Lifecycle Variation, Errors-in-Variables Bias and Nonlinearities in Intergenerational Income Transmission: New Evidence from Canada

Wen-Hao Chen[1], Yuri Ostrovsky[2] and Patrizio Piraino[3]

5 June 2015

ABSTRACT

This paper uses Canadian administrative data to test the impact of lifecycle earnings variation and errors-in-variables bias on estimates of intergenerational earnings and income mobility. We find lower levels of mobility compared to previous studies, with a new estimate of the father-son intergenerational earnings elasticity of 0.32. Our analysis also shows that the father-daughter elasticity is much less sensitive to these biases. We investigate whether improved measures of father/child permanent earnings may have a distinct impact on the estimated intergenerational persistence at different parts of the distribution. Taking advantage of exceptionally high sample sizes, we find that the impact of lifecycle bias is more pronounced at the top of the income distribution. We also document that much of the average intergenerational income persistence in Canada can be accounted for by limited mobility at the top, while mobility is significantly higher among children born to low-income fathers. These nonlinear patterns resemble those found in Northern Europe and are different from those observed in the United States.

JEL classification: J62, D31, D63

Keywords: Earnings inequality, intergenerational mobility

1.   Introduction

Comparative studies of intergenerational earnings and income mobility largely regard Canada as one of the most mobile countries among advanced economies (Bjorklund & Jantti, 2010; Corak, 2013; Solon, 2002). A popular chart (known as the Great Gatsby Curve), which is often used to show the relationship between income inequality and social mobility, depicts Canada as one of the countries with the highest mobility, at similar levels as Denmark, Finland and Norway.[1] The assertion that Canada is a highly mobile society is primarily drawn from intergenerational income elasticity (IGE) estimates reported in Corak and Heisz (1999) as well as in Fortin and Lefebvre (1998). While these are careful studies that offer the best estimates of the IGE based on the data available at the time of their writing, recent literature suggests that such estimates may be subject to bias due, in particular, to lifecycle variation (Mazumder, 2005; Haider and Solon, 2006; Bohlmark and Lindquist, 2006).

In the absence of data on lifetime earnings, IGE estimates may be affected by measurement error on both sides of the equation. Lifecycle bias (left-hand side measurement error) may arise in a father-child analysis when children's permanent earnings, proxied by yearly or short-run average earnings, are observed at a time of the child's working career when earnings do not closely correspond to their lifetime values. Haider and Solon (2006) show that this type of bias usually results in lower estimates of the IGE if earnings from the early (or too late) stages of the working careers are used. Their finding is consistent with reviews of the literature showing that the smallest IGE estimates are most often found in studies where sons' earnings are observed at younger ages (Solon, 1999). This raises a concern that the Canadian IGE—usually found to be in the neighbourhood of 0.2—may have been underestimated in earlier studies, as the evidence was based on data covering sons' earnings at a rather young age (29-32).

In addition to lifecycle bias, the IGE estimates may also be affected by the classical errors-in-variables bias (right-hand side measurement error), which occurs when fathers' lifetime earnings are not adequately measured. Corak and Heisz (1999), for instance, addressed this problem by using a 5-year average instead of single-year earnings to approximate fathers' lifetime earnings. While such practice reduces the errors-in-variables bias, it may not eliminate it entirely if the number of years used in the calculation of the averages is insufficient. Mazumder (2005) shows that even estimates based on 5-year average earnings are still subject to a significant errors-in-variable bias.

The primary objective of this study is thus to re-examine the extent of intergenerational earnings and income mobility in Canada in light of the estimation issues raised in the literature. Using longitudinal earnings data from a significantly higher number of years compared to previous research, the study highlights the sensitivity of IGE estimates to lifecycle and errors-in-variables bias. Importantly, the paper also revisits the issue of nonlinearity in the intergenerational transmission of earnings. Improved measures of fathers and child permanent earnings may have a different impact on the estimated intergenerational persistence at different parts of the distribution. Both polynomial and quantile regression specifications are employed to investigate this possibility. With nearly a quarter million observations, we are able to examine differences in the degree of intergenerational mobility across the full spectrum of the distribution.

---

[1] See (http://www.whitehouse.gov/sites/default/files/six_challenges_for_the_statistical_community.pdf). The term Great Gatsby Curve was first introduced in 2012 by Alan Krueger, the chairman of the U.S. Council of Economic Advisers, who used data from the work of Miles Corak.

The empirical analysis is based on an augmented version of the Intergenerational Income Database (IID), a high-quality dataset linking administrative records of parents and children in Canada. The data allow us to introduce two essential novel elements compared to the previous literature. First, the updated IID maintains the exceptionally high sample size of the previous versions while offering a significantly longer panel of tax records for children, who are now observed from their late teen years well into their mid-forties. This information is essential for testing the impact of lifecycle bias on the IGE estimates, as the literature shows that this bias can be greatly mitigated by comparing fathers' and offspring's earnings near their mid-career (Grawe, 2006; Haider and Solon, 2006, Gouskova *et al*, 2010). Second, instead of being averaged over an arbitrarily defined calendar time period, income in this study is averaged over a specified age range. With up to 22 years of valid data, a father's lifetime income is approximated by averaging the incomes he received between the ages of 35 and 55. This approach should significantly reduce (if not eliminate) the attenuation bias arising from measurement error in fathers' permanent incomes.

The results from our analysis suggest that the extent of intergenerational earnings and income mobility in Canada was overestimated in the early studies. The new estimate of the father-son earnings elasticity is about 0.32, which is noticeably higher than the values previously reported in the literature (in the neighbourhood of 0.2). Failing to account for lifecycle bias explains about two-thirds of the difference between the current and previous estimates, while errors-in-variables bias contributes to another one-third of the discrepancy. Interestingly, our analysis also shows that the father-daughter elasticity is much less sensitive to these biases. Finally, the paper finds that the impact of such biases is more pronounced at the top of the income distribution and documents a clear pattern of nonlinearity in the intergenerational transmission of earnings and income in Canada. In particular, the path to the top of the distribution appears to be quite challenging for sons born to low-income fathers. On the other hand, these same sons appear to have significant chances of moving into the middle class.

The remainder of the paper is organized as follows. The next section describes the data and the calculation of permanent earnings and incomes. Section 3 examines the sources of bias arising from imperfect measures of lifetime income and how these may affect estimates of the IGE. Section 4 presents the new estimates of intergenerational mobility in Canada for both father-son and father-daughter pairs. Section 5 addresses the issue of nonlinearity, while Section 6 concludes.

2.   Data and calculation of lifetime income

Our empirical analysis is based on matched parent-offspring tax records from the updated version of the Intergenerational Income Data (IID) file. A detailed description of the original IID file, which covered the period from 1978 to 1995, can be found in Corak and Heisz (1999). In essence, the IID consists of three sub-samples of children aged 16 to 19 in (i) 1982, (ii) 1984 and (iii) 1986 whose tax records are linked to the tax records of their parents based on their Social Insurance Numbers (SINs) and information from Statistics Canada's T1 Family File (TIFF). To improve coverage, child-parent pairs are drawn from the TIFF in all years between 1982 and 1986 and, if such links are available in multiple years, the earliest ones are retained. Once the child-parent link is established, it is possible to track children's and parent's annual earnings from 1978 to the latest available year (1995 in the original IID and 2008 in the updated file) using annual tax files (T1) and individuals' unique longitudinal identifiers based on their SINs.

One of the limitations of the original version of the IID was the relatively short number of adult years over which sons and daughters could be observed. For instance, even the oldest cohort of children—those born in 1963/66—could only be observed up to ages 29-32 (in 1995). Observations on the two other cohorts of children were limited to even younger ages. The recently updated version of the IID has the same structure as the original file—same three cohorts of children—but it extends the sample period up to 2008. Hence, the kids from the 1963/66 cohort can now be observed up to ages 42-45. The child cohorts born in 1965/68 and in 1967/70 can now be observed up to ages 40-43 and 38-41, respectively. In our analysis, we mainly focus on the 1963/66 cohort of children and their linked fathers.

It should be mentioned that not all existing father-child pairs can be identified in the IID. First, a generational link cannot be established if a child, still living with his/her parents, did not file a tax return in any year from 1982 to 1986. Second, generational links cannot be established for children linked to families that had no fathers between 1982 and 1986. Finally, father-child links cannot be established for children whose records could not be linked to any family.[2] Similar to Solon (1992) and Corak and Heisz (1999), we use only the oldest sons (daughters) when more than one son (daughter) is matched to the same father.

The analysis looks at three different income measures: earnings, market income and total income. Earnings are measured as a sum of wages/salaries from T4 slips issued by employers and other employment income including tips, gratuities and director's fees. Market income further includes rental income, self-employment income as well as asset income.[3] Total income refers to market income plus all government transfers, such as unemployment insurance benefits and pension benefits, but excluding taxes.[4] All monetary amounts are expressed in constant 2010 dollars.

An improved measure of lifetime earnings/income for fathers

It is well recognized in the literature that the use of current or single-year income as a proxy for permanent income can result in significant errors-in-variables biases in the estimation of the IGE (Solon, 1999). A common remedy is to use multi-year averages to reduce the transitory component of income. However, in the absence of a full history of lifetime data, the number of years over which researchers can observe incomes is rather limited, usually 3 to 5 years. While the literature seems to agree that multi-year averages are better measures of permanent income than single-year records, there has been relatively little discussion on the exact number of years required for the computation of the averages. Is taking 5-year averages sufficient to approximate for lifetime income? Based on a unique (but rather small) survey-administrative-linked data, Mazumder (2005) concludes that short-term proxies for fathers' permanent income may still be susceptible to bias as the variance of the transitory component of income varies considerably by age.

---

[2] As noted in Corak and Piraino (2011), the algorithm used to create the data leads to an under-representation of children from lower-income backgrounds and from the major metropolitan areas (Montreal, Toronto, and Vancouver). Studies that investigate the underreporting (Corak and Heisz, 1999; Oreopoulos, 2003; and Oreopoulos, Page, and Stevens, 2008) find little evidence of bias in their analytical results. In the empirical analysis below, we make use of Census-based weights created to account for this underreporting.

[3] Self-employment income includes 5 broad categories: business income, commission income, professional income, farming and fishing income. Asset income covers dividends and investment/interest income.

[4] See IID guide (2010) for a detailed list of income sources.

