.37) defines a Census family as "a married couple and the children, if any, of either and/or both spouses; a couple living common law and the children, if any, of either and/or both partners; or a lone parent of any marital status with at least one child living in the same dwelling and that child or those children. All members of a particular census family live in the same dwelling. A couple may be of opposite or same sex. Children may be children by birth, marriage, common-law union or adoption regardless of their age or marital status as long as they live in the dwelling and do not have their own married spouse, common-law partner or child living in the dwelling. Grandchildren living with their grandparent(s) but with no parents present also constitute a census family."

education, and family environment at birth and during adolescence. However, the relatively small sample size of the LISA and the retrospective identification of parent-child pairs may impose limitations on its potential.

The objective of this study is two-fold. First, it replicates the results of four key papers that have significantly contributed to the measurement of intergenerational income mobility in Canada: Corak and Heisz (1999), Chen et al. (2017), Corak (2017) and Connolly et al. (2019). This allows for the assessment of the reliability of LISA intergenerational mobility measures. Second, it uses the comparison drawn with these studies to produce a set of recommendations for researchers interested in using the LISA to study the mechanisms underlying the intergenerational transmission of income. This is intended to guide users of the LISA intergenerational files as they make decisions about their estimations strategies and build their parental and child income measures.

The paper is divided in the following sections. First, the paper provides more details about how children are linked to their parents using T1FF data and focuses on sources of selection specific to the LISA data, which might be a source of difference in the estimates obtained between the LISA and the IID. The section justifies why specific birth cohorts are selected in the sample, and what factors might be driving sample selection in any intergenerational file of the type used in the LISA and IID data. The second section presents the measures of income used in the literature on the intergenerational transmission of income to address various measurement biases. In the third section, the LISA intergenerational files are used to replicate the estimates in four existing studies based on the IID. Results show some limited difference in most of the estimates, which suggests that the much smaller sample size of the LISA and the specific sample selection features do not strongly bias the results.

The study does identify three important differences that researchers need to consider when producing estimates of intergenerational mobility using LISA data:

- Various studies set inclusion conditions requiring observations to report a mean income level above a specific threshold over a given period of time, or on every year over that period. Estimates from specifications implementing the most restrictive inclusion conditions lead to dropping a large part of the sample. This appears to have an especially important impact on the size of the estimates based on LISA data.
- Consistent with the existing literature in Canada and the US, LISA estimates show relatively consistent evidence of a life cycle bias when permanent income is calculated using income reported before 30 years old. At the same, there is also a greater level of volatility than other research in the estimates based on permanent income calculated by averaging annual income over different age ranges at older ages.
- Comparing change in intergenerational mobility over time is sensitive to the number of observations in the units (groups of birth cohorts) compared, and to the inclusion of certain cohorts in the overall sample. Any researcher using LISA data to study change in intergenerational dynamics over time should pay close attention to the properties of the LISA data described in section three.

Lifecycle effects are further explored in the fourth section. The results show that the rank-rank estimates of intergenerational mobility are sensitive to lifecycle dynamics across various specifications based on different definitions of parental and child income. Including observations for which income is measured at a young age (below 30) systematically leads to estimates of smaller size (attributed to lifecycle bias, as explained in the next section).

In section four, the study also reports three noteworthy facts.

- First, the association between parental and child income is the strongest for total income and the weakest for labour earnings.
- Second, the association between parental and child income is stronger for measures of income adding the incomes of couples.
- Finally, while the rank-rank association is stronger for father-son pairs than for father-daughter pairs (and vice-versa for mother-son and mother-daughter pairs), the rank-rank association between the parent with the highest individual income (either father or mother) and their child, regardless of its gender, yields higher mobility estimates than other types of individual pairs.

Section five explores a measure of child income that is specific to the needs of researchers interested in using the administrative LISA file in combination with its survey component. When using survey variables from the LISA in conjunction with intergenerational tax data, users might favour a measure of income close to the survey date, so that the time at which income and the survey variables are measured correspond. This requires measuring the income of children from different birth cohorts at different ages (the age they had around survey date). This paper finds that researchers can comfortably pool a large number of cohorts to allow for a sample size that can support multivariate analysis if the sample lower age bound is above 30 years old in 2013, given that they applied the appropriate correction on the child income measure (described in section five). Due to lifecycle effects, the inclusion of younger cohorts appear to bias the estimates even in models correcting for the age difference of sample observations.

In the last substantive section, the paper provides evidence of non-linearities in the association between parental and child income rank. The LISA estimates are consistent with IID estimates. This is followed by a conclusion.

The results presented in this paper suggest that the LISA data is well adapted for several types of analyses of intergenerational mobility after consideration of some important limitations. In this sense, the LISA intergenerational files demonstrate potential for further explorations of the mechanisms underpinning the intergenerational transmission of income.

Data and income measures

Sample construction and selection

The LISA and IID intergenerational samples differ from one another due to their respective sample construction. In this section, these differences are discussed in order to identify what may drive discrepancies between LISA and IID estimates reported in the replication results.

The original version of the IID presented in Corak and Heisz (1999) used the population of all tax filers 16-19 years of age in 1982 as their base sample. All men born between 1963-1966 who had filed personal income taxes between 1982-1986 were drawn into the sample. A parent-child link was then established for those who lived with a father over that period. To be identified as living with a father, both the child and the parent had to have filed their taxes in the same year at least once. Hence, father-son links were observed when the children born in 1963 were 19-23 years old and when children born in 1966 were 16-20 years old. Only the first observed parent was kept. The same process was later repeated using 1984 T1FF data to draw children from the 1965-1968 birth cohort, and using 1986 T1FF data, to draw children for the 1967-1970 birth cohort (see Chen et al., 2017; Corak, 2017). Finally, the sample was further extended to include children from the 1971-1985 birth cohorts (Connolly et al., 2019). The current version of the IID includes parents and children of both sexes.

Conversely, the LISA is a representative survey sample of around 24,000 Canadians drawn in 2012 (approximately 19,000 respondents in Wave 2, the sample used for this project), and parent-child links were established retrospectively. The first step is the linkage of LISA respondents to their own T1FF records. This is achieved by a linkage between LISA data and data from the two years prior to a given LISA wave. For example, Wave 2 of LISA, which includes 2014 survey data, is linked to the T1FF records from 2012-2013. Second, if at least one linkage is established, respondents are linked retrospectively to their T1FF for all years where they filed their own taxes, starting in 1982.[2]

In both the IID and the LISA, a parent-child link is established if two conditions are met: (1) the child and the parent both filed their taxes at least once on the same year, and (2) in at least one of those years, they were residing at the same address.[3] The oldest birth cohort included was born in 1963. In the LISA, that cohort was linked with a parent at the earliest in 1982, at age 19, similar to the IID lower bound. The youngest cohort included was born in 1995.

2. The record linkage approach between LISA respondents and their T1FF records is described in more detail in Hemeon (2016). Note that that the T1FF is a standard Statistics Canada dataset. LISA users may be granted access to these files as well as some processed files developed under this project and facilitating the identification of parents.
3. The intergenerational record linkage approach is described in more detail in Statistics Canada (2018b), available to Research Data Centre users.

As is the case for the IID, parent-child links in the LISA depend on children filing their taxes in at least one year before leaving the parental home.[4] However, the intergenerational LISA sample counts are a fraction of the observations the IID contains. For instance, there are more than 2 million observations for children born in 1963-1970 in the IID, compared to less than 2,000 in the LISA sample. Furthermore, attrition may impact the LISA sample, depending on the wave users wish to use for their analysis. The baseline sample is representative of the Canadian population in 2012, and 6,600 sample members were lost to attrition after the first wave. Of the 32,100 sample members in Wave 1 (23,900 respondents), 24,300 sample members (18,500 respondents) continued on to Wave 2 (700 temporary sample members also entered the sample). Survey weights are recalculated at every wave based on observed characteristics to ensure the representativeness of the sample. However, attrition may lead to biases in studies of the intergenerational transmission of income if unobserved characteristics that are associated with attrition are also associated with intergenerational mobility.[5]

Finally, the retroactive parent-child linkages in the LISA implies that the intergenerational sample is only representative of a given cohort to the extent that members of that cohort were still present in 2012. For instance, individuals born in 1963 would have been 49 years old in 2012. Therefore, the sample for that cohort may underrepresent individuals who were born with health problems or in socio-economic circumstances associated with higher rates of morbidity, or individuals with a greater probability of permanent or temporary emigration.

Parental income measures

A fairly extensive literature concerned with the limitations of intergenerational data has investigated the effect of the timing of income measurement, both for parents and children. In particular, authors have identified two issues that have resulted in the overstatement of intergenerational mobility in earlier work: errors-in-variables and life-cycle biases (Atkinson, Maynard, and Trinder, 1983; Solon, 1992; Zimmerman, 1992; Haider and Solon, 2006). This section reviews these two issues and discusses how they are addressed in this paper, in light of the existing technical literature.

Errors-in-variables bias refers to the bias that results from using annual income or the average of a small number of annual income values instead of true permanent income. Life-cycle bias refers to the bias introduced by measuring income too early in parents and/or children's careers. This is particularly problematic if the incomes of parents and children are measured at different points in their respective life-cycles, or if children from higher income parents are more likely to have a steep age-income profile.

