Intergenerational Mobility between and within Canada and the United States

Marie Connolly, Miles Corak, Catherine Haeck*

Université du Québec à Montréal, University of Ottawa, Université du Québec à Montréal

January 2017, second draft not for distribution or citation

Abstract

Intergenerational mobility is about twice as great in Canada than in the United States, but varies significantly within each country. Our sub-national analysis of six different indicators finds that the national border only partially distinguishes the close to one thousand regions we analyze within these two countries. The Canada-US border clearly divides Central and Eastern Canada from the Great Lakes regions and the Northeast of the United States. But these differences drive only part of the national differences in mobility. While some Canadian regions have more in common with the low mobility southern parts of the United States than with the rest of Canada, the fact that they represent a much smaller fraction of population is the other reason why overall mobility is lower in the United States.

KEYWORDS: Intergenerational mobility, equality of opportunity, geography

JEL Classification: *

*We are respectively affiliated with the department of economics at the Université du Québec à Montréal, the Graduate School of Public and International Affairs at the University of Ottawa, and the department of economics at the Université du Québec à Montréal. The first draft of this paper was written for the NBER and Employment and Social Development Canada conference on “Public Policies in Canada and the United States,” held in Ottawa Canada on October 27 and 28, 2016. We are grateful to Yuri Ostrovsky and Winnie Chen for their support in constructing the Canadian data upon which our analysis is based. This was possible with the cooperation of the Social Analysis and Modelling Division of Statistics Canada through a cost recovery project, and with financial support from the fonds de recherche du Québec - Société et culture (grant 2016-PU-195586 to Connolly and Haeck) and from the Social Sciences and Humanities Research Council of Canada (Insight grant 230936-191999-2001 to Corak). We also gratefully acknowledge research assistance from Charles-A. P.-Demontigny and the support of Martin Leblond-Létourneau, and thank David Card, Emmanuel Saez, and Jann Speiss for their comments and suggestions. The analysis was in part conducted at the Quebec Interuniversity Centre for Social Statistics, which is part of the Canadian Research Data Centre Network. These services are made possible by the financial or in-kind support of the Social Sciences and Humanities Research Council, the Canadian Institutes of Health Research, the Canada Foundation for Innovation, Statistics Canada, the Fonds de recherche du Québec - Société et culture, and the Fonds de recherche du Québec - Santé. The views expressed in this paper are those of the authors, and not necessarily those of the Canadian Research Data Centre Network or its partners. A preliminary draft was also presented in October 2016 as an invited paper to the Sixth Annual NYU/UCLA Tax Policy Symposium on Tax Policy and Upward Mobility held at UCLA, in November 2016 as an invited paper to the Centro de Estudios Espinosa Yglesias and Fundación Espinosa Rugarcía sponsored Social Mobility Summit held in Mexico City, and to departmental seminars at the Graduate Center, City University of New York, and the University of California, Berkeley. Comments and suggestions are welcomed at connolly.marie@uqam.ca, @MilesCorak, and haeck.catherine@uqam.ca.
1 Introduction

to be determined and written.

2 Lessons from theory for cross-country comparisons

It is natural to compare intergenerational mobility over time within a country, but theory suggests this may be less informative than generally thought. And while this caution also applies to cross-country comparisons, there is a case to be made that a Canada-US comparison may be particularly apt, or certainly more-so than comparisons between either of these countries to any other rich country. In particular, there are “small differences” in the way Canadians and Americans value and understand intergenerational mobility, and this may help to bring into relief important differences in the nature of labour markets and public policies.

Time series variation in the intergenerational earnings elasticity is often viewed as policy relevant because it is the first step in uncovering a causal relationship. This is certainly implicit, if not explicit, in discussions that relate income inequality to intergenerational income mobility. Alan Krueger (2012), in his role as Chairman of the Council of Economic Advisers, cited the positive cross-country correlation between income inequality and the intergenerational earnings elasticity—a relationship he christened as the “Great Gatsby Curve”—to motivate the possibility that rising income inequality in the United States will in the coming decades move the country in the direction of less social mobility. Chetty, Hendren, Kline, Saez, et al. (2014) examine temporal variation of intergenerational mobility. Their analysis of cohorts born since the 1980s may lead some readers to conclude that in spite of rising inequality there has not been much, if any, change in intergenerational mobility. Lee and Solon (2009) use data for many earlier cohorts, and also find no statistically significant trend in mobility.\(^1\)

The obvious caution is that there are very long lags in intergenerational processes that, after all, are determined by many different factors that may be changing in possibly offsetting ways. Solon (2004; 2015) adapts the workhorse model due to Becker and Tomes (1986; 1979) and Loury (1981) to illustrate the challenges of making comparisons across time and space. A simplified version of the model in Solon (2004)—one that puts aside an explicit utility function reflecting the influence of parental altruism on child investments, and also the influence of the progressivity of government investment—is captured in the following equations:

\[ y_t = \rho h_t \]  

---

\(^1\)Clark (2013) uses rare surnames of elite families to argue that there is essentially no, and there never has been, variation of intergenerational mobility across time, or for that matter between countries. His analysis does not refer directly to intergenerational income mobility, but rather to an underlying latent indicator that he implicitly relates to genetic endowments. Aaronson and Mazumder (2008) use US Census data, and find a rise in the intergenerational earnings elasticity that they argue coincides with a period of rising inequality, as indicated by increases in the return to education.
\[ h_t = \theta y_{t-1} + e_t \]  
\[ e_t = \lambda e_{t-1} + \epsilon_t \]  

where \( t \) indexes generations within a family dynasty, \( y \) refers to the logarithm of permanent income, \( h \) to human capital, which determines income (having a return \( \rho \)), and in turn may be determined by parental income if there are credit constraints in human capital investment, but otherwise determined by \( e \), an unobserved endowment not influenced by family investment decisions, and mechanically transmitted across generations according to a first order auto-regression governed by parameter \( \lambda \), the “inheritability of endowments,” with \( \epsilon \) being a random variable representing “luck.”\(^2\) Solon (2004) shows that in a steady state the population regression of earnings across generations has an elasticity \( \beta = (\rho \theta + \lambda) / (1 + \rho \theta \lambda) \).

With perfect capital markets, that is \( \theta = 0 \), the structure of labour markets and the institutions that determine human capital investment do not come into play, and differences in earnings mobility over time or across space reflect—in some loosely defined way—differences in populations and environments that imply differences in the nature and transmission of endowments. This is hardly expected to vary much over time within a country, and offers in the manner of Clark (2013) an explanation for the lack of variation in the intergenerational mobility even with rising inequality, or the passing of many generations.

But if capital markets are not perfect, then differences in the returns to human capital, and the causal role of parental income determining the amount of human capital, must also be part of the story. Inequality may be on the rise, as reflected in higher values of \( \rho \), but individual behaviour or public investments easing \( \theta \) may be countervailing forces. For example, in the United States income inequality and the returns to schooling have certainly risen since the 1980s, but at the same time high school drop out rates and teen-age fertility have fallen, while government support directed to the schooling of children from lower income families has increased. In other words, even in this simple model, variations, or for that matter, the lack of variation in \( \beta \) over time, need not be informative, offering no specific basis for uncovering the causal impact of labour market inequality.

Besides, all of this discussion refers to comparative statics, and Nybom and Stuhler