Moreover, the age at which fathers' incomes are averaged is sometimes overlooked in the literature. Corak and Heisz (1999), for instance, calculated fathers' permanent income by averaging their annual incomes over the period between 1978 and 1982. However, some fathers in the sample may be too young or too old during the observed period and therefore the averages may not properly capture their permanent income. Table A1 displays the age distribution of fathers in the IID. Indeed, nearly one quarter of fathers in the sample are arguably either too young or too old between 1978 and 1982. About 5.3% of the fathers were 34 or under at the beginning of 1978 and nearly 18% of fathers were 56 or more by the end of 1982. As expected, the earnings of these fathers are significantly lower compared to those of their prime-age counterparts. Even within the prime-age, there is some degree of variation: the mean earnings tend to be higher among fathers who were between 41 and 45 years old during the 1978-82 period.

To improve the measure of fathers' lifetime income, we average their annual incomes between the ages of 35 and 55, conditional on having positive values ($500 and over in constant 2010 dollars) in at least 10 of these 21 years. The age restriction ensures that the income averages for fathers in the sample are calculated at a similar stage of their lifecycles and therefore are less affected by the larger transitory components typical of the early/late stages of individuals' working careers. Restricting the sample to fathers with 10 or more (non-successive) years of positive income further reduces variation driven by few high or low income years within the prime age. Note that for each father in the IID, earnings and income data are available from 1978 to 1999. This implies that different cohorts of fathers will have different years over which their earnings can be averaged (see Table A2 for an illustration).[5] In order to have at least 10 positive annual records that satisfy both the age (35-55) and the calendar time (1978-1999) restrictions, only fathers born between 1932 and 1955 are included.[6] It is important to point out that the proposed measure is regarded as an improved proxy for lifetime income, rather than its "true" measure. The final sample consists of 356,321 fathers (with positive lifetime earnings) that can be matched to sons/daughters records.[7] Among these fathers, 56.4% are from the 1932-1938 cohorts and 40.2% are from 1939-1945 cohorts. Only 3.4% were born between 1946 and 1955.

3.   Empirical testing for lifecycle bias and errors-in-variables bias

Using U.S. data, Haider and Solon (2006) show that the bias arising from lifecycle variation can be greatly reduced if children' earnings are measured at their mid-careers. However, determining the age that minimizes the lifecycle bias is an empirical task. Importantly for our purposes, such age may be country-specific. In this section, Haider and Solon's generalised errors-in-variables model is applied to the IID data. It should be emphasized that the purpose of this exercise is not to formally examine the association between annual and lifetime earnings, as done by Haider and Solon (2006), since this would

---

[5] For instance, the incomes of a father born in 1943 can be averaged from 1978 to 1998 (when he was between 35-55 years old). A younger (older) father, such one born in 1950 (1935), was 35-55 years old in 1985-2005 (1970-1990). In this case, only up to 15 (13) years of data will be used to compute the average, provided that positive annual earnings are observed in at least 10 years.

[6] The oldest fathers included in the analytical sample are those born in 1932, as they were 46 years old in 1978 (the first earnings year available) and they will be in the sample only if they received positive earnings in *each* year between the ages of 46 and 55. Similarly, the youngest fathers that could be included are those born in 1955, as they were 44 years old in 1999 (the last earnings year available) and they will be in the sample only if they received positive earnings in *each* year between the ages of 35 and 44.

[7] In the case of market income and total income, the number of included fathers is much higher, at 489,410 and 502,580 respectively.

require full life-long earnings histories not available in the IID.[8] In addition, the analysis in this study is conditional on the sample of men who were inter-generationally linked at a certain point in their life, which may not be comparable to the broader samples used in studies that focused mainly on lifecycle earnings variation. As a consequence, this section is only intended to illustrate the bias in IGE estimates when lifetime incomes of fathers and offspring are not adequately measured.

Following the literature standards, a simple model describing the relationship between fathers' and children's incomes can be written as:

$$Y_i^s = \alpha + \beta Y_i^f + \varepsilon_i \tag{1}$$

where Y_i^s (Y_i^f) is the log of child's (father's) lifetime income, ε_i is a random error uncorrelated with Y_i^f, and β captures the intergenerational income elasticity (IGE). Since permanent income is usually not available for either fathers or children, replacing them by yearly data or even multi-year averages could result in biased estimates of the IGE due to measurement errors on both sides of the equation. These are generally known as lifecycle bias and errors-in-variables bias. To illustrate this, the analysis follows Haider and Solon (2006) and defines the left-hand side income measure (children's earnings in our case) as:

$$Y_{it} = \lambda_t Y_i + u_{it} \tag{2}$$

where Y_i is the lifetime income proxied by Y_{it} (annual earnings at age t), u_{it} is a random disturbance uncorrelated with Y_i and ε_i, and λ_t is the slope coefficient in the so-called "forward regression" of Y_{it} on Y_i. Suppose one wishes to estimate (1). Substituting (2) into (1) gives

$$Y_{it} = a + \lambda_t \beta Y_i^f + (\lambda \varepsilon_i + u_{it}) \tag{3}$$

so $\lambda_t \beta$ (instead of β) is the probability limit of the estimated coefficient on Y_i^f. As a result, the OLS estimator of (1) is consistent only if $\lambda_t = 1$. A lifecycle bias thus arises when $\lambda_t \neq 1$, which may vary with the age at which incomes are observed. Haider and Solon (2006) demonstrate that for the U.S. λ_t profiles vary notably across the lifecycle: the profile begins at 0.24 at age 19, rises to about 1 at age 32, and then declines towards the end of the working career.

Figure 1 presents estimates of λ_t using the IID data described above. They are slope coefficients in the forward regression of log annual earnings at age t on log lifetime earnings (equation 2). As explained above, lifetime earnings are calculated by taking average earnings between the ages of 35 and 55 (conditional on positive values in at least 10 years). This proxy may be considered an upper bound estimate of lifetime earnings since it excludes low-earnings years (i.e. early and late parts of the lifecycle). Nonetheless, to the extent that the prime-age earnings are more representative of the lifetime values (Haider and Solon, 2006; Bohlmark and Lindquist, 2006) the coefficients from the forward regressions can provide some guidance on the possible presence of lifecycle bias in previous estimates of the IGE in Canada. The results in Figure 1 indicate that λ_t does not equal 1 throughout the lifecycle. It

---

[8] A related Canadian study is Baker and Solon (2003) who used a representative sample from tax file to examine life-cycle patterns in earnings dynamics. They show that the variance of transitory log-earnings varies significantly over the life cycle.

begins at 0.47 at age 25, increases gradually and reaches unity in the early forties. It continues to increase to around 1.15 at age 52 and falls back to below 1 after age 56. An important implication of Figure 1 is that the estimates of the IGE could be significantly underestimated when sons' earnings are observed at younger ages. This finding appears to justify our intention to revisit the previous Canadian estimates in Corak and Heisz (1999) as their data was based on sons' earnings measured between ages 29 and 32. Figure 1 also suggests that any bias arising from lifecycle variation may be mitigated when sons' earnings are measured somewhere around the late thirties or early forties (taking into account the fact that our permanent income measure may overestimate lifetime earnings).

Figure 1: Estimates of λ_t (lifecycle bias)
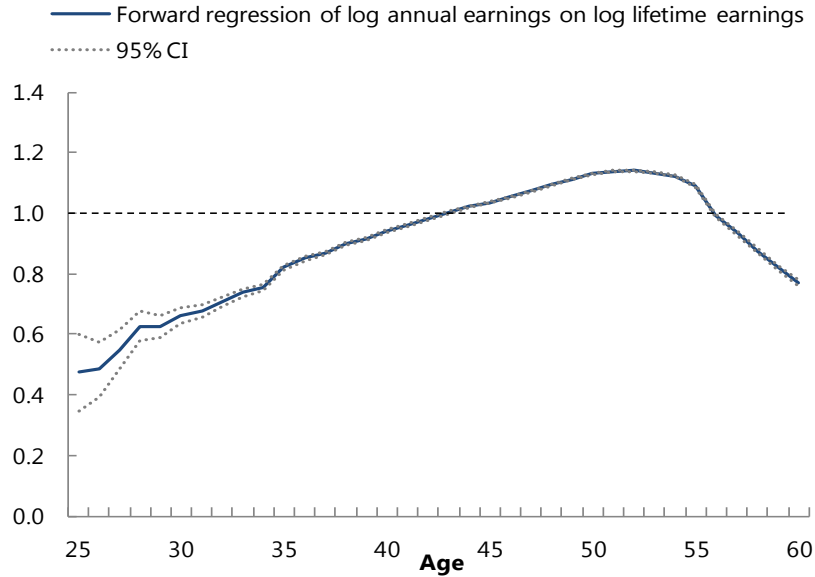
Note: The sample includes individuals with positive lifetime earnings. For each individual in the sample, lifetime earnings are calculated by averaging one's income between the ages of 35 and 55, conditional on having positive income ($500 and over) in at least 10 years. All earnings are CPI adjusted in 2010 constant dollars.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

Haider and Solon (2006) also address the right-hand side measurement error. The unobserved regressor, Y_i, (fathers' lifetime earnings in our case) is proxied by annual earnings at age t, Y_{it}:

$$Y_i = \theta_t Y_{it} + v_{it}. \qquad (4)$$

Equation (4) is called the "reverse regression" of Y_i on Y_{it}, and the probability limit of the estimated slope coefficient is

$$plim\hat{\beta} = \frac{Cov(Y_{it}^f, Y_i^s)}{Var(Y_{it}^f)} = \theta_t \beta \qquad (5)$$

where

$$\theta_t = \frac{Cov(Y_{it}^f, Y_i^f)}{Var(Y_{it}^f)} = \frac{\lambda_t Var(Y_i^f)}{\lambda_t^2 Var(Y_i^f) + Var(v_{it})}. \qquad (6)$$

When $\lambda_t = 1$, θ_t is simply the textbook case of attenuation bias. However, θ_t will also depend on the value of λ_t. Haider and Solon (2006) argue that in rare cases, θ_t can turn out to be an amplification rather than attenuation bias.

We estimate the trajectories of θ_t for Canadian men using again our measure of lifetime earnings. As expected, using annual earnings to approximate lifetime values on the right-hand side of the regression results in a significant attenuation bias (solid line). The bias is especially pronounced ($\theta_t = 0.25$) when current earnings are observed early or late in the lifecycle and remains large (about 0.55) if earnings are measured at the mid-career years. These findings are in line with Bjorklund (1993) and Bohlmark and Lindquist (2006) for Sweden as well as with Haider and Solon (2006) for the United States.