Both errors-in-variables and life-cycle biases result in underestimates of intergenerational income correlation. In the Canadian context, Chen et al. (2017) demonstrated how measuring parental and child income at similar ages, well into their working years, and including at least ten years of data for parents, produces estimates of the intergenerational elasticity of income (IGE) around 0.32, nearly 50% higher than previous findings by Corak and Heisz (1999). The latter had measured parental income irrespective of parents' age, and used, at most, five years of parental income data.

4. In another paper, we document the impact of that requirement on the way samples are selected in data sources constructed from tax records (Simard-Duplain and St-Denis 2019b).
5. Survey non-response and attrition are accounted for by the weights included as part of the LISA data files. In the IID, weights based on the Census were developed to account for selection due to non-filing.

To allow for comparisons with previous work, this study measures parental income by following the standard approach in the literature (Corak and Heisz, 1999; Corak, 2017; Connolly et al., 2019). Income is averaged over the five-year period when the child is 15-19 years of age,[6] which corresponds approximately to the time of the first parental link.[7] Parents are allowed to have zero income or no tax records in one or more years (coded as zero income). However, parental observations can be dropped from the sample if their income of each year or their average income falls below a certain threshold.[8] For replication purposes, this paper also reproduces the measure introduced in Chen et al. (2017), which is described in greater detail below. This measure aims to better capture permanent income and correct for lifecycle bias when parental income is measured at a different age than child income. Henceforth, this paper respectively refers to the two measures of parental income as COR and COP. All values are CPI adjusted in 2013 constant Canadian dollars.

Child income measures

The choice of child income measures can also contribute to lifecycle and errors-in-variables biases. Corak and Heisz (1999) measure child income in 1995, when the 1963-1966 cohort was between 29-32 years of age. Using similar log-log models, Chen et al. (2017) compare IGE estimates when child income is measured at 30 years of age, 40 years of age, and averaged between 38-42 years of age (inclusion criterion: three or more years with income of at least $500). The largest difference in the estimates appears between measures of income at 30 and 40 years old. Corak (2017) averaged child income over 31-32 and 38-45 age cohorts (inclusion criterion: average income of at least $500). Connolly et al. (2019) follow a similar approach based on five-year intervals from 25-39 years old. The last two papers use rank-rank estimates, which appear less sensitive to life cycle biases. The difference between these models are discussed below.

This paper reproduces the log-log and rank-rank estimates from the papers discussed above, and presents original rank-rank estimates when respondents are 25-29, 29-31 and 30-34 years old. It also introduces a measure of income specifically for LISA, based on the income observed close to the LISA data collection date. This measure is provided as a comparison based on the premise that this is the most appropriate measure of child income for studies using variables from LISA survey data. Indeed, respondent characteristics such as education, marital status, and job characteristics are measured at the moment the LISA is conducted. Using a measure of income that precedes data collection by several years would be theoretically inconsistent, especially when considering that these characteristics can highly vary over time even late in one's life course.

Reproduction of existing estimates from the Canadian literature

The complexity of the intergenerational transmission phenomenon has led to the production of several indicators that seek to describe it, including transition matrices, probabilities for children to be in bottom and top income quintiles, conditional on their parents' ranking, log-log regressions, and rank-rank regressions. Becker and Tomes (1986)'s work pioneered log-log regressions, whereby the log value of child income is regressed on the log value of parental income, yielding an estimate that corresponds to the elasticity of child income with respect to parental income. In that context, the coefficient on the log value of parental income reflects the extent to which positive or negative deviation from the mean among parents translates into deviation from the mean among children. Following the work of Dahl and Deleire (2008), Chetty et al. (2014a, b) also popularized the use of rank-rank regressions, which instead regress the child's percentile income rank on their parent's percentile income rank. This approach presents two advantages over the classical log-log regression: it does not require arbitrary decisions about how to treat zero and near-zero income values, and Corak (2017) and Connolly et al. (2019) found them much less sensitive to the way income is measured. For instance, Corak (2017) compared estimates of intergenerational mobility using child income averaged between 29-32 and between 38-45 years of age and found no significant difference. While income level changes over the life course, income ranking is more stable.

6. The T1FF data on which the LISA intergenerational files are based on start in 1982, rather than 1978 for the IID. In 1982, respondents born in 1982 were already 19 years of age. Therefore, the parental income measure for the 1963 birth cohort respondents is computed by averaging parental income over three years when the child was 19-21, decreasing the upper age bound by one year for each subsequent birth cohort until we reach 19 years old for the 1966 birth cohort. Parents of children in the 1963-1966 birth cohort have their income averaged over three years (four years for parents of children in the 1967 birth cohort). This leads to minor inconsistencies with the approaches in Corak and Heisz (1999) and Corak (2017).
7. In Corak and Heisz (1999), parental income is measured over a five-year interval spanning from 1978-1982, for children who were 16-19 years of age in 1982. Therefore, children born in 1963 were 15-19 years old between 1978-1982, and children born in 1966 were 12-16 years old in 1978-1982.
8. Corak and Heisz (1999) experimented with different exclusions based on annual income thresholds and found that the exclusion of individuals without any income in any of the five-year interval over which parental income is measured leads to a large increase in the estimates of intergenerational elasticity.

As a starting point, this study reproduces the results presented in Corak and Heisz (1999), Chen, Ostrovsky and Piraino (2017), Corak (2017) and Connolly, Haeck and Lapierre (2019). The design of these studies are formally reported in the Appendix. The replication of their results using LISA data is presented in tables 1 to 3. All LISA estimates are obtained from models using sample weights that adjust for attrition and representativeness. Overall, the replication results suggest that data from the IID and the LISA intergenerational files have similar properties, despite differences in design and sample size. There are plausible explanations for most of the differences between our estimates and those found elsewhere in the literature. This is an indication that the LISA dataset is representative of the segment of the Canadian population included in the IID, and that further research using the LISA dataset can be conducted with confidence.

Corak and Heisz (1999)

Corak and Heisz (1999) were the first to use tax data to obtain estimates of intergenerational mobility in Canada. They focused on father-son pairs and found that income correlation increased when only using years with positive income in their calculation of permanent income for parents. They also observed that earnings mobility was greater than market income mobility, and documented important non-linearities in the experience of mobility across the parental income distribution.

Father income was averaged over 1978-1982, a five-year period over which sons born in 1963-1966 were 12-19 years old (depending on the birth cohort). This corresponds broadly to our measure where parental income is averaged over a period of five years when a child is 15-19 years of age. In Corak and Heisz (1999), the measure for child income came from a single year, 1995, when sons were between the ages of 29-32. This measure is reproduced and the results are reported in table 1.

For labour earnings elasticities, Corak and Heisz (1999) report log-log coefficients of 0.131 when only observations with positive mean parental income are included (minimum $1), and 0.228 when all five years of parental income between 1978-1982 have a positive value (0.242 if the minimum in all years is set to $3,000).[9] For market income, their coefficients range from 0.194 to 0.236 (also 0.236 if the minimum in all years is set to $3,000).

As reported in Table 1, coefficients of 0.143 were obtained for labour earnings, and 0.141 for market income for the specification with a restriction based on mean parental income, when child income is measured in 1995. We obtain coefficients of 0.279 for labour earnings and 0.175 for market income when restricting the parents to those with positive income on all years. Our market income mobility estimates are smaller than in Corak and Heisz (1999). Also in contrast with their results, we obtain smaller coefficients for market income than labour earnings. When measuring child income at 30 or 31 years old, we obtain larger coefficients for market income than labour earnings (not shown).

Overall, results for father-son elasticities are consistent with those reported in Corak and Heisz (1999): the size of the coefficients are relatively similar, and fall within the large confidence intervals (95% level) of our point estimates.[10] The confidence intervals are particularly large and include zero for models requiring positive income on all years over which parental income is averaged. Differences may emerge as a result of small samples of around 300 observations, but also when a more restrictive inclusion condition is applied and observations with less than $1 of parental income on each year are excluded from the sample. This sample restriction leads to the loss of 31.9% of fathers compared to 19.4% with the less restrictive inclusion condition for labour earnings and 17.0% of fathers compared to 6.0% with the less restrictive inclusion condition for market income (first two rows of Table A1).[11] The results in Table A1 show that 32% of all fathers had at least one year without labour earnings over the period when the child was 15 to 19. It was the case for 17.0% of fathers when it comes to market income. In contrast, 19.4% of fathers did not have any employment income over that period, and 6.0% did not have any market income. It is possible that in the context of relatively small samples, LISA estimates may be particularly affected by restrictive sample inclusion conditions that require a positive value for parental income on all years.

9. Solon (1992) demonstrated that the noise in measures of permanent income that use one or few values of annual income leads to downward bias in the estimation of intergenerational income correlation (overestimation of mobility). Findings in Corak and Heisz (1999) are consistent with this. When a single year of parental income is measured instead of an average of five years, the log-log coefficients fall as low as 0.100 for labour earnings and 0.119 for market income depending on the year when parental income is measured (1978-1982).
10. The IID includes several hundred thousand observations. Confidence intervals for income elasticities tend to be below 0.01 log points.
11. Since 1993, the T1FF captures approximately 96% of Canadians (Frenette, Green and Milligan, 2007; Frenette, Green and Picot, 2004). If different individuals have missing T1FF observations from year to year, requiring five years of positive income values may exclude as much as 20% of Canadians. This is particularly concerning if those that are excluded are people with weaker labour market attachment.