Figure 2: Estimates of θ_t (attenuation bias)
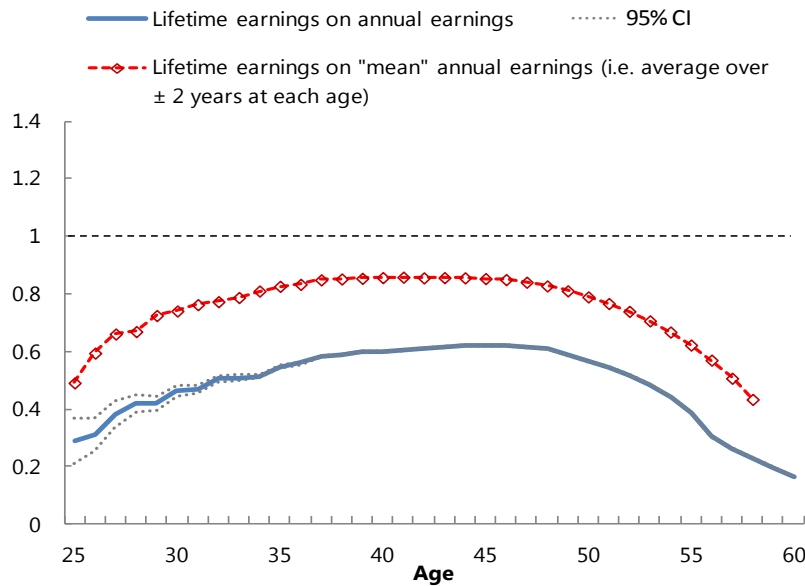
Note: Sample includes those with positive lifetime earnings. For each individual in the sample, lifetime earnings are calculated by averaging one's income between the ages of 35 and 55, conditional on having positive income ($500 and over) in at least 10 years. All earnings are CPI adjusted in 2010 constant dollars.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

In practice, researchers often use multi-year averages instead of yearly earnings as a proxy for lifetime earnings. To what extent do such practices reduce the errors-in-variables bias? To answer this question, the reverse regressions are re-estimated with current annual earnings replaced by *mean* annual earnings within a five-year window centered at any given age.[9] The dashed line in Figure 2 reveals that the attenuation bias can be significantly mitigated when 5-year averages are used to proxy for lifetime earnings instead of annual earnings. The estimated θ_t can be as high as 0.85 when mid-career multi-year

---

[9] For instance, mean earnings at age 30 will be a 5-year average between the ages of 28 and 32. Similarly, mean earnings at age 55 are averaged over the ages of 53-57.

earnings averages are used. Figure 2 also appear to support Mazumder (2005), who argues that even estimates based on 5-year averages may still be subject to non-negligible errors-in-variables bias. All in all, results from both Figures 1 and 2 seem to suggest that the Canadian IGE is likely to have been underestimated in previous studies.

International comparison of bias profiles

In this subsection, the estimated Canadian bias profiles are compared with those of other countries. The comparison may shed some light on the sources of cross-country differences in the intergenerational transmission of earnings. There are a limited number of studies investigating the association between current and lifetime incomes. Haider and Solon (2006) was the first to offer empirical estimates of λ_t and θ_t, based on nearly career-long earnings histories from a U.S. panel. Two other papers known to the authors have been able to replicate Haider and Solon's approach for other countries: these are Bohlmark and Lindquist (2006) for Sweden and Brenner (2010) for Germany, using data from Swedish tax records and the German Socio-Economic Panel (GSOEP) respectively.[10]

The estimates of λ_t and θ_t from these studies are shown in panel (A) and panel (B) of Figure 3 respectively. To improve comparability, we restrict Canadian men to those born between 1939 and 1945—similar to the German and Swedish cohorts but still about a decade younger than the American cohort. The Canadian profile is shorter as only up to 21 years of earnings data are available in the IID. In general, with respect to the estimated λ_t, all four studies show that the textbook scenario of $\lambda_t=1$ throughout the lifecycle does not apply. As a consequence, using annual earnings from early career stage to proxy for lifetime earnings as the dependent variable would lead to an attenuation bias.

There is also a similarity in terms of where (or at what age) the lifecycle bias is minimized. The estimate of λ_t seems to approach 1 at the age of 40 for Canadian men but at a somewhat younger age—at around mid-thirties—for the other 3 countries. This may be due to the fact that our measure of lifetime earnings for Canada is overestimated by definition (i.e. because individuals' earnings are averaged over their prime age). However, 95% confidence intervals around these profiles do overlap over the mid-career, close to the age of 40-41. After 41, the bias profiles for each country become more distinct. For the United States, the lifetime earnings are underestimated from the mid-career on, while the opposite seems to be the case for Germany where λ_t continues to grow until the age of 44 and remains above unity throughout the rest of the career. The Canadian and Swedish profiles are somewhat similar in the sense that the estimated λ_t in both countries peaks at around age 50 and then declines thereafter with a more, pronounced drop for Canadian men (which, again, may be due to the particular measure used here).

These patterns are insightful for international comparisons of the intergenerational transmission of earnings. First, even if sample and methodological differences are accounted for, cross-national variation in IGE estimates may still arise due to differences in the age at which sons' earnings are observed, everything else being equal.[11] For instance, using sons' earnings corresponding to the ages

---

[10] Nynom and Stuhler (2011) also offer estimates of bias profiles for Sweden using the same data as Bohlmark and Lindquist (2006).

[11] To address cross-country comparisons of the intergenerational mobility, there was some attempt in Corak (2006) to recognize methodological differences as well as age differences in children. Nynom and Stuhler (2011) further detail a few additional reasons related to lifecycle income profiles that make cross-country comparisons of intergenerational mobility difficult. In particular, they argue that inconsistencies may stem from the interaction of

between mid-forties and mid-fifties tend to induce an attenuation bias in the U.S, but an amplification bias in the other three countries. On the other hand, lifecycle bias seems to be greatly reduced when sons' earnings are measured around the age of 40 in all countries.

For the trajectory of θ_t (Panel B), all countries exhibit a similar inverted U-shaped profile over the lifecycle. Using current earnings to proxy for lifetime earnings as the independent variable will lead to downward bias when earnings from early and late stages of the working careers are used. In all cases, the bias is smaller around the mid-career, but θ_t remains far below unity, suggesting that the bias arising from the right-hand side measurement error cannot be eliminated completely at any stage of the working life. It is interesting to note that since λ_t is approximately 1 for all four countries at about age 40, θ_{40} captures the attenuation bias due to classical measurement error. Again, cross-national differences in θ_t are more visible at the beginning and end of the working career.

Figure 3: Cross-national comparison of bias profiles for men

(A) Estimates of λ_t
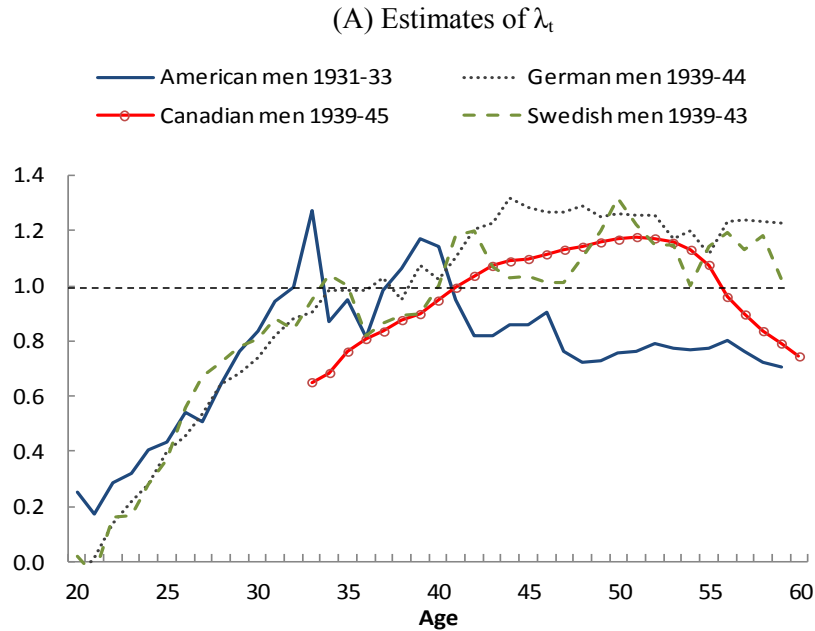
two factors: unobserved heterogeneity in income profiles (e.g. returns to schooling) and idiosyncratic deviations from average profiles that are correlated with individual and family characteristics (e.g. stochastic income shocks to consumption).

(B) Estimates of θ_t
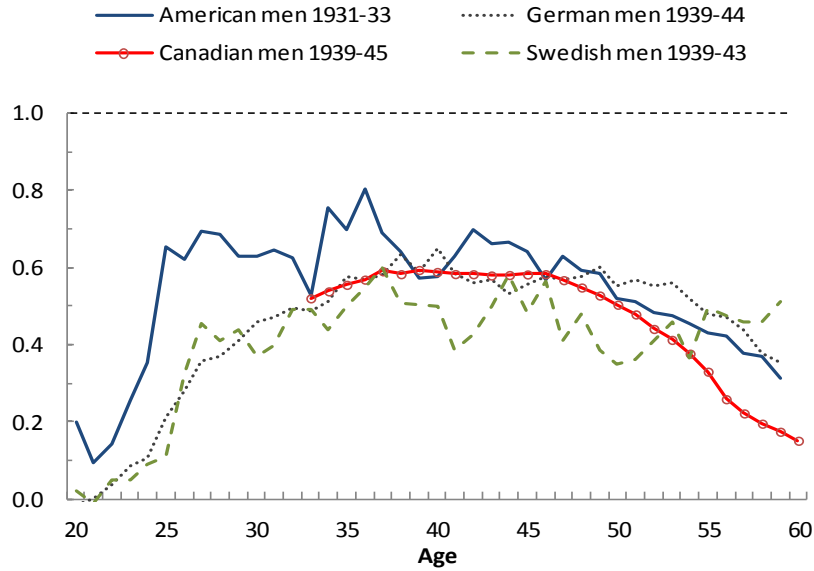
Source: Haider and Solon (1996) for the U.S., Brenner (2010) for Germany, Bohlmark and Lindquist (2006) for Sweden, Canadian results are from Statistics Canada's Intergenerational Income Data (IID).
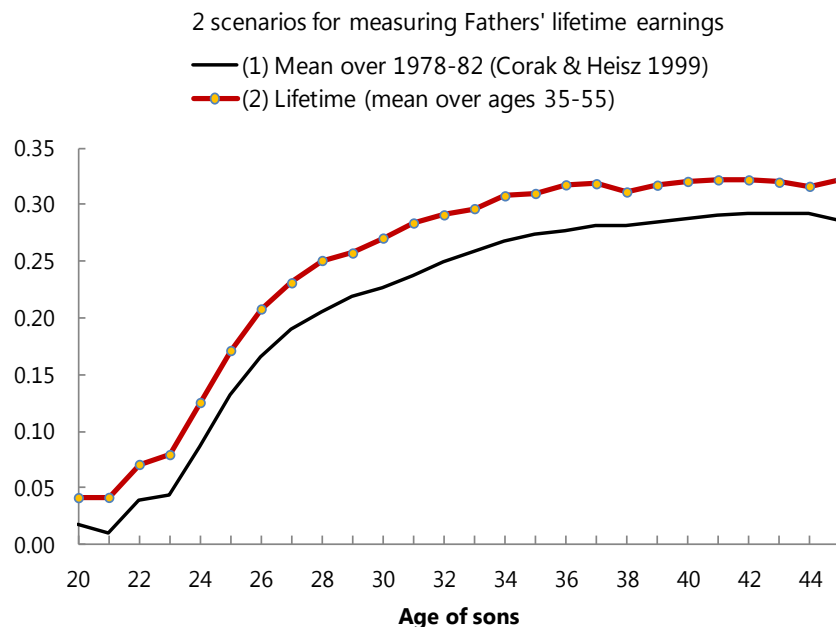
4.   New evidence of intergenerational earnings/income mobility in Canada