Chen, Ostrovsky and Piraino (2017)

This paper then reproduces the study by Chen et al. (2017), who compared log-log estimates obtained from specifications where parental income is averaged over a five-year period (the COR measure) with those averaged over a 21-year period (the COP measure). The measure developed by Chen et al. (2017) is intended to better capture true permanent income, or lifetime income by averaging annual income over the 21 years when the parent is between ages 35-55. By averaging over a greater number of years, the approach further reduces errors-in-variables bias. Its explicit focus on the age at which income is measured among parents also limits life-cycle bias. Parental observations with less than 10 years of income at or above $500 are excluded. Child income is also measured at different ages to account for lifecycle bias.

The restrictions imposed on parental age and annual income values in the COP measure leads to the exclusion of a large proportion of observations for the 1963-1966 cohort: the sample falls to around 100 observations.[12] In replicating their estimates, this paper uses the 1963-1970 cohort to overcome potential sample size limitations. The parental inclusion condition leads to dropping up to 60.1% of all fathers, with about half of fathers being dropped because they were not in a birth cohort where they spent 10 years of their life in the 35-55 age range between 1982 and 2013 (Table A1). Among those falling in the adequate age range, the proportion of fathers dropped is between the two proportions reported in the first rows of Table A1, which is expected given that positive values are required for approximately half of all years over the 11-years period over which father's income is averaged.

The LISA estimate for the COR measure shows differences with the IID estimates of various sizes and directions. Nevertheless, the IID point estimates fall within the confidence intervals of the LISA estimates in all cases, suggesting that the differences are not statistically significant. Despite these important differences that are likely due to a combination of restrictive inclusion conditions and small sample sizes, our results are qualitatively consistent with those in Chen et al. (2017). COR estimates of intergenerational income elasticity are smaller when children are 30 years of age than when they are 40. Averaging the income of fathers over a 21-year period (the COP estimate) yields higher estimates in most cases (comparing the two bottom panels of table 2).[13] In sum, researchers using COP estimation with LISA data should carefully assess the impact of their sample design and should consider the possibility of outlier values, especially if conducting analysis on small samples or subgroups with small cell counts.

Corak (2017)

This paper also reproduces the log-log coefficients for the 1963-1970 birth cohort from Corak (2017). To implement this approach, all parents and children (men and women) are pooled in the same sample, and total family income is used as the measure of income. In Corak (2017), parental income is averaged over the five years when the child was 15-19 years old. It sums the income of both spouses when parents form a couple (common law union or marriage, according to tax records) in the first year when parents are observed, regardless of couple dissolutions over the following years used to calculate income (likewise, no change is made to account for couple formation over that period). Child income is measured as the average of total family income between 38-45 years old, following the same approach to identify couples. Observations are excluded if they have an average income of less than $500.[14] This exclusion is done on spouses separately before calculating family income. Finally, this study replicates the rank-rank estimates from Corak (2017).

In table 3, the log-log estimate is 25% larger than Corak's, at 0.255 compared to 0.201. The LISA rank-rank estimates when child income is measured at 31-32 years of age exceed those of Corak using the IID by 0.074 points. The difference is only 0.035 points when income rank is measured at 38-45 years of age. Again, the IID coefficients fall within the 95% confidence intervals of LISA estimates for all specifications. As in Corak (2017), the estimates reported in table 3 are relatively similar regardless of the age when child income is measured, especially compared to log-log estimates in Table 2. This leads to conclude that the lifecycle bias appears muted in rank-rank models using respondents over the age of 30, an observation also made by Chetty et al. (2014a) and Connolly et al. (2019).

12. Unsurprisingly, estimates based on this sample (results available upon request) take extreme values and have wide confidence intervals.
13. Chen et al. (2017) found differences of 0.04 log points or less between results based on COR and COP measures. These are unlikely to show up systematically in small sample estimates with wide confidence intervals.
14. As shown in the last row of Table A1, this condition leads to the exclusion of only 7.8% of parents.

Connolly, Haeck and Lapierre (2019)

Next, this study reproduces the results from Connolly et al. (2019), reported in Table 4. Connolly et al. (2019) use the latest version of the IID to study change in intergenerational mobility over time. They also assess the stability of rank-rank estimates of intergenerational mobility across different age groups. In other words, they document life-cycle bias in rank-rank models of income transmission in the same way as Corak (2017). The main differences are their inclusion of observations with income measured before 30 years of age[15] and the use of a wider range of birth cohorts. The age ranges used in Connolly et al. (2019) are five-year intervals: 25-29 years old, 30-34 years old, and 35-39 years old. Parental income is based on total family income, unadjusted for couples, and child income is measured as individual income.

They find a decrease in intergenerational mobility in Canada between the 1963 and 1985 birth cohorts. Using parent's total income (including spouse) measured when the child was 15-19 years of age and child individual income measured when the child was 25-29 years of age, their rank-rank estimates go from 0.189 for the 1963-1966 IID cohort to approximately 0.234 for the 1982-1985 IID cohort (Connolly et al., 2019). The results on lifecycle bias in their study align with findings from Corak (2017) as little difference was found in the estimates when child income is averaged over 30-34 years old and 35-39 years old over the 1963-66, 1967-70 and 1972-75 periods (the only periods when this comparison is possible). In addition, they do find evidence of lifecycle bias earlier in the life course of these cohorts, consistent with Chetty et al. (2014b), as evidenced by lower rank-rank estimates for income averaged over 25-29 years old compared to other age ranges.

Table 4 presents results from the replication exercise conducted in this study.[16] Rank-rank estimates using the LISA's smaller sample are similar to those of Connolly et al. (2019) for all cohort groups, given wide confidence intervals (as reported in parentheses in table 4 next to each rank-rank point estimate). Importantly, the LISA estimates also show life-cycle bias, with lower estimates when child income is averaged over the 25-29 age interval compared to the 30-34 and 35-39 (if available) intervals, except for the 1977-1979 LISA cohort for which the point estimates for income averaged over 30-34 years old is lower than the point estimate for income averaged over 25-29 years old. Unlike Connolly et al. (2019) however, the estimates for the 35-39 age range are consistently lower than those obtained for the 30-34 age range for birth cohorts where that comparison is possible.

A second substantive difference between the LISA estimates in table 4 and those from the IID reported in Connolly et al. (2019) is the absence of a trend over time in intergenerational mobility (when considering average income over 25-29 years old, which is available for 1963-1984 cohorts in the LISA). Instead, the results show important year-to-year fluctuation in the estimates and no clear trend over time. This absence of trend and the level of volatility in the rank-rank estimates may be due to sample size (400 to 1000 for each birth cohort group), translating into wide confidence intervals. In addition, the high estimates for the baseline cohorts (1963-1966) might be due to differences in the measurement of parental income for these cohorts (see footnote 6) and the fact that in the 1982 data, parent-child links could only be observed starting at 19 years old in the 1963 cohort, 18 in the 1964 cohort, 17 in the 1965 cohort, and 16 in the 1966 cohort. They can be observed starting at 15 years old for all subsequent cohorts.

In Table 5, LISA birth cohorts are grouped to form larger sub-samples and the 1963-1966 cohorts are excluded to ensure comparability over time and to mute estimate volatility that might be due to small samples. The results show evidence of a decrease in intergenerational mobility between the 1967-1976 and 1977-1982 (or 1977-1984) cohort groups although confidence intervals largely overlap. These results also demonstrate that the estimates are sensitive to the inclusion of the 1983-1984 cohort in the sample.[17] When excluding the 1983-1984 cohort, the estimates are closer to those in Connolly et al. (2019) when child income is measured at 25-29 years of age.

15. This is motivated by their use of 1963-1985 birth cohorts to study trends over time in the intergenerational transmission of income. For the youngest cohorts, income measures are available only when children were in their mid- to late-twenties.
16. Because the LISA T1FF data used in this paper stops in 2013 rather than 2014, some cohort groups exclude young respondents that are included in the cohorts created in Connolly et al. (2019). Results for these restricted cohort groups are identified in parentheses in the first column of table 4.
17. Those born in 1984 were 29 years old in 2013. The 1983 and 1984 cohorts are therefore excluded from the sample using mean child income over the 29-31 age interval.

The results for trends in intergenerational mobility suggest that researchers seeking to use the LISA to study change over time should use aggregate cohort groups that offer a sample large enough to mute any short-term volatility in the estimates, and conduct checks for the sensitivity of their estimates to the inclusion of younger or less comparable birth cohorts. Users should also be aware of lifecycle bias dynamics when selecting a measure of child income.

Individual and family income measures for various income sources

With the replication exercise concluded, this next section reports rank-rank regression results based on a wide variety of specifications in a systematic fashion. These baseline specifications are intended to provide a portrait of the sensitivity of intergenerational estimates to different modelling decisions. Important decisions include whether to use individual or family income, whether to restrict the sample to males or to females, and which source of income to consider.

Studies have adopted different approaches to these issues. For example, Corak (2017) presents estimates of intergenerational family income mobility using parents of both sexes and adding the income of spouses for couples. This contrasts with previous studies that mostly focused on father-son pairs (but also provided some basic results on father-daughter, mother-son, and mother-daughter pairs). In addition, there are different ways to measure family income. In Corak (2017), total family income is divided in half if two parents are present in the first year a parent is observed (but does not account for family transformation thereafter). This is done to facilitate the comparison between households with different structures, and it assumes an equal sharing of income between two adults in a family. Connolly et al. (2019) use total income without an adjustment for couples to capture the level of resources available to children born in those families. This is a more direct measure of the overall income level of a family, but does not account for the fact that two-parent families, on average, spend relatively less to fulfill their basic needs due to economies of scale. None of these measures are able to account for the interaction between labour supply, marital status, and fertility decisions. Connolly et al. (2019) use child individual income in most of their specifications, and present some of their results separately for sons or daughters, but not fathers and mothers. A measure of individual income can be used for parents without restricting the sample to fathers (or to mothers) exclusively. In this case, the parent with the highest income is selected. This might better capture the difference in the socioeconomic status that children are able to achieve on their own while accounting for children of lone-mothers or of parents where the mother was the highest income earner.

Rank-rank regression results using the LISA intergenerational data and following different configurations of family and individual income measures are reported in table 6 (sample sizes are reported in Table A2, row 1, showing the number of children in the cohort sample with at least one parent with an income of $500 or more). It reports estimates for the 1963-1979 birth cohorts, for all income sources: total income, market income, employment income, and labour earnings.[18]

First, the results show the differences between rank-rank estimates for 25-29 and 30-34 age groups, regardless of the source of income and whether or not couple income is added when a spouse is present. Second, the association between parental and child income is strongest for total income and weakest for employment income and labour earnings. Third, when using the measure of parental income based on the main parent (highest-income parent when a couple is observed), the estimates are lower than all other specifications except for the specification using parental family income rank adjusted for couples and child individual income rank (rows 7-9 in Table 6).

Finally, table 6 shows that the association between parental and child income is stronger for measures of income where the incomes of couples are added, with the exception of the 25-29 age group (more likely to be single, but also greater likelihood of being single for more highly educated, high-income individuals). This is consistent with the assortative mating literature in Canada (Blanden, 2005). It suggests that children born from higher income percentile parents are more likely to form a couple in their thirties with an income-earning spouse (likely a high earning spouse), than children born from lower income percentile families. This is true regardless of whether the income of couples is divided by two to adjust for the pooling of resources. Nevertheless, the adjustment for couples weakens the association between parental and child income.

18. Market income is the total income before tax excluding income from the government (transfers). This includes income from employment, investments, retirement pensions, superannuations and annuities, and other money income. Employment income includes labour earnings and self-employment income. Labour earnings includes income from paid employment as reported in a T4 slip.

Finally Table 7 reports results for the same cohort for individual parent-child pairs with different gender configurations (sample sizes reported in Table A2). An important finding is that the use of the main parent, instead of selecting pairs based on the gender of the parent, yields higher intergenerational transmission estimates for all income sources. This is true for pairs with either male or female children. Second, the income rank of fathers is more strongly associated with the income rank of sons than daughters, a finding consistent with the existing literature (Chen et al., 2017). Third, the income rank of mothers is only weakly correlated with the income rank of their sons, but is strongly correlated, in relative terms, with the income rank of their daughters. This association seems to be driven by the correlation in mother-daughter employment income and labour earnings, suggesting that an intergenerational transmission of employment participation may have occurred. Finally, there does not appear to be an impact of life cycle bias for daughters (see Chen et al., 2017, for similar evidence). This is consistent with the fact that women experience flatter earnings growth over their life course (Beach and Finnie, 2004).

Child income measurement in pooled birth cohorts samples for multivariate and subgroup analysis

Although existing bivariate estimates from previous studies could be replicated using samples with a very small number of observations, a small sample size is maladapted to subgroup analysis or multivariate analysis using a large number of covariates. Pooling birth cohorts might be attractive for researchers to obtain a sample sufficiently large for this type of analysis.

Using covariates from the LISA survey data to conduct subgroup analyses or estimate multivariate models comes with an additional challenge. Researchers might prefer using a measure of income close in time to the period when other covariates are measured. This is especially true when using variables that represent time-variant variables such as education or job characteristics. In practice, subgroup or multivariate analysis using LISA intergenerational data may therefore require pooling a large number of birth cohorts and measuring the income of children from different cohorts at very different ages. For example, an analysis pooling the 1963-1979 cohorts yields a sample of around 3000 observations. However, the age of children in the youngest cohort will be 34, while the age of those in the oldest cohort will be 48. The implications of this type of measure has not been explored in the existing literature.

Therefore, this study presents results for a measure of child income based on child close to Wave 2 data collection, which occurred in early 2014. To account for the impact of error-in-variable bias, we use a measure of child income that averages income over 2009-2013. This type of measure is useful if researchers want to use covariates available from LISA survey data.

Children of different birth cohorts have different ages over that period; for instance, children born in 1963 were 50 years old in 2013 and those born in 1979 were 34 in 2013. These are wide age ranges, which may lead to an important source of bias if the percentile rank is calculated based on the income distribution for a sample pooling all birth cohorts selected for analysis. When a large number of cohorts are pooled and the age range is wide (for example, results for the 1963-1979 birth cohorts), older individuals having reached the maximum point of their age-income curve might rank higher, on average, than younger individuals situated at the beginning of their age-income curves. In this case, income rank will be correlated with the birth cohort of observations.

One way to limit the potential effect of the life-cycle bias is calculating the income percentile separately for each birth cohort, or for narrow birth cohort intervals. Alternatively, this can be done by residualizing income on a quadratic age term before calculating percentiles.[19] Note that those two corrections will fail to account for a life cycle bias emerging from large differences in age-income profiles across children born from parents with different income ranks when individuals are in their thirties or forties (see also Haider and Solon, 2006). Instead, the corrections aim to account for the correlation between rank and birth cohort that might emerge from pooling several cohorts of children at different points of their age-income curve.

19. In this context, residualizing means extracting the residuals of a regression of child permanent income on child age, squared. The residual extracted this way is the component of child permanent income that is uncorrelated with child age.

Results for the 1963-1974, 1963-1979 and 1963-1984 cohorts are presented in Table 8.[20] The following corrections are applied and compared to the results for a model where child income percentile is calculated based on the income distribution of the pooled cohort, without accounting for potential life cycle effects:

1. Percentile ranks are calculated separately by single-year birth cohorts;
2. Percentile ranks are calculated separately for each birth cohort, using the income distribution of a 3-year pooled sample. This sample includes the cohort born the year before and after each given cohort. This allows for a larger number of observations when calculating percentile thresholds;
3. Percentile ranks are calculated on the residuals of a regression of child income on child age, squared.

The first and second corrections provide the rank of an observation relative to other observations in their birth cohort at the age they had in 2013. The residuals method ranks observations relative to the distribution income net of age for the full pooled cohorts sample in 2013.

First, table 8 shows that samples pooling the 1963-1984 cohorts almost double the sample size compared to one restricted to the 1963-1974 cohorts, from 2473 to 4043 observations when using total income. Table 8 also shows little difference in the results across specifications for the 1963-1974 and 1963-1979 estimates (the residualization method may lead to slightly higher estimates). This conclusion is reached by comparing the estimates in the first r column to the other estimates on its row. A difference of 0.02 to 0.04 points is systematically observable for the 1963-1984 pooled sample between the uncorrected and corrected percentile rank calculation methods. This suggests relatively flat age-income profiles over the 34-50 age interval, but important lifecycle bias effects when younger cohorts are included.

More generally, it can be concluded from results in Table 8 that age differences when using a measure of child income close to the Wave 2 (2014) data collection date does not bias the estimates for cohorts born before 1980, to the extent that this bias may be driven by a correlation between age and income. When pooled analysis uses younger cohorts, the corrections appear to address some of that bias.

However, these corrections are unlikely to address the effect of differences in age-income slopes for children with different levels of parental income. This might account for the smaller magnitude of intergenerational mobility estimates for pooled cohorts including individuals born after 1980, even after correction methods are applied. Indeed, if children born from higher income parents are more likely to pursue a higher level of education, a large share of those age 24-29 might have just entered the labour market or still be in full-time education. A steeper initial income growth compensating for these differentials would lead to the income gap between this group and the group of children born from low-income parents, who may be more likely to pursue a lower level of education and promptly enter the labour market, in turn leading to higher earnings.[21] Correcting for the age difference between younger and older cohorts in pooled cohort samples is an insufficient adjustment in that case.[22]

Overall, the results presented in this section show that users of the LISA conducting analyses based on child income measured close to the data collection date should be especially careful when using pooled cohort samples where a large number of cohorts were in their twenties or early thirties when child income is measured. Various corrections are unlikely to account for differences in the age-income profile of children born from parents of different income ranks. The results are consistent across sources of income and income definitions.[23]

20. Children born in 1963 were 46 years old in 2009 and 50 years old in 2013. Those born in 1974 were 35 in 2009 and 39 in 2013, those born in 1979 were 30 in 2009 and 34 in 2013, and those born in 1984 were 25 in 2009 and 29 in 2013.
21. There is important selection by parental socioeconomic status into fields of study by labour market return (Zarifa 2012a), and into graduate or professional education (Zarifa 2012b).
22. Results for the 1980-1989 birth cohorts (24-33 years old in 2013) show a rank-rank estimate of 0.11 for total income and 0.07 for employment income when using a measure of parental family income, unadjusted, and child individual income measured in 2013. When restricting the sample to the 1985-1989 cohort, results show no association between parental and child income (results available upon request).
23. Results for labour earnings and market income, and for models with a single year of child income (2013) as well as for adjusted family income are available upon request.