To what extent are the estimates of the intergenerational elasticity in Canada affected by the use of inadequate proxies for both fathers' and sons' lifetime earnings? In order to answer this question, two different scenarios using alternative measures of fathers' lifetime earnings are presented in this part of the analysis. To compare our results to previous findings, the first scenario follows Corak and Heisz (1999) and defines fathers' earnings as five-year averages over the period from 1978 to 1982.

The solid blue line in Figure 4 presents estimates of the IGE obtained by running an OLS regression defined by (1) for different ages at which son's earnings are observed. Overall, the estimated IGE follows a concave trajectory, reflecting lifecycle differences in the magnitude of the bias from left-hand side measurement error. When sons' earnings are observed at around age 30, as in Corak and Heisz (1999), the model produces an estimate of β (IGE) that is nearly identical to the one in their study—about 0.227.[12] The estimated elasticity continues to increase as sons' earnings are measured at older ages. It rises to 0.29 when earnings are observed in the early forties (where λ_t approaches 1), indicating a higher degree of intergenerational earnings persistence.

---

[12] This can be compared to 0.227-0.237 in Corak and Heisz (1999, Table 3A, specifications 3-4). Note that in our study, we restrict the sample to fathers whose earnings in each of five years (1978-82) are equal or greater than $500 constant dollars. In a more recent paper Corak *et al.* (2014) pointed out that 0.25 is the preferred Canadian estimate based on an updated IID where sons' earnings are measured between 33 and 36 years of age.

Figure 4: Estimates of β (IGE) by the age of sons when earnings are observed

2 scenarios for measuring Fathers' lifetime earnings
- (1) Mean over 1978-82 (Corak & Heisz 1999)
- (2) Lifetime (mean over ages 35-55)

Note: For scenario (1), lifetime earnings are calculated by averaging one's income between the years/ages indicated, conditional on having positive income ($500 and over) in at least 3 years. *see footnote in Figure 1.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

In addition to lifecycle bias, estimates of the IGE also depend on how the fathers' lifetime earnings are calculated. The second scenario uses the measure for father's lifetime earnings introduced in this study in an attempt to minimize the right-hand side measurement error. In general, the second scenario (blue line) produces higher earnings elasticity compared to the first specification. The estimated β now increases by another 3 points to 0.32 when sons' earnings are observed in their mid-career. Figure 4 confirm that both left- and right-side measurement errors can produce a downward bias in the IGE estimates. In particular, the left-hand side lifecycle bias can be substantial if the estimated IGEs are based on sons' earnings from the very early years of their working careers.

4.1   Are Canadians as mobile as previously suggested?

Table 1 presents new estimates of the intergenerational elasticity for Canada. Again, different scenarios for measuring lifetime earnings of fathers and sons are offered in order to grasp the possible impact of left- and right-hand side measurement errors on the estimated coefficients. The baseline scenario (1) uses a definition of lifetime earnings similar to the one in Corak and Heisz (1999); scenario (2) mitigates left-hand side lifecycle bias by observing sons' earnings at age 40; scenario (3) further reduces the right-hand side attenuation bias by using fathers' earnings averaged over their prime age (from 35 to 55); and the last scenario—our preferred one—improves the precision of sons' lifetime earnings with the use of 5-year averages around their mid-careers (i.e. ages 38-42). In addition to earnings, the results are presented for both market and total incomes (before taxes, after transfers).

Focusing on scenarios (3) and (4), the father-son intergenerational earnings elasticity in Canada is about 0.32—about 46% (or 10 points) higher than the commonly quoted estimate of 0.22 from Corak and Heisz (1999). Both sources of bias (lifecycle and errors-in-variables) contribute to the under-estimation in the early studies. In general, failing to account for the lifecycle bias explains about two-thirds of the difference between the current and previous estimates, while the use of a less accurate proxy for fathers' lifetime earnings contributes to another one-third of the discrepancy. Comparing specifications (4) with (3) in Table 1 also suggests that when sons' earnings are observed at the "right" age (i.e. around their forties), replacing them by 5-year averages does not significantly change the estimate of the IGE. This may suggest that there is a lesser need to approximate sons' lifetime earnings with multi-year averages as long as earnings data at their mid-careers can be observed.

Table 1. Estimates of intergenerational elasticity, fathers and sons

	β (Fathers-Sons)		
	Earnings	Market income	Total income
Proxy for lifetime incomes			
(1)			
Kids: at age 30	0.227	0.230	0.222
Fathers: mean over 1978-82	(0.003)	(0.003)	(0.002)
(2)			
Kids: at age 40	0.287	0.301	0.317
Fathers: mean over 1978-82	(0.004)	(0.003)	(0.003)
(3)			
Kids: at age 40	0.321	0.349	0.359
Fathers: mean over ages 35-55[1]	(0.004)	(0.003)	(0.003)
(4) preferred model			
Kids: mean over ages 38-42[2]	**0.318**	**0.343**	**0.359**
Fathers: mean over ages 35-55[1]	**(0.004)**	**(0.003)**	**(0.003)**
Father-child pairs (N)	**196,422**	**246,350**	**261,871**

Note: Standard errors in parentheses.
[1] Fathers with positive earnings/incomes (>=$500) in at least 10 years during the ages of 35-55.
[2] Sons with positive earnings/incomes (>=$500) in at least 3 years during the ages of 38-42.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

4.2 Market and total income mobility

In addition to earnings, the IID also includes other income sources that allow researchers to examine the intergenerational transmission of market as well as total income. Understanding *income* mobility can offer additional insight into transmission mechanisms across generations. Previous research, for instance, has shown clear evidence of the intergenerational transmission of jobs (Kramarz and Skans, 2014) self-employment (Lentz and Laband, 1990, Dunn and Holtz-Eakin, 2000, Sorensen, 2004), chief executive officer positions (Perez-Gonzalez, 2006), liberal professions (Aina and Nicoletti, 2013) and employers (Corak and Piraino, 2011). Aside from financial transfers, these articles also emphasize the importance of other transmission mechanisms, such as transmission of human capital, networking, and in some cases, nepotism. Dunn and Holtz-Eakin (2000), for example, point out that parents' own entrepreneurial experience and business success has a significant effect on the propensity of becoming self-employed.

This literature suggests that intergenerational persistence should be higher for market income, since this includes self-employment and asset incomes. Column 2 of Table 1 confirms this expectation. The sample here further includes father-son pairs who have positive self-employment or asset income. The intergenerational elasticity for fathers and sons now increases by another 3 points to around 0.35 in the preferred model. It is also noteworthy that estimates based on scenario 1 significantly underestimate the dynamics of market income. A possible explanation is that self-employed children may have not yet started (or just started) their own business when they are in their early thirties. Alternatively, or perhaps in addition to that, asset income may grow faster over time for children from more affluent backgrounds. As a result, the bias arising from lifecycle variation can be more pronounced for market income than for earnings.

Moving on to total income, which also includes government transfers (except taxes), brings us a step closer to measuring the extent of intergenerational transmission of economic well-being in Canada. Individuals, both fathers and children, who are less attached to the labour market are now included in the analysis as long as they received transfers from governments at any level (i.e. local or national). The literature has shown a strong intergenerational correlation in the receipt of government assistance (Corak *et al.*, 2004; Page, 2004). If children of low-income fathers are more likely to receive government assistance, this may be reflected in the IGE estimates. However, the last column of Table 1 reports that the intergenerational persistence is only marginally higher for total incomes, being now close to 0.36.

To sum up, the findings presented in Table 1 show that measuring earnings of both fathers and sons in a way that minimizes the impact of transitory earnings and lifecycle bias leads to an estimated IGE for Canada of about 0.32. The elasticity is even higher when market and total income are considered. It is important to emphasize that while the new finding would put Canada in the middle of the international spectrum of IGE estimates (Bjorklund & Jantti, 2010; Corak, 2013; Solon, 2002), very few countries have data allowing an analysis similar to the one presented here. Probably the most comparable estimates to the coefficients presented in this paper are the results in Mazumder (2005). He reports an estimated IGE just above 0.6 in the United States when applying sample restrictions similar to the ones used here. Compared to Mazumder's results, our estimates confirm that Canada remains significantly more mobile than the United States. However, the "true" rate of intergenerational transmission may not be as low as previously thought.

4.3   Gender differences in intergenerational income mobility

Table 2 replicates the analysis using father-daughter pairs. In general, the intergenerational transmission of earnings and income is weaker for daughters than for sons. The estimated elasticity is about 0.23 for earnings and 0.25-0.26 for income. This result seems to suggest that daughters' outcomes are less dependent on the earnings/incomes of their fathers. The results from the baseline scenario (1) can be compared with early Canadian studies (Fortin and Lefebvre 1998, Corak 2001), which also found the intergenerational elasticities of earnings/income for fathers and daughters to be around 0.2. Interestingly, unlike the father-son IGE, the father-daughter IGE does not seem to be affected by lifecycle variation. In fact, estimates based on daughters' earnings at age 30 (scenario 1) are very similar to those based on daughters' earnings at age 40 (scenario 2).