Non-linearities in the association between parental and child income

Existing studies in Canada (Corak and Heisz, 1999; Corak, 2017; Connolly et al., 2019) have identified non-linearities in the association between parental and child income. Non-linearity would mean that there are different levels of mobility at different points of the parental income distribution. This can be expressed by using transition matrices or by plotting the average percentile of children born in different parental income deciles, vintiles, or percentiles.

Table 9 reports quintile transition matrices for different measures of income for the 1963-1979 cohort.[24] The rows represent the quintiles of origin (parental income quintile), and the columns, the quintiles of destination (child income quintile). Values in a row sum to 100% and show the distribution of child income quintile for children born in a given parental income quintile. In a quintile transition matrix, perfect intergenerational mobility would translate into having a value of 20% for each cell.

The results show that children born from bottom or top quintile parents are more likely to remain in that income quintile as adults than those born in other parental income quintiles. Across transition matrices based on different income definitions, the probability of being in the bottom quintile conditional on being born in the bottom quintile is around 30%, and the conditional probability of being in the bottom two quintiles for that group is consistently around 55%. The opposite is true for those born in the top parental income quintile. Their conditional probability of being in the top income quintile is above 33%, and their conditional probability of being in the top two income quintiles is around 55%, with some variation depending on the income definition. The estimates are consistent with those in Corak (2017, p.18).[25]

These results show that there is a relatively strong level of mobility among children born in the middle three quintiles, although those born in the second and third quintile have a lower conditional probability of being in the top quintile as adults. Likewise, those born in the fourth parental income quintile have a lower conditional probability of being in the bottom two quintiles of total income. Overall, the results show non-linearities in the association between parental and child income, with the association being stronger at the tails of the distribution. However, the results also demonstrate that a large proportion of children experience some degree of intergenerational mobility.

The second way to represent non-linearities is to plot mean child income percentile against parental income vintiles. In Chart 1, this data is reported for the 1963-1979 cohort using total income and employment income, including spouse income, when present. Results show evidence of a steeper slope at the tails of the parental income distribution. Differences in mean child income percentile are small between children born in middle parental income vintiles, although there is some idiosyncratic volatility at some points of the parental income distribution. This suggests little difference in intergenerational mobility among children born to parents at the middle of the income distribution.

Overall, the evidence presented in this section shows important non-linearities in the association between parental and child income, a feature of the LISA intergenerational data consistent with IID data.

24. The results for all parental and child income measures and all sources of income, for parent-child pairs of all genders, are available upon request.
25. Based on a measure of child income averaged over 38-45 years of age for the 1963-1970 cohorts, Corak (2017) finds that the conditional probability of being in the top quintile is 32.3% for children born from top quintile parents and 11.4% for those born from bottom quintile parents. The conditional probability of being in the bottom quintile for children born in the bottom quintile is 30.1%. Finally, the conditional probability of being in the third quintile for children born from parents in the third quintile is 21.9%.

Conclusion

This paper is intended for potential users of the LISA intergenerational files, based on a linkage between Wave 2 (2014) of the LISA and T1FF data on survey respondents and their parents from 1982-2013. Researchers might seek to use this source of data to study how parental income is associated with different child outcomes over their life course and what characteristics measured by the survey variables might account for the intergenerational transmission of income or intergenerational mobility dynamics. This paper presents a detailed assessment of the LISA data by replicating existing studies on intergenerational mobility using the IID, an administrative dataset containing a much larger number of observations and exploring the sensitivity of the estimates of intergenerational mobility to different specification decisions.

The results show that estimates obtained from the LISA are broadly consistent with those presented in studies using the IID. Overall, the LISA data is well adapted for the study of intergenerational mobility, and that the joint weighted distribution of parental and child income is similar to the IID, despite a much smaller sample size. However, there is a large variation in the estimates across specifications, when using a limited number of birth cohorts yielding a small number of observations (e.g. father-son pairs over 1963-1970).

In addition, the results suggest that users should be aware of life cycle bias and of the possible variability of estimates over the life cycle even after 30 years old, and should be careful when using LISA data to study changes in intergenerational mobility dynamics over time.

Finally, this study discusses ways to measure child income close to data collection date. This is relevant for researchers who wish to conduct subgroup analysis or estimate multivariate models that combine child income variables from the T1FF data and child characteristics from LISA survey variables. In this case, it might be important to have corresponding time points for the measure of income and the measures of respondent characteristics. Researchers might also pool a large number of birth cohorts together in order to obtain a large enough sample to conduct their analysis. The results presented in this study suggest that measuring child income close to data collection date does not lead to biases due to age differences between children of different birth cohorts when income is measured on a given year or year interval. One exception is when the 1982-1984 cohort is included. This cohort was in their twenties for the whole 2009-2013 period including the five most recent years of T1FF data prior to LISA Wave 2 (2014) data collection, used in this paper. This bias is evident when calculating child income percentiles based on the pooled sample of respondents. This paper presents an approach to partially correct for this source of bias. It relies on calculating percentiles separately by single-year birth cohort or using a variable for income residualized on age in order to remove the effect of pooling respondents at different points in their age-income profile. However, this does not adjust for variation in the slopes of the age-income profiles of children born from parents in different income percentiles.

New research integrating LISA survey variables will be able to assess whether some of the challenges identified in this paper have implications for studies focusing on variables associated with the intergenerational transmission of income. This is explored in a follow-up paper on the role of education and other child outcomes in the intergenerational transmission of income (Simard-Duplain and St-Denis, 2019a). In addition, a companion paper focuses on a potentially important source of bias: the selection of respondents based on the ability to identify a parent-child linkage in the T1FF data (Simard-Duplain and St-Denis, 2019b). This provides a complementary assessment of the suitability of LISA data for the analysis of intergenerational mobility and of the characteristics of the data. The availability of a large number of survey variables on the characteristics of the respondents also allows for an exploration of the characteristics of selected observations in intergenerational databases, including the IID.

Tables and chart

Table 1
Replication of father-son intergenerational income elasticities (log-log) in Corak and Heisz (1999), 1963 to 1966 cohorts

Parent income measure	Inclusion criterion	IID estimates	LISA estimates	(95% C.I.)	Difference (IID-LISA)	Number
Labour earnings, five years mean	Mean ≥ 1	0.131	0.143	(0.014,0.272)	-0.012	297
Market income, five years mean	Mean ≥ 1	0.194	0.141	(0.025,0.257)	0.053	365
Labour earnings, five years mean	All years ≥ 1	0.228	0.279	(-0.079,0.637)	-0.051	252
Market income, five years mean	All years ≥ 1	0.236	0.175	(-0.014,0.364)	0.061	324

Note: The IID estimates are from Corak and Heisz (1999). The measure of parental income in the IID is the average income between 1978 and 1982. The measure in the LISA is the average parental income when the child was 15 to 19 (and shorter age ranges for 1963 to 1966 birth cohorts - see text). Following Corak and Heisz (1999), child income is measured in 1995. Confidence intervals for LISA estimates are calculated at the 95% level using bootstrap survey weights (normal approximation).
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013. For IID estimates: Corak and Heisz (1999).

Table 2
Replication of father-son intergenerational income elasticities (log-log) in Chen et al. (2017)

Parent income measure	Inclusion criterion	Child age	IID	LISA	(95% C.I.)	Difference
COR measure, 1963 to 1966 birth cohorts						
Labour earnings, five years mean	All years ≥ 500	30	0.227	0.098	(-0.105,0.302)	0.129
		40	0.287	0.203	(0.040,0.366)	0.084
Market income, five years mean	All years ≥ 500	30	0.230	0.072	(-0.173,0.316)	0.158
		40	0.301	0.315	(0.130,0.500)	-0.014
Total income, five years mean	All years ≥ 500	30	0.222	0.302	(0.028,0.576)	-0.080
		40	0.317	0.461	(0.221,0.702)	-0.144
COR measure, 1963 to 1970 birth cohorts						
Labour earnings, five years mean	All years ≥ 500	30	0.227	0.227	(0.079,0.374)	0.000
		40	0.287	0.352	(0.213,0.490)	-0.065
Market income, five years mean	All years ≥ 500	30	0.230	0.199	(0.027,0.370)	0.031
		40	0.301	0.423	(0.261,0.585)	-0.122
Total income, five years mean	All years ≥ 500	30	0.222	0.323	(0.125,0.520)	-0.101
		40	0.317	0.532	(0.366,0.698)	-0.215
COP measure, 1963 to 1970 birth cohorts						
Labour earnings, mean over age 35 to 55	10 years ≥ 500	40	0.321	0.216	(0.099,0.334)	0.105
		38 to 42	0.318	0.530	(0.201,0.860)	-0.212
Market income, mean over age 35 to 55	10 years ≥ 500	40	0.349	0.304	(0.164,0.444)	0.045
		38 to 42	0.343	0.378	(0.198,0.558)	-0.035
Total income, mean over age 35 to 55	10 years ≥ 500	40	0.359	0.460	(0.309,0.612)	-0.101
		38 to 42	0.359	0.389	(0.252,0.525)	-0.030

Note: All IID estimates are from Chen et al. (2017). They are based on a sample of observations from the 1963 to 1966 birth cohorts. As indicated in the table, some LISA estimates are based on a sample of respondents in the 1963 to 1970 birth cohorts in order to obtain a sufficiently large sample size. The "Diff." column shows the difference between IID estimates and LISA estimates. The COR measure of parental income in the IID is the average income between 1978 and 1982. The COR measure in the LISA is the average parental income when the child was 15 to 19 (and shorter age ranges for 1963 to 1966 birth cohorts - see text). Parents are excluded if their income is lower than 500 on any of the five years over which income is averaged. When child income is average between 38 to 42 years old, children are excluded if their income falls under $500 for three or more years. Confidence intervals for LISA estimates are calculated at the 95% level using bootstrap survey weights (normal approximation).
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013. For IID estimates: Chen, Ostrovsky and Piraino (2017).