While this may appear puzzling at first, there may be several factors that can help explain this result. Typically, women are more likely than men to experience career breaks related to child-bearing and child-rearing during the early stages of their working life. More generally, women are less attached to the labour market compared to men. The findings are also consistent with results from other countries regarding the role of "assortative mating", which has been identified in the literature as one of the possible reasons for lower IGEs for daughters (see Chadwick and Solon 2002, Ermisch *et al.* 2005, Raaum *et al.* 2007). In the presence of marital sorting, daughters with high earnings potential are more likely to marry to high-earnings husbands. Such daughters could choose to work fewer hours or accept lower pay in exchange for better work-family balance. In the presence of assortative mating, fathers' lifetime earnings may be more closely tied to daughters' *family* (including spousal) earnings and not to their own earnings.

This could help explain why the estimated father-daughter IGE does not rise as much as it does for sons when earnings are measured at their forties. Although the estimates in scenario (1) may be biased downward due to lifecycle variation, they may be less affected by assortative mating as many daughters remain as single at age 30. On the other hand, at age 40 most daughters are already married so low estimated coefficients in scenario (2) points to the possibility that assortative mating may be playing a role in the intergenerational transmission of income between fathers and daughters in Canada. They also suggest that future studies focusing on the extent of intergenerational transmission of income for daughters may need to consider *family* incomes.

Table 2. Estimates of intergenerational elasticity, fathers and daughters

	β (Fathers-Daughters)		
	Earnings	Market income	Total income
Proxy for lifetime incomes			
(1)			
Kids: at age 30	0.191	0.212	0.213
Fathers: mean over 1978-82	(0.004)	(0.003)	(0.003)
(2)			
Kids: at age 40	0.195	0.207	0.222
Fathers: mean over 1978-82	(0.004)	(0.004)	(0.003)
(3)			
Kids: at age 40	0.221	0.234	0.249
Fathers: mean over ages 35-55[1]	(0.005)	(0.005)	(0.004)
(4) preferred model			
Kids: mean over ages 38-42[2]	**0.228**	**0.241**	**0.254**
Fathers: mean over ages 35-55[1]	**(0.004)**	**(0.004)**	**(0.003)**
Father-child pairs (N)	**160,462**	**205,921**	**216,214**

Note: Standard errors in parentheses.
[1] Fathers with positive earnings/incomes (>=$500) in at least 10 years during the ages of 35-55.
[2] Daughters with positive earnings/incomes (>=$500) in at least 3 years during the ages of 38-42.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

5.   Nonlinearities

The results from the linear regression analysis above may mask nonlinear patterns in the intergenerational transmission of earnings. Various hypotheses have been put forward to explain nonlinearities in the relationship between fathers' and son's log earnings. For instance, the human capital model proposed by Becker and Tomes (1986) implies a concave pattern of intergenerational earnings transmission, as families at the bottom of the earnings distribution may be more likely to be borrowing constrained. Empirical evidence related to this hypothesis, however, is rather mixed. In fact, studies of the Nordic European countries find instead a pattern of convexity. Bratsberg *et al.* (2007), for instance, show that intergenerational earnings persistence in the Nordic countries is highly nonlinear, with greater mobility found at the bottom of the distribution. Similarly, Björklund *et al.* (2012) report a very high intergenerational persistence of top incomes in Sweden with an estimated elasticity of about 0.9 compared to an average value of 0.26.

The following section presents results related to nonlinearities in intergenerational persistence in Canada. The analysis in the previous section indicates that the channels of intergenerational income transmission seem to be more complicated for daughters than for sons and may require taking spousal incomes into account. Therefore, the analysis below is restricted to the father-son sample only. With nearly 200,000 father-son pairs in the IID, it is possible to examine potential nonlinearities at even finer level of detail than the previous Canadian studies.

5.1   *Descriptive correlation*

The issue of nonlinearity is first examined using a descriptive and intuitive method well-suited to illustrate the relationship between fathers' and sons' earnings across the entire spectrum. Fathers are ranked and divided into percentiles according to their lifetime earnings. This generates 100 data points for scatter plots. For each percentile of fathers' earnings, Figure 5 shows the mean log earnings (Panel A) and the earnings shares of sons and fathers (Panel B). The pattern in Panel (A) clearly shows that the relationship between fathers and sons earnings in Canada is not linear. In particular, it reveals a somewhat convex pattern. The estimated slope coefficient is 0.36, but the profile is flat at the bottom of the distribution, increases monotonically over the main part of the distribution and becomes steeper at the top. This pattern has some interesting implications. First, earnings mobility seems to be high among children born to the lowest 15 percentiles of the fathers' earnings distribution as many of them find themselves in higher earnings percentiles than their fathers. Second, earnings persistence tends to be very high among sons born to the top 10% of the fathers' distribution, suggesting a significant degree of intergenerational transmission of advantages at the top. Third, between the 15th and 90th percentiles of the fathers' distribution the intergenerational earnings persistence is moderate.

The mean earnings in a percentile tell us little about earnings variation within the percentile. Panel (B) presents earnings shares held by sons and fathers in each percentile of the fathers' earnings distribution. If sons' lifetime earnings are independent of their fathers', we would expect earnings shares held by sons in each of the fathers' percentiles to be close to 1%. Conversely, a profile that looks like a 45-degree straight line would indicate a high degree of persistence between fathers and sons' earnings. Panel (B) generally confirms the nonlinear pattern of intergenerational transmission of earnings noted above. Again, mobility appears to be substantial at the bottom of the distribution. For instance, the

earnings share held by fathers in the lowest 1% of the distribution is only 0.1%, while the earnings share of sons from these families amounted to nearly 0.8% of total sons' earnings. The correlation is stronger in the middle part of the distribution as both sons' and fathers' earnings shares tend to move in the same direction.

Figure 5: The relationship between fathers' and sons' earnings
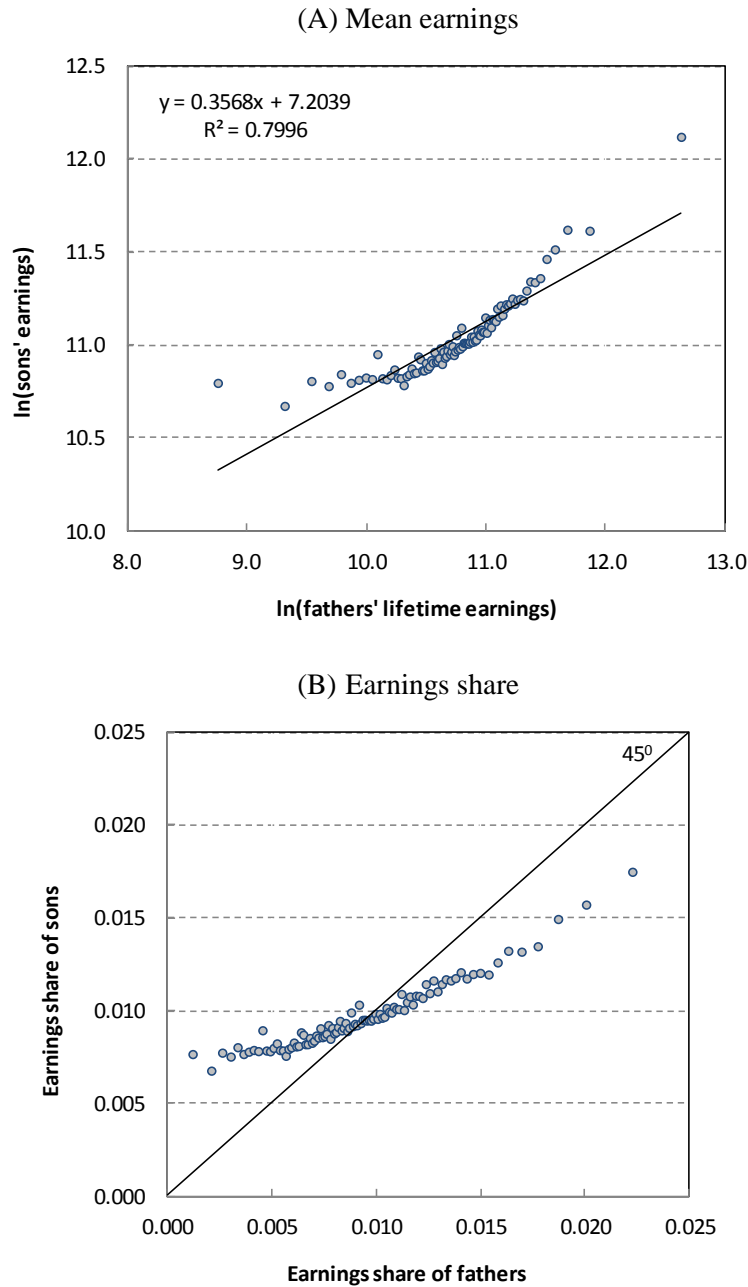
(A) Mean earnings

$y = 0.3568x + 7.2039$
$R^2 = 0.7996$

(B) Earnings share

Note: Lifetime earnings are calculated by averaging fathers' (sons') earnings between the ages of 35 and 55 (38 and 42), conditional on having positive income, $500 and over, in at least 10 (3) years.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

The patterns of nonlinearities observed in Figure 5 do not appear consistent with the borrowing constraints model in Becker and Tomes (1986) or Mulligan (1997), which would imply a concave relationship between children's and parent's earnings. These studies suggest that parents in the lower half of the distribution are more likely to be credit-constrained. Under some additional assumptions, this would imply sub-optimal investments in children's human capital at the bottom of the fathers' earnings distribution with a resulting stronger correlation (i.e. steeper profile) between father and son earnings. Credit constraints are then assumed to gradually relax at higher percentiles of the distribution and, as a result, sons' earnings would become more independent from fathers' earnings in the upper half of the distribution.

A possible reason why a pattern of concavity may not be found in the data is the violation of some of the model's assumptions. Han and Mulligan (2001), for instance, show that heterogeneity in children's innate earnings potential, as well as in parents' altruism, make testing for the existence of credit-constrained families difficult. The pattern of nonlinearity observed in the Canadian data, however, seems to be more in line with the Nordic evidence of a convex intergenerational earnings relationship. Bratsberg *et al.* (2007) argue that institutional factors may explain why mobility is higher at the bottom. In particular, they suggest that that educational and welfare systems in the Nordic countries help the upward mobility of young people with few parental resources. Interestingly, our findings for Canada are different from the patterns estimated in the United States, which exhibit an almost perfectly linear relationship between children's and parents' ranks in the income distribution (Chetty *et al*. 2014).[13]

In sum, the nonlinear patterns in Figure 5 suggest that a single IGE coefficient may be insufficient to offer an accurate picture of intergenerational earnings mobility in Canada. Corak and Heisz (1999) reach a similar conclusion using income transition matrices and then proceed to employ a nonparametric technique to explore the nature of these nonlinearities. In what follows, we use instead two different parametric approaches—i.e. higher-order polynomial and quantile regressions—to address the issue of nonlinearities. Our parametric results will be briefly compared to the findings of Corak and Heisz (1999) in the last section.

5.2   Nonlinear regressions: higher-order polynomial

As shown by Bratsberg *et al.* (2007), the intergenerational earnings model can be estimated using more flexible functional forms (i.e. higher-order polynomial) to fit the data. Here we estimate the coefficients from a fourth-order polynomial on the IID data and use them to calculate the IGE at each percentile of the fathers' earnings distribution. The results, presented in Figure 6, confirm that the degree of intergenerational earnings mobility in Canada is characterized by a marked nonlinear pattern. In fact, Figure 6 shows a logarithmic growth trajectory. In particular, the elasticity is quite low for sons at the bottom two percentiles of the fathers' earnings distribution (~ 0.1), which suggests a significant degree of upward mobility for sons born to very low-earnings fathers.

---

[13] In addition to linearity of the rank-rank relationship in the income distribution, Chetty *et al.* (2014) also show that the relationship between college attendance rates and parent income ranks is approximately linear. The latter finding would imply that all families are to some extent credit-constrained, as the chance of attending a college increases with parental incomes at a similar rate throughout the income distribution.

Further up the fathers' earnings distribution, the degree of mobility starts to decline. The estimated IGE sharply rises, reaching 0.32—the same value as the estimated coefficient from the linear specification—at around the 23rd percentile. Thereafter, earnings persistence increases monotonically with fathers' earnings. The estimated elasticity tops at 0.46 for the 98th and 99th percentiles of fathers' earnings, suggesting a strong intergenerational correlation of earnings among the very high-earnings families.

A salient feature of Figure 6 is that the estimates from the linear model (0.32) seem to understate the intergeneration transmission of earnings for the bulk of the population. These results support the conclusions in Bratsberg et al. (2007) that international comparisons of intergenerational mobility based on more flexible specifications may be more meaningful. Indeed, comparing the results in this study with the findings in Bratsberg et al., suggests that Canada has a pattern of intergenerational mobility that is quite similar to the one observed in the Scandinavian countries: a flat intergenerational relationship in the lower segments of the fathers' distribution and an increasingly positive correlation in middle and upper segments. Interestingly, this pattern is not observed in the U.S. or the U.K. (Bratsberg et al., 2007).

Figure 6: Forth-order polynomial estimates of β (IGE), evaluated at each percentile of fathers' earnings, father-son pairs
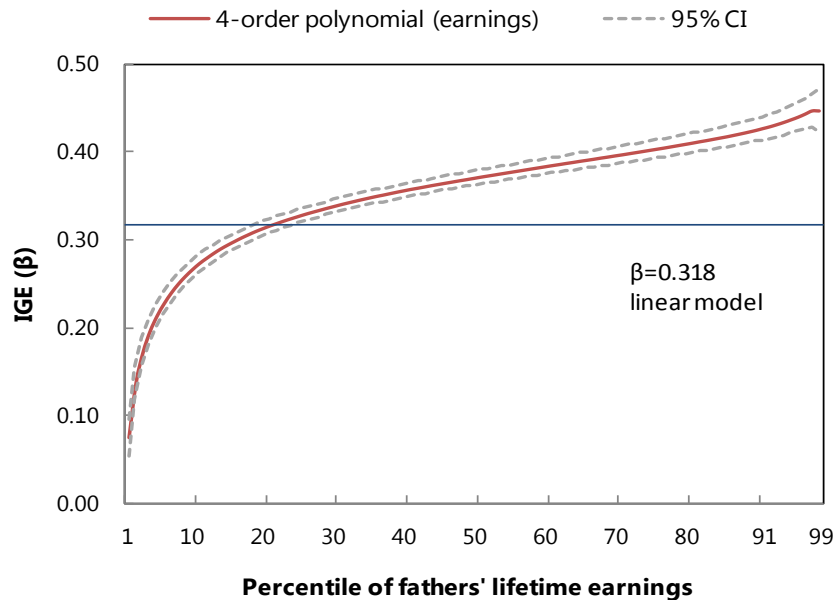
Note: Lifetime earnings are calculated by averaging fathers' (sons') earnings between the ages of 35 and 55 (38 and 42), conditional on having positive income, $500 and over, in at least 10 (3) years.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

5.3   Quantile regressions

While the results above present evidence of nonlinearity in the IGE across the *fathers'* earnings distribution, it is also interesting to examine whether mobility differs across the distribution of *sons'* earnings. As pointed out in Mulligan (1997) and Corak and Heisz (1999), the optimal amount of human capital investment made by fathers may also depend—positively—upon the child's ability. These authors

note that the parents who are more likely to be credit-constrained are the low-income parents with high-ability children. This raises the question of how the earnings of high-ability children from low-income backgrounds compare to those of their counterparts from high-income families. The answer to this question may be informative for policy discussions related to "equality of opportunity" in Canada. This notion refers to the existence of similar economic opportunities for children with similar abilities, regardless of their family of origin (Roemer, 1998).

Grawe (2004) suggests using quantile regressions to investigate this type of question. For instance, high elasticities in the top percentiles of the children's earnings distribution would suggest that high-earnings children come almost exclusively from families with high-earnings fathers. Unless one makes particular assumptions regarding the heritability of innate ability, this can be interpreted as evidence for inequality of opportunity. That is, it would suggest that high-ability children from low-earnings backgrounds have little chances to realize their full potential and become high-income earners.

Following the approach in Grawe (2004), Figure 7 presents the results from quantile regressions estimated on our father-son pairs. The estimates are produced for each percentile (1st, 2nd,…, 99th). The estimated 95% confidence intervals are also reported. The quantile regression results depict a clear pattern of nonlinearity over the distribution of sons' earnings. The estimated earnings elasticity is relatively low (~ 0.2) at the bottom two percentiles of the sons' earnings distribution, rises to 0.35 at the 10th percentile and stays at around 0.33-0.34 throughout the lower half of the distribution. It then drops gradually reaching 0.27 for the 80-85th percentiles, and reverses its course by rising again to 0.41 in the 99th percentile.

Three distinct patterns of intergenerational earnings transmission emerge from the quantile regression analysis. First, quantile regression coefficients are quite low for sons in the bottom 2 percentiles of the distribution. That is, a large fraction of sons with the lowest earnings do not have fathers who were themselves at the very bottom of the distribution in their generation. This implies that the sons of the lowest-earnings fathers are relatively mobile, which is consistent with our results from the polynomial model above. Second, between the 10th and 85th percentiles of the sons' earnings distribution, persistence is fairly high in the beginning (the estimated elasticity is above 0.35 at the 10th percentile), declining slowly over the lower half, and dropping more notably over the upper-middle part of the distribution. A rather high elasticity over the lower-middle part of the sons' distribution indicates that a considerable fraction of moderate-earnings sons have fathers with similar economic outcomes. Such correlation, however, becomes less obvious as sons' earnings increase. Earnings mobility starts to rise for sons in the upper-middle part of the distribution. The estimated elasticity declines from 0.33 at the median to 0.27 at the 85th percentile. This suggests that sons from a wide range of earnings backgrounds have a good chance of becoming adults with above-average earnings. The third pattern from Figure 7 is the reversal of the quantile regression coefficients over the top (85th to 99th) percentiles, which reveals a stronger degree of earnings persistence for high-earnings sons. The estimated coefficients are 0.38 and 0.41 for the top two percentiles. That is, a significant fraction of those who make it to the highest earnings groups have a high-earnings father.

Finally, Figure 7 also suggests that Canadian mobility may be characterized by rather complex transmission mechanisms. In particular, distinct channels of parental influence may be at work at different parts of the distribution. At the bottom of the distribution, institutional factors—as suggested in the

Nordic studies—may play a role in facilitating upward mobility for children from very low-earnings backgrounds.[14] The estimated profile over the bulk of the distribution would be broadly in line with the credit constraints hypothesis: the relationship between moderate-earnings sons and moderate-earnings fathers is expected to be stronger because these fathers are more likely to be credit-constrained and may not be able to invest optimally in the human capital of their children. The profile corresponding to the upper-middle part of the distribution may also be consistent with less binding credit constraints for fathers whose earnings are increasingly sufficient to invest in their children's human capital. Finally, channels of transmission may be even more complex at the top, as suggested in the literature. Factors other than human capital investment, such as networking or family-specific capital, may be more salient for the intergenerational earnings transmission among top-earnings families (Bjorklund *et al*, 2012; Corak and Piraino, 2011; Kramarz and Skans, 2014).

A few caveats, however, need to be borne in mind in interpreting the dynamics at the bottom end of the distribution. It is possible that at this end of the distribution the mother is carrying the load in terms of market activities. Also note that our analytical sample excludes all those sons raised in fatherless (i.e. single mother) households who are more likely to be in low income situations. Moreover, data quality may also be a concern as measurement error often tends to be more pronounced at the bottom end of earnings or income distribution. Further discussion and identification on channels of influence is required but is beyond the scope of this paper.