Table 3
Replication of intergenerational mobility estimates in Corak (2017)

Parent income measure	Inclusion criterion	Child age	IID	LISA	(95% C.I.)	Difference	Number
Intergenerational income elasticity (log-log)							
Total family income, five years mean	Mean ≥ 500	Mean 38 to 45	0.201	0.255	(0.184,0.326)	-0.054	1,724.00
Relative mobility (rank-rank)							
Total family income, five years mean	Mean ≥ 500	Mean 31 to 32	0.240	0.314	(0.226,0.402)	-0.074	775.00
		Mean 38 to 45	0.242	0.277	(0.210,0.343)	-0.035	1,724.00

Note: All estimates use average parental income when the child was 15 to 19 (and shorter age ranges for 1963 to 1966 birth cohorts in LISA data - see text). The "Diff." column shows the difference between IID estimates and LISA estimates. Estimates using a child income measure based on mean income over age 38 to 45 are for the 1963 to 1970 birth cohorts. Estimates using a child income measure based on mean income over age 31 to 32 are for the 1967 to 1970 birth cohorts. Total family income divided by two if spouse is present. Both members of a couple need to be above the inclusion threshold for each of them to be included. See Corak (2017) for details. Confidence intervals for LISA estimates are calculated at the 95% level using bootstrap survey weights (normal approximation).
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013. For IID estimates: Corak (2017).

Table 4
Replication of intergenerational mobility estimates from Connolly et al. (2019)

Source	25 to 29 years old		30 to 34 years old		35 to 39 years old	
	Estimates	(95% C.I.)	Estimates	(95% C.I.)	Estimates	(95% C.I.)
			1963 to 1966			
IID	0.189	..	0.210	..	0.210	..
LISA	0.215	(0.111, 0.319)	0.255	(0.159, 0.351)	0.237	(0.137, 0.336)
			1967 to 1970			
IID	0.188	..	0.214	..	0.214	..
LISA	0.192	(0.103, 0.281)	0.222	(0.135, 0.309)	0.207	(0.123, 0.292)
			1972 to 1975			
IID	0.216	..	0.232	..	0.230	..
LISA	0.208	(0.109, 0.307)	0.210	(0.109, 0.311)	..	..
LISA (1972 to 1974)	0.184	(0.080, 0.288)	0.232	(0.127, 0.338)	0.190	(0.085, 0.295)
			1977 to 1980			
IID	0.215	..	0.235	..	..	..
LISA	0.271	(0.164, 0.377)	..	..	..	..
LISA (1977 to 1979)	0.232	(0.112, 0.351)	0.196	(0.078, 0.314)	..	..
			1982 to 1985			
IID	0.234	..	..	..	..	..
LISA (1982 to 1984)	0.182	(0.068, 0.296)	..	..	..	..

.. not available for a specific reference period

Note: All estimates use average parental total family income (sum of both parents for couples) when the child was 15 to 19. Observations are included if their average income is $500 or more. Both members of a couple need to be above inclusion threshold for each of them to be included. IID estimates are rank-rank estimates reported in Connolly et al. (2019), table 3. Birth cohorts correspond to the ones used in Connolly et al. (2019) unless otherwise specified in parentheses in the Source column. Confidence intervals not provided; standard errors are all 0.001 or lower. Confidence intervals for LISA estimates (reported in parentheses) are calculated at the 95% level using bootstrap survey weights (normal approximation).
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013; IID estimates from Connolly et al. (2019).

Table 5
Trend in intergenerational mobility by income definition, aggregate birth cohort groups, 1967 to 1984

Parental income measure	Child income measure	1967 to 1976		1977 to 1982		1977 to 1984	
		Estimates	(95% C.I.)	Estimates	(95% C.I.)	Estimates	(95% C.I.)
				25 to 29 years old			
Parents family income, unadjusted	Child individual income	0.191	(0.132, 0.250)	0.221	(0.134, 0.309)	0.212	(0.138, 0.286)
Parents family income, unadjusted	Child and spouse income, unadjusted	0.172	(0.115, 0.230)	0.244	(0.159, 0.328)	0.207	(0.133, 0.280)
Parents family income, adjusted	Child individual income	0.123	(0.060, 0.186)	0.155	(0.066, 0.245)	0.146	(0.070, 0.221)
Parents family income, adjusted	Child and spouse income, adjusted	0.135	(0.070, 0.200)	0.174	(0.084, 0.263)	0.148	(0.070, 0.225)
Main parent individual income	Child individual income	0.150	(0.090, 0.210)	0.179	(0.089, 0.268)	0.162	(0.088, 0.236)
				29 to 31 years old			
Parents family income, unadjusted	Child individual income	0.195	(0.138, 0.252)	0.212	(0.124, 0.299)	..	..
Parents family income, unadjusted	Child and spouse income, unadjusted	0.208	(0.150, 0.266)	0.267	(0.184, 0.350)	..	..
Parents family income, adjusted	Child individual income	0.159	(0.101, 0.218)	0.173	(0.081, 0.264)	..	..
Parents family income, adjusted	Child and spouse income, adjusted	0.195	(0.133, 0.258)	0.211	(0.121, 0.301)	..	..
Main parent individual income	Child individual income	0.169	(0.111, 0.227)	0.175	(0.084, 0.265)	..	..

.. not available for a specific reference period

Note: All IID estimates are from Chen et al. (2017). They are based on a sample of observations from the 1963 to 1966 birth cohorts. As indicated in the table, some LISA estimates are based on a sample of respondents in the 1963 to 1970 birth cohorts in order to obtain a sufficiently large sample size. The "Diff." column shows the difference between IID estimates and LISA estimates. The COR measure of parental income in the IID is the average income between 1978 and 1982. The COR measure in the LISA is the average parental income when the child was 15 to 19 (and shorter age ranges for 1963 to 1966 birth cohorts - see text). Parents are excluded if their income is lower than 500 on any of the five years over which income is averaged. When child income is average between 38 to 42 years old, children are excluded if their income falls under $500 for three or more years. Confidence intervals for LISA estimates are calculated at the 95% level using bootstrap survey weights (normal approximation).
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013. For IID estimates: Chen, Ostrovsky and Piraino (2017).

Table 6
Intergenerational mobility by income source and child age, 1963 to 1979 birth cohorts

Parent income definition	Child income definition	Child age	Total income		Market income		Employment income		Labour earnings	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Parents family income rank, unadjusted	Child individual income rank	25 to 29	40.18	0.204	40.54	0.197	41.75	0.173	42.24	0.163
		29 to 31	39.46	0.218	41.13	0.185	41.78	0.172	42.06	0.167
		30 to 34	39.24	0.223	40.30	0.201	41.05	0.187	41.64	0.175
Parents family income rank, unadjusted	Child family income rank, unadjusted	25 to 29	40.49	0.197	40.66	0.194	41.75	0.173	42.28	0.162
		29 to 31	38.17	0.244	39.26	0.222	40.18	0.204	40.60	0.196
		30 to 34	37.17	0.264	38.11	0.245	39.09	0.225	39.98	0.208
Parents family income rank, adjusted	Child individual income rank	25 to 29	42.60	0.156	43.29	0.142	44.63	0.116	44.99	0.108
		29 to 31	40.76	0.192	42.99	0.148	43.64	0.135	43.77	0.133
		30 to 34	39.91	0.209	41.60	0.176	42.80	0.152	43.38	0.140
Parents family income rank, adjusted	Child family income rank, adjusted	25 to 29	41.60	0.176	41.97	0.168	43.61	0.136	43.90	0.130
		29 to 31	38.73	0.233	40.46	0.198	41.73	0.173	42.15	0.165
		30 to 34	37.79	0.251	39.25	0.222	40.51	0.197	41.04	0.187
Parent (main) individual income rank	Child individual income rank	25 to 29	41.47	0.178	41.55	0.176	42.81	0.152	43.51	0.138
		29 to 31	40.40	0.200	41.89	0.170	42.52	0.157	42.64	0.155
		30 to 34	39.63	0.215	40.66	0.194	41.57	0.176	41.98	0.168

Note: All estimates use average parental income (sum of both parents for couples) when the child was 15 to 19. Unadjusted family income is the sum of the income of both spouses, when present. Adjusted family income divides the sum of the income of both spouses by two, for couples. Results using main parent's individual income rank uses the income of the parent with the highest income, if two parents are observed. Observations are included if their average income has a value of $500 or more. Both members of a couple need to be above inclusion threshold for each of them to be included.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Table 7
Intergenerational mobility by parent and child sex and child age, 1963 to 1979 birth cohorts