Figure 7: Quantile regression estimates of β (IGE), earnings of fathers and sons
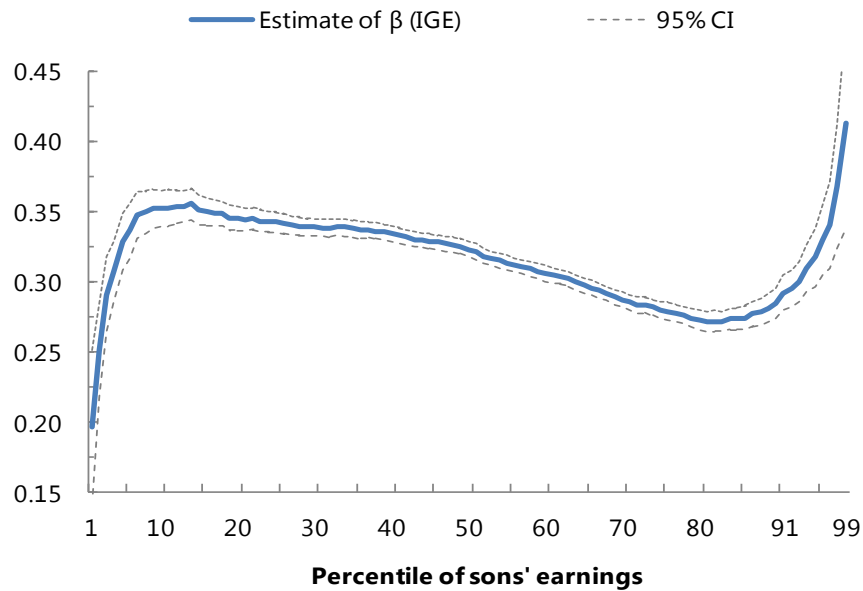
Note: Lifetime earnings are calculated by averaging fathers' (sons') earnings between the ages of 35 and 55 (38 and 42), conditional on having positive income, $500 and over, in at least 10 (3) years.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

---

[14] Higher mobility at the very bottom would also be consistent with human capital models if fathers of low-ability children are reluctant to invest in the human capital of their children (Becker and Tomes, 1986).

5.4   Market income and total income

Figure 8 shows nonlinear patterns in intergenerational income persistence for market (dashed line) and total (dotted line) income. Both polynomial (Panel A) and quantile regression (Panel B) results are presented. In general, despite differences in the magnitude, the pattern of nonlinearity across the distribution seems quite similar for all three income measures. As in the linear model, persistence tends to be higher for market or total income than for earnings. This is especially the case at the very top: sons of fathers with the highest income are much more likely to have the highest incomes as adults.

Differences in generational persistence for different income sources, however, are more visible in the quantile regression results (Panel B). For instance, the correlation between low-income sons (e.g., those in the 10th percentile) and low-income fathers is stronger than between low-earnings fathers and sons. Moreover, higher mobility over the upper-middle part of the distribution is somewhat limited for market and especially total income. Unlike the results for earnings, where some of the higher-earnings sons (e.g., those at the 85th percentile) can come from a rather diverse earnings background, the higher-income sons are more likely to have fathers with a similar position in the income distribution. Income persistence is especially pronounced at the top end of the distribution, at 0.51 or higher for the top 2 percentiles of the sons' total income. These results are broadly consistent with a growing literature showing that affluent families are able to improve the income potential of their children in various ways—including better child care, private schooling, transfer of entrepreneurial skills and social connections (e.g. Duncan and Murnane, 2011).

Figure 8: Nonlinear estimates of β (IGE) for fathers and sons, by income source
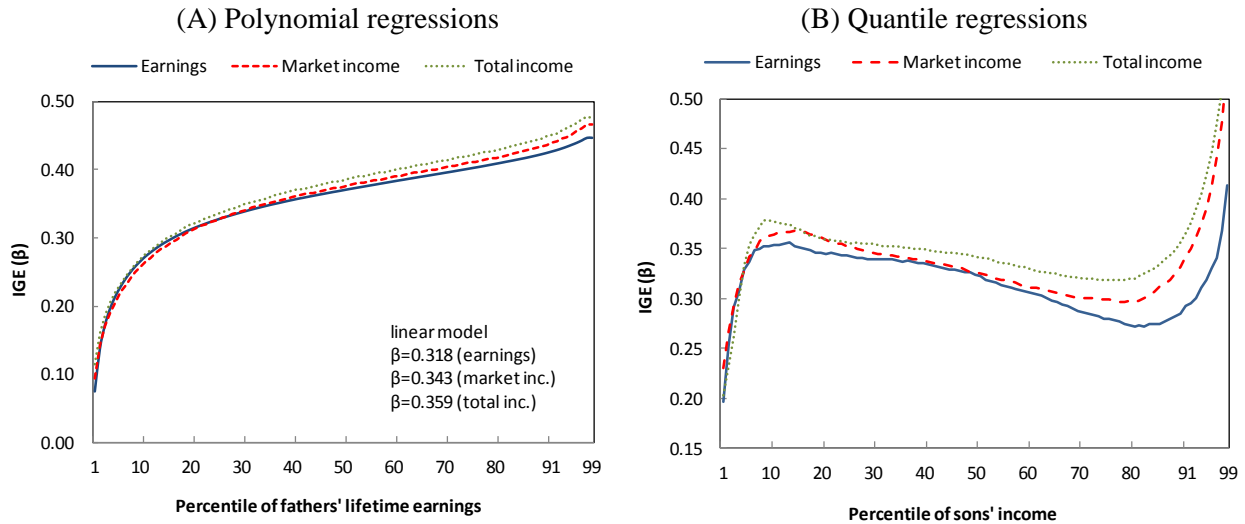
(A) Polynomial regressions   (B) Quantile regressions

Note: Lifetime earnings are calculated by averaging fathers' (sons') earnings between the ages of 35 and 55 (38 and 42), conditional on having positive income, $500 and over, in at least 10 (3) years.

Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

## 5.5   *Lifecycle, errors-in-variables, and nonlinearity*

As it was the case in the simple linear model, the patterns of nonlinearity may not be estimated correctly if the lifecycle and errors-in-variables biases are not properly addressed. The distortion is likely to be greater in the upper part of the distribution, since most children with high lifetime earnings have not yet reached their full earnings potential at younger ages. To confirm this intuition, the analysis in the previous section is repeated with two alternative definitions for fathers' and sons' lifetime earnings. The first scenario defines fathers' lifetime earnings as five-year averages over the period from 1978 to 1982, while sons' earnings are drawn from the year corresponding to age 30 (dashed line). The second scenario maintains the same earnings definition for fathers but measures sons' earnings are at age 40 (dotted line). Based on the results in the previous section, both lifecycle and errors-in-variables biases are likely to be present in the first scenario, while the former bias can be minimized in the second scenario. The results are presented in Figure 9, which also provides the preferred estimates for reference (solid line).

Figure 9 indeed reveals a very different pattern of nonlinearity when sons' earnings at age 30 are used. For polynomial regressions, this leads to an underestimation of intergenerational persistence throughout virtually the entire distribution of fathers' earnings. Because of lifecycle variation, the extent of the correlation between father's and son's earnings appears to be particularly underestimated in the upper part of the distribution. For instance, the estimated elasticity at the 95th percentile is only 0.24— about 45% lower than the estimate (0.44) from the preferred definition. When lifecycle bias is minimized by using sons' earnings at age 40, the pattern of nonlinearity is almost identical to the one from the preferred model for much of the distribution. However, the estimates for the top percentiles are still lower than in the preferred model due to the measure of fathers' earnings being less accurate.

The quantile regression results also show that not accounting for lifecycle variation may result in significantly biased patterns of nonlinearity. The decline in earnings persistence between the 15th and 95th percentiles of sons' earnings is much steeper compared to the preferred model, giving a wrong impression that mobility increases until the 95th percentile of the son's distribution. The estimated elasticity at 95th percentile is as low as 0.14 when sons' earnings at age 30 are used as a proxy for lifetime earnings, 56% lower than the preferred estimate of 0.32. Again, the pattern of nonlinearity becomes more similar to the one from the preferred model when sons' earnings at age 40 are used. It is also interesting to note that using five-year earnings averages from a given time period for fathers also leads to a downward bias in the quantile regression estimates for large segments in the middle of the sons' earnings distribution.

Figure 9: Lifecycle and errors-in-variables induced biases and nonlinearities
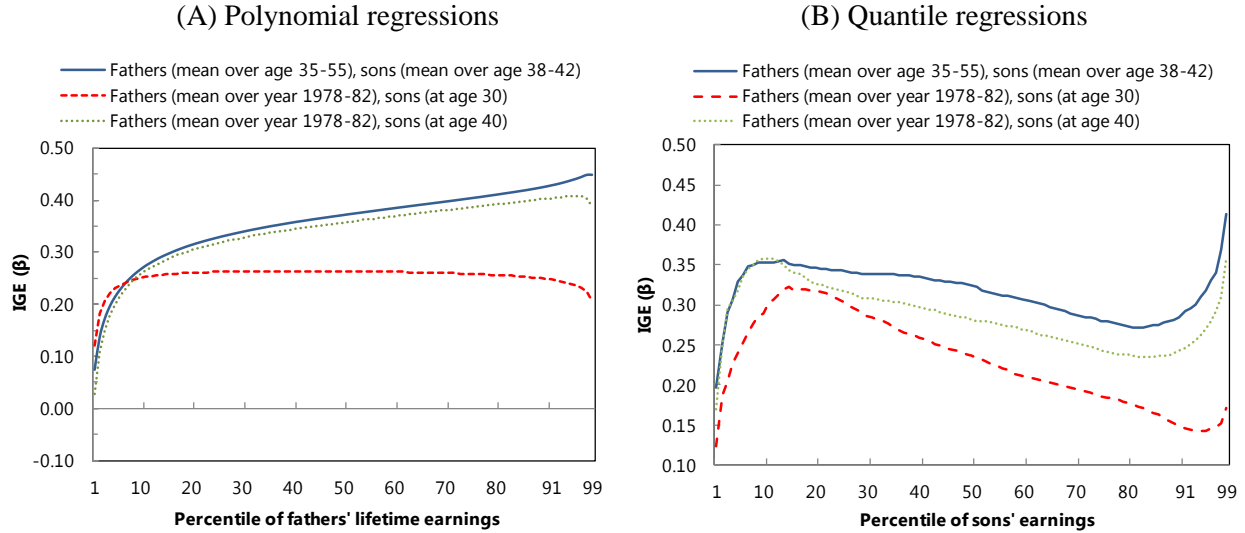
(A) Polynomial regressions

(B) Quantile regressions

Note: Lifetime earnings are calculated by averaging fathers' (sons') earnings between the ages of 35 and 55 (38 and 42), conditional on having positive income, $500 and over, in at least 10 (3) years.

Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

Finally, note that Corak and Heisz (1999) also examined the nature of nonlinearities in intergenerational earnings persistence in Canada. They estimate a flexible nonparametric model on the older version of the IID data. While a direct comparison with our study is not possible, due to the different functional forms employed, it is interesting to note that Corak and Heisz (1999) also show a considerable rise in the estimated elasticity at the very top of the distribution—above the 99th percentile. Unlike our results, however, they find less persistence just below the 99th percentile compared to the IGE estimated for most of the upper-part of the distribution. It is difficult to say whether this difference is due to the different functional forms employed in the two studies or to the fact that our improved dataset can better account for lifecycle variation in earnings. However, the results depicted in Figure 9 suggest that inaccurate proxies for lifetime earnings can significantly distort the estimated patterns of nonlinearity in the intergenerational transmission of earnings.

6. Conclusions

Understanding the extent of intergenerational earnings and income mobility is informative for economic and social policies. In particular, estimates of the intergenerational income elasticity (IGE) are often seen as broad indicators of equality of opportunity. However, because full career histories of parents and children are generally not available in the data, many existing IGE estimates for various countries may be affected by biases arising from inadequate proxies for lifetime earnings. This paper re-examines the extent to which lifecycle variation and errors-in-variables can bias IGE estimates both at the mean as well as across the percentiles of the income distribution. The new augmented Intergenerational Income

Data (IID) from Canada, with nearly full history of career data, enables us to advance our understanding on this subject.

Our analysis shows that lifecycle bias may be present at any stage of the working career. The bias tends to be higher when annual earnings from early career years are used to proxy for lifetime earnings as the dependent variable, which is in line with the existing international evidence. However, we also find cross-country differences with respect to the age at which lifecycle bias may be minimal. This finding highlights the difficulty of making appropriate international comparisons even when sample and methodological differences are accounted for. That is, cross-national variation in IGE estimates may still arise due to differences in the age at which sons' earnings are observed.

The intergenerational earnings elasticity for Canada is estimated at around 0.32. This is higher than the 0.22 estimate obtained in previous Canadian literature. Accounting for lifecycle bias in the early estimates explains about two-thirds of the discrepancy, while errors-in-variables induced bias contributes to the remaining difference. We show that the extent of the bias is larger for market and total income as compared to earnings alone. The results also reveal significant gender differences with regards to the effect of theses biases on the estimated IGEs. The father-daughter elasticity remains quite modest irrespective of the ages at which daughters' earnings are measured. It is possible that the lower IGE for daughters may be driven by gender differences in labour force participation and/or by estimation issues related to marital sorting. This highlights the need for a closer look into these patterns for future research.

Using data that are less affected by measurement error compared to previous studies, the paper shows a distinct pattern of nonlinearity in the intergenerational transmission of income in Canada. The results show that the relationship between fathers' and sons' earnings exhibits a rather convex pattern, similar to the one found in the Nordic European studies and in contrast to the linear pattern observed in the U.S. literature. Both polynomial and quantile regressions reveal high mobility rates at the very bottom of the distribution and low mobility at the top. We conjecture that social institutions may help explain these findings. Moreover, we demonstrate that the patterns of nonlinearity can be significantly misread when the lifecycle and errors-in-variables biases are not adequately addressed, especially over the upper part of the distribution.

References

Aina, Carmen and Cheti Nicoletti (2013). "The intergenerational transmission of liberal professions: nepotism versus abilities," Unpublished memo.

Becker, G.S., and G. Solon (2003). "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records," *Journal of Labor Economics* 21(2):289-321.

Becker, G.S., and N. Tomes (1986). "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics* 4(3):S1-S39.

Bjorklund, Anders (1993). "A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89," *Review of Income and Wealth,* 39(4):377-86.

Bjorklund, A., & Jantti, M. (2010). Intergenerational income mobility and the role of family background. In W. Salverda, B. Nolan, & T. Smeeding (Eds.), Handbook of economic inequality. Oxford: Oxford University Press.

Björklund, A., J. Roine, and D. Waldenström (2012). "Intergenerational top income mobility in Sweden: Capitalistic dynasties in the land of equal opportunity?" *Journal of Public Economics*, 96, 474-484.

Böhlmark, Anders and Matthew J. Lindquist (2006). "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden," *Journal of Labor Economics,* 24(4):879-896.

Bratsberg, B., K. Røed, O. Raaum, R. Naylor, M. Jäntti, T. Eriksson and E. Österbacka (2007). "Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons," *The Economic Journal*, 117:C72-92.

Brenner, Jan (2010). "Life-Cycle Variations in the Association between Current and Lifetime Earnings: Evidence for German Natives and Guest Workers," *Labour Economics,* 17:392-406.

Chadwick, L. and G. Solon (2002). "Intergenerational Income Mobility among Daughters," *The American Economic Review*, 92(1): 335-344.

Chetty, Raj, N. Hendren, P Kline and E. Saez (2014). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 1-71.

Corak, Miles (2001). "Are the kids all right? Intergenerational mobility and child well-being in Canada," *The review of Economic Performance and Social Progress*, 273-292.

Corak, Miles (2006). "Do poor children become poor adults? Lessons from a cross country comparison of generational earnings mobility," in: John Creedy and Guyonne Kalb (Ed.), *Research on Economic Inequality*, Vol. 13, 143-188.

Corak, Miles (2013). "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," Discussion Paper No. 7520. IZA.

Corak, Miles and Andrew Heisz (1999). "The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data," *Journal of Human Resources*, 34(3): 504-33.

Corak, Miles, Björn Gustafsson and Torun Österberg (2004). "Intergenerational Influences on the Receipt of Unemployment Insurance in Canada and Sweden," in Corak, Miles (Ed.), Generational income mobility in North America and Europe. Cambridge University Press, Cambridge.

Corak, Miles and Patrizio Piraino (2011). "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, 29(1): 37-68.

Corak, Miles, Matthew Lindquist and Bhashkar Mazumder (2014). "A Comparison of Upward and Downward Intergenerational Mobility in Canada, Sweden and the United States," *Labour Economics*, 30: 185-200.

Duncan, Greg J. and Richard J. Murnane (2011). "Introduction: The American Dream, Then and Now." In Greg J. Duncan and Richard J. Murnane (editors). Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances. New York: Russell Sage Foundation.

Dunn, Thomas, and Douglas Holtz-Eakin (2000). "Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links," *Journal of Labor Economics* 18:282–305.

Ermisch, John, Marco Francesconi and Thomas Siedler (2005). "Intergenerational Economic Mobility and Assortative Mating," IZA Discussion Paper No. 1847.

Fortin, N., and S. Lefebvre (1998). "Intergenerational Income Mobility in Canada," in *Labour Markets, Social Institutions, and the Future of Canada's Children,* ed. M. Corak. Catalogue #89-553. Ottawa: Statistics Canada.

Gouskova, E., Chiteji, N., Stafford, F., 2010. Estimating the intergenerational persistence of lifetime earnings with life course matching: Evidence from the PSID. Labour Economics 17 (3), 592–597.

Grawe, Nathan D. (2004). "Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints," *Journal of Human Resources*, vol. 39(3), pp. 813–27.

Grawe, Nathan D. (2006). "Lifecycle bias in estimates of intergenerational earnings persistence," *Labour Economics*, Vol. 13, 551-570.

Haider, Steven and Gary Solon (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *The American Economic Review*, 96(4), 1308-1320.

Han, Song and Casey B. Mulligan (2001). "Human Capital, Heterogeneity and Estimated Degree of Intergenerational Mobility," *The Economic Journal*, 111, 207-243.

IID Guide (2010). "Intergenerational Income Data Reference Manual," Unpublished Manuscript, Statistics Canada.

Kramarz, F. and O. N. Skans (2014). "When Strong Ties are Strong: Networks and Youth Labour Market Entry," *The Review of Economic Studies,* 81:1164–1200

Lentz, B. F., and D. N. Laband (1990). "Entrepreneurial success and occupational inheritance among proprietors," *Canadian Journal of Economics,* 23:563–79.

Mazumder, Bhashkar (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data," *The Review of Economics and Statistics*, 87(2), 235-255.

Mulligan, Casey B. (1997). "Parental Priorities and Economic Inequality," University of Chicago Press, Chicago.

Nybom, Martin and Jan Stuhler (2011). "Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation," IZA Discussion Paper No. 5697.

Oreopoulos, Philip. 2003. The long-run consequences of growing up in a poor neighborhood. Quarterly Journal of Economics 118:1533–75.

Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens. 2008. The intergenerational effects of worker displacement. Journal of Labor Economics 26:455–84.

Page, Marianne E. (2004). "New Evidence on Intergenerational Correlations in Welfare Participation," in Corak, Miles (Ed.), Generational income mobility in North America and Europe. Cambridge University Press, Cambridge.

Perez-Gonzalez, Francisco (2006). "Inherited control and firm performance," *American Economic Review* 96:1559–88.

Raaum, O., B. Bratsberg, K. Røed, E. Österbacka, T. Eriksson, M. Jäntti, and R.A. Naylor (2007). "Marital sorting, household labor supply and intergenerational earnings mobility across countries," *The B.E. Journal of Economic Analysis and Policy*, Article 7.

Roemer, J.E. (1998). *Equality of opportunity*, Harvard University Press, Cambridge.

Solon, Gary (1992). "Intergenerational Income Mobility in the United States," *American Economic Review*, 82(3):393-408.

Solon, Gary (1999). "Intergenerational Mobility in the Labor Market," in Orley C. Ashenfelter and David Card, eds., *Handbook of labor economics*, Vol. 3A. Amsterdam Elsevier Science, North-Holland, 1761-1800.

Solon, G. (2002). Cross-country differences in intergenerational earnings mobility. Journal of Economic Perspectives, 16(3), 59–66.

Sorensen, Jesper B. (2004). "Closure and Exposure: Mechanisms in the Intergenerational Transmission of Self-employment," Unpublished memo.

Table A1. Age distribution of fathers in 1978, IID

Age in 1978	N	%	Mean earnings over 1978-82
34 or less	27,001	5.34	46557.4
35-40	137,263	27.13	53508.7
41-45	143,230	28.31	55282.1
46-50	108,582	21.46	54201.6
51 and over	89,910	17.77	49252.5
Total	505,986	100	53155.1

Note: The sample includes all fathers (of sons and daughters) associated with the IID cohort 1982.
Source: Authors' calculations from Statistics Canada's Intergenerational Income Data (IID).

Table A2. Illustration of sample selection criteria for fathers

Linked earnings years for fathers: 1978-1999

ID	Father cohort	year at 35	year at 55	Age (35–55)	# records bt 35-55	# years $500+	Record used
A	1925	1960	1980		3	3	No
B	1935	1970	1990		13	13	Yes
C	1940	1975	1995		18	16	Yes
D	1943	1978	1998		21	21	Yes
E	1944	1979	1999		21	15	Yes
F	1950	1985	2005		15	15	Yes
G	1950	1985	2005		15	7	No
H	1955	1990	2010		10	10	Yes
I	1960	1995	2015		5	5	No

■ years for which earnings records are available   ■ earnings <$500 (2010 constant)