Parent income definition	Child income definition	Child age	Total income		Market income		Employment income		Labour earnings	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Main parent	Son	25 to 29	40.92	0.19	42.21	0.16	43.80	0.13	44.00	0.13
		29 to 31	38.03	0.25	39.85	0.21	40.37	0.20	40.16	0.20
		30 to 34	39.50	0.22	41.58	0.18	42.40	0.16	42.20	0.16
Main parent	Daughter	25 to 29	40.61	0.19	39.14	0.22	40.08	0.21	41.58	0.18
		29 to 31	39.51	0.22	39.03	0.23	40.39	0.20	42.14	0.16
		30 to 34	39.46	0.22	39.68	0.21	40.36	0.20	41.33	0.18
Father	Son	25 to 29	42.38	0.16	42.69	0.15	44.17	0.12	44.24	0.12
		29 to 31	39.67	0.21	39.73	0.21	40.81	0.19	41.43	0.18
		30 to 34	41.53	0.18	42.32	0.16	43.44	0.14	43.16	0.14
Father	Daughter	25 to 29	44.02	0.13	42.33	0.16	43.58	0.14	45.59	0.10
		29 to 31	42.96	0.15	42.41	0.16	43.62	0.13	45.68	0.09
		30 to 34	42.98	0.15	43.13	0.14	44.36	0.12	45.87	0.09
Mother	Son	25 to 29	46.77	0.07	47.42	0.06	49.43	0.02	49.00	0.03
		29 to 31	44.02	0.13	46.28	0.08	47.85	0.05	48.23	0.04
		30 to 34	43.24	0.14	45.42	0.10	47.00	0.07	47.42	0.06
Mother	Daughter	25 to 29	41.15	0.18	40.86	0.19	41.43	0.18	41.43	0.18
		29 to 31	40.87	0.19	40.18	0.20	41.02	0.19	41.55	0.18
		30 to 34	41.36	0.18	41.08	0.19	40.56	0.20	41.10	0.18

Note: All estimates use average parental income when the child was 15 to 19. Results using main parent's individual income rank uses the income of the parent with the highest income, if two parents are observed. Observations are included if their average income has a value of $500 or more.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Table 8
Intergenerational mobility estimates in pooled birth cohorts samples when child income is measured at LISA data collection date, by percentile calculation method

	Pooled cohorts		By single birth year		By cohort, 3-year groups		Residualized on age		Number
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
1963 to 1974 birth cohorts (39 to 50 years old in 2013)									
Employment income									
Child individual income	40.99	0.19	40.58	0.19	40.60	0.19	40.46	0.20	2,272.00
Child and spouse income, unadjusted	39.85	0.21	39.20	0.21	39.31	0.22	39.37	0.22	...
Total income									
Child individual income	38.62	0.23	38.59	0.23	38.55	0.23	38.07	0.25	2,473.00
Child and spouse income, unadjusted	36.75	0.27	36.54	0.27	36.54	0.27	36.38	0.28	...
1963 to 1979 birth cohorts (34 to 50 years old in 2013)									
Employment income									
Child individual income	41.23	0.18	40.45	0.19	40.69	0.19	40.71	0.19	2,995.00
Child and spouse income, unadjusted	39.25	0.22	37.98	0.23	38.41	0.23	38.63	0.23	...
Total income									
Child individual income	39.73	0.21	39.38	0.21	39.44	0.21	39.07	0.23	3,244.00
Child and spouse income, unadjusted	37.33	0.26	36.63	0.26	36.80	0.27	36.74	0.27	...
1963 to 1984 birth cohorts (29 to 50 years old in 2013)									
Employment income									
Child individual income	42.46	0.16	40.85	0.18	41.11	0.18	40.90	0.19	3,753.00
Child and spouse income, unadjusted	40.82	0.19	38.48	0.22	38.96	0.22	38.98	0.23	...
Total income									
Child individual income	40.74	0.19	39.53	0.21	39.70	0.21	39.27	0.22	4,043.00
Child and spouse income, unadjusted	39.01	0.23	37.29	0.25	37.60	0.25	37.44	0.26	...

... not applicable

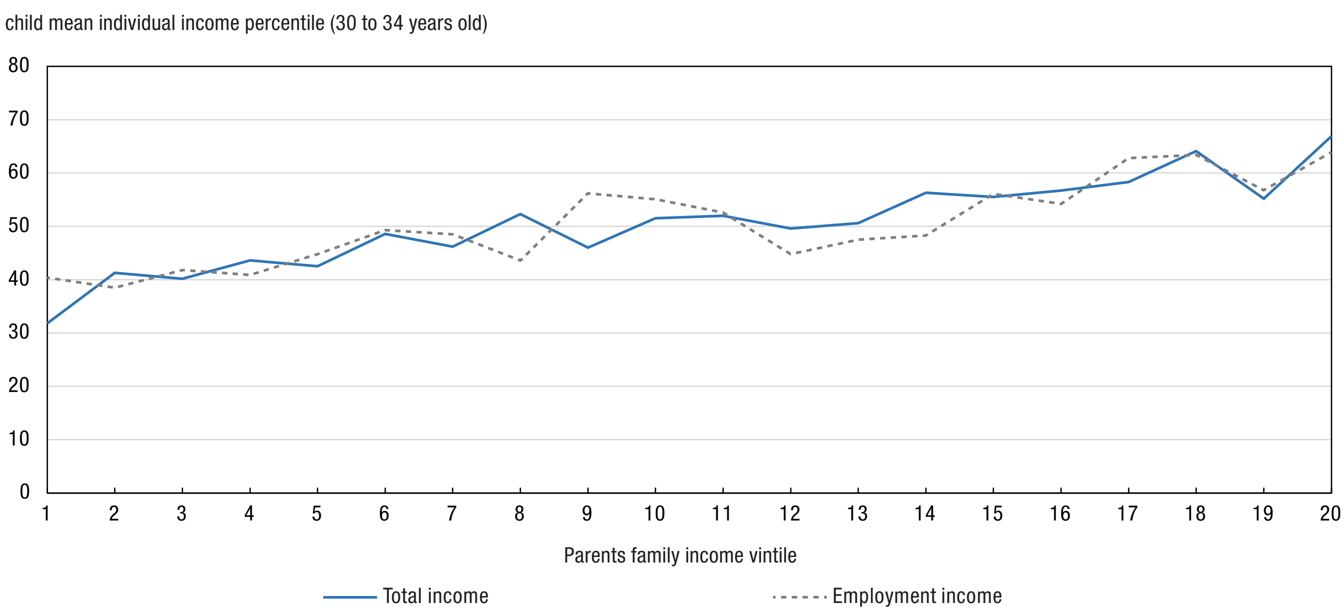
Note: All estimates use average parental family income when the child was 15 to 19, unadjusted for family size and child income averaged over five years between 2009 to 2013. Unadjusted family income is the sum of the income of both spouses, when present. Observations are included if their average income is positive ($1 or more), before family income is calculated.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Table 9
Quintile transition matrices, 1963 to 1979 birth cohorts

Parental income quintile	Child income quintile (mean child income at 30 to 34 years old)					
	Bottom	2nd	3rd	4th	Top	Total
Parents family income and child and spouse income, total income						
Bottom	31.9	24.0	20.1	14.0	10.0	100.0
2nd	20.4	24.8	18.3	22.6	13.9	100.0
3rd	19.7	20.7	20.3	22.4	16.9	100.0
4th	15.2	15.9	24.3	21.0	23.6	100.0
Top	12.8	14.5	17.2	19.9	35.5	99.9
Parents family income and child individual income, total income						
Bottom	31.8	23.7	19.8	14.3	10.5	100.1
2nd	18.7	23.0	18.7	22.3	17.3	100.0
3rd	17.8	20.8	24.9	19.3	17.2	100.0
4th	15.6	17.5	21.7	23.3	21.9	100.0
Top	16.2	15.0	14.9	20.9	33.0	100.0
Parents family income and child and spouse income, employment income						
Bottom	30.9	23.3	20.9	15.4	9.6	100.1
2nd	21.4	25.2	20.0	17.5	15.9	100.0
3rd	17.3	20.2	18.1	24.5	19.9	100.0
4th	15.8	20.6	23.3	22.6	17.7	100.0
Top	14.9	10.5	17.7	20.0	36.9	100.0
Parents family income and child individual income, employment income						
Bottom	28.9	27.4	18.2	13.8	11.6	99.9
2nd	18.3	22.3	20.2	23.1	16.1	100.0
3rd	18.9	19.7	22.4	20.3	18.7	100.0
4th	19.2	14.8	22.3	24.0	19.7	100.0
Top	14.7	16.0	16.7	19.0	33.6	100.0

Note: All estimates use average parental income (sum of both parents for couples) when the child was 15 to 19. Unadjusted family income is the sum of the income of both spouses, when present. Adjusted family income divides the sum of the income of both spouses by two, for couples. Observations are included if their average income has a value of $500 or more. Both members of a couple need to be above inclusion threshold for each of them to be included.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Chart 1
Non-linearity in the association between parental and child income rank by source of income, 1963 to 1979 birth cohorts



Note: Child mean individual income rank at 30 to 34 years old by source of income, 1963 to 1979 birth cohorts.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Table A1
Proportion of parent-child pairs dropped, by parental income inclusion condition

Condition	Labour earnings	Market income	Total income
Corak and Heisz (1999), 1963 to 1966 birth cohorts (father-son pairs)			
All years ≥ 1	31.9	17.0	…
Mean ≥ 1	19.4	6.0	…
Chen et al. (2017), 1963 to 1970 birth cohorts (father-son pairs)			
Parents in adequate age range (10+ years in dataset)	47.3	47.3	…
Among those in age range, 10+ years with income ≥ 500	24.3	10.0	…
Both inclusion conditions (10+ years in dataset with income ≥ 500)	60.1	52.6	…
Corak (2017), 1963 to 1970 birth cohorts (all parents-child pairs)			
Individual mean ≥ 500	…	…	7.8

.. not applicable

Note: The specifications are the same as in Tables 1 to 3. Parent-child pairs with children without income included. Excluding children without an income yields very similar results.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Table A2
Sample size of intergenerational income mobility models by child sex (child income averaged over 30 to 34 years old), 1963 to 1979 birth cohorts

Child	Labour earnings	Employment income	Market income	Total income
Both sexes	2,729.00	2,945.00	3,016.00	3,207.00
Daughter	1,400.00	1,503.00	1,544.00	1,658.00
Son	1,329.00	1,442.00	1,472.00	1,549.00

Note: Includes all observations matched with at least one parent of any gender. Observations are included if their average income has a value of $500 or more.
Source: Longitudinal and International Study of Adults, 2014, and T1FF, 1982 to 2013.

Appendix

Immigrants who arrived in Canada at age 15 or older

For most cohorts of interest, the first child-parent link is observed between 15-21 years of age. As a consequence, foreign-born LISA respondents who arrived in Canada after the age of 15 are less likely to be linked to a parent. The number of excluded respondents in the intergenerational sample is very small, given the low likelihood that they would meet the conditions to appear in the sample (both a parent and the child filing their taxes on the same year while residing at the same address), especially given that immigrants who arrived as adults might have arrived in Canada without their parents altogether or have never shared an address with them. However, studies focusing on immigration should consider excluding all immigrants who arrived in Canada after 15 years of age from their sample to account for the much lower linkage rate for this group of respondents. Unlike previous studies, the LISA data includes information on the year of arrival to Canada for foreign-born respondents, which can be used to perform that exclusion.

Summary of existing studies

Here, we systematically summarize sample design, how parents are identified and how measures of parental and child permanent income are operationalized in each of the four studies we attempt to replicate.

Corak and Heisz (1999)

Child selection

- Male children from the 1963-1966 birth cohorts.

Parent selection

- Father observed with their sons in 1982-1986 (among sons who were 16-19 years old in 1982). Only the first observed father is retained, if more than one are linked to a child over 1982-1986.

Parental income

- Employment and market income in 1978-1982, averaged.
- Exclusion option 1: Father excluded if average income (1978-1982) falls below $1.
- Exclusion option 2: Father excluded if income falls below $1, $100, $1,000 or $3,000 in any of the five years over which income is averaged (1978-1982).

Child income

- Employment and market income in 1995.

Chen, Ostrovsky and Piraino (2017)

Child selection

- Male children from the 1963-1966 birth cohorts (females also included in separate results).

Parent selection

- Father observed with their children in 1982-1986 (among sons who were 16-19 years old in 1982). Only the first observed father is retained, if more than one are linked to a child over 1982-1986.

Parental income

- Measure 1:

  Employment, market and total income in 1978-1982, averaged.

  Father excluded if income falls below $500 in any of the five years over which income is averaged (1978-1982).

- Measure 2:

  Employment, market and total income when father is between 35-55 years old averaged.

Father excluded if income less than ten years with income above $500 in the 21 (potential) years over which income is averaged (35-55 years old). Fathers who are not observed at certain ages because of left- or right-censoring are included in the sample (a father observed between 45 and 55 years old, with 10 years of income above $500, is included in the sample).

Child income

- Measure 1:

  Employment and market income at 30 or 40 years old.

- Measure 2:

  Employment and market income at 38-42 years old, averaged.

  Child excluded if income falls below $500 in three or more of the five years over which income is averaged (38-42 years old).

Corak (2017)

Child selection

- Children of any gender from the 1963-1966 and 1967-1970 birth cohorts.

Parent selection

- 1963-1966 birth cohorts: parent(s) observed with their children in 1982-1986 (among children who were 16-19 years old in 1982).
- 1967-1970 birth cohorts: parent(s) observed with their children in 1986-1990 (among children who were 16-19 years old in 1986). Only the first observed parent (or parents, if more than one is observed on the same year) is retained.
- If two parents are observed on the same year, they are considered to be a couple for the whole period over which parental income is calculated. Changes in marital status are not accounted for.

Parental income

- Total family income when children were 15-19 years old, averaged.
- For parents in a couple, their income is summed and the total divided by two.
- Parents are excluded if average income (when child is 15-19 years old) falls below $500. Exclusions are applied on parents separately before the sum is calculated.

Child income

- Total family income when children where 38-45 (1963-1970 birth cohorts) or 31-32 years old (1966-1970 birth cohorts).
- For children in a couple, their income is summed with their spouse's income and the total divided by two.
- Children or their spouses are excluded if average income (at 38-45 or 31-32 years old) falls below $500. Exclusions are applied on children and their spouses separately before the sum is calculated.
- If a child is first observed with a spouse, they are considered to be a couple for the whole period over which child income is calculated. Changes in marital status are not accounted for.

Connolly, Haeck and Lapierre (2019)

Child selection

- Children of any gender from the 1963-1985 birth cohorts separated in groups bounded by three year intervals (1963-66, 1967-70, 1972-75, 1977-80, 1982-85).

Parent selection

- Parent(s) of any gender observed when children were 16-19 years old on the following years: 1982, 1984, 1986, 1991, 1996, 2001.

Parental income

- Total family income when children were 15-19 years old, averaged.
- For parents in a couple, their income is summed.
- Parents are excluded if average income (when child is 15-19 years old) falls below $500. Exclusions are applied on parents separately before the sum is calculated.

Child income

- Total family or individual income when children were 25-29 years old, averaged (the analysis of the 1972-1975 birth cohort also includes results when children are 30-34 and 35-39 years old).
- When calculating family income for children in a couple, their income is summed with their spouse's income.
- Children or their spouses are excluded if average income over five year periods falls below $500. Exclusions are applied on children and their spouses separately before the sum is calculated.

References

Atkinson, A. B., Maynard, A. K., and Trinder, C. G. (1983). Parents and Children: Incomes in Two Generations. London: Heineman Educational Books.

Beach, C. M. and Finnie, R. (2004). A Longitudinal Analysis of Earnings Change in Canada. Analytical Studies Branch Research Paper Series. Catalogue no. 11F0019MIE – No. 227.

Becker, G. S., and Tomes, N. (1986). Human capital and the rise and fall of families. Journal of Labour Economics, 4(3, Part 2), S1-S39.

Blanden, J. (2005). Love and money: Intergenerational mobility and marital matching. Analytical Studies Branch Research Paper Series, Family and Labour Studies Division, Ottawa.

Chen, W. H., Ostrovsky, Y., and Piraino, P. (2017). Lifecycle variation, errors-in-variables bias and nonlinearities in intergenerational income transmission: new evidence from Canada. Labour Economics, 44, 1-12.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014a). Where is the land of opportunity? The geography of intergenerational mobility in the United States. Quarterly Journal of Economics, 129(4), 1553-1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014b). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. American Economic Review, 104(5), 141-147.

Connolly, Marie, Catherine Haeck and David Lapierre (2019), “Social mobility trends in Canada: Going up the Great Gatsby Curve.” Research Group on Human Capital Working Paper Series No. 19-03.

Corak, M., and Heisz, A. (1999). The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data. The Journal of Human Resources 34(3), 504-533.

Corak, M. (2017), “Divided Landscapes of Economic Opportunity: The Canadian Geography of Intergenerational Income Mobility”, Human Capital and Economic Opportunity Working Paper No. 2017-043.

Dahl, M. W., and DeLeire, T. (2008). The association between children’s earnings and father’s lifetime earnings: estimates using administrative data. University of Wisconsin-Madison, Institute for Research on Poverty.

Haider, S. and Solon, G. (2006), “Life-Cycle Variation in the Association between Current and Lifetime Earnings,” American Economic Review 96(4): 1308-1320.

Hemeon, J. (2016). Historical Data Linkage Quality: The Longitudinal and International Study of Adults, and Tax Records on Labour and Income. Ottawa: Statistics Canada (Catalogue no. 89-648-X).

Simard-Duplain, Gaëlle and Xavier St-Denis (2019a), “An exploration of the role of education in intergenerational income mobility in Canada: Evidence from the Longitudinal and International Study of Adults.” Working paper.

Simard-Duplain, Gaëlle and Xavier St-Denis (2019b), “Sample selection in tax data sets of intergenerational links: Evidence from the Longitudinal and International Study of Adults.” Working paper.

Solon, Gary (1992). Intergenerational income mobility in the United States. American Economic Review, 82(3), 393-408.

Statistics Canada (2018a). Dictionary, Census of Population, 2016. Ottawa: Statistics Canada (Catalogue no. 98-301-X2016001).

Statistics Canada (2018b). Longitudinal and International Study of Adults 2016 (Wave 3) User Guide. Ottawa: Statistics Canada.

Zarifa, D. (2012a). Choosing fields in an expansionary era: Comparing two cohorts of baccalaureate degree-holders in the United States and Canada. Social Stratification and Mobility 30(2012): 328-351.

Zarifa, D. (2012b). Persistent Inequality or Liberation from Social Origins? Determining Who Attends Graduate and Professional Schools in Canada's Expanded Postsecondary System. Canadian Review of Sociology 49(2): 109-137.

Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. American Economic Review, 82(3), 409-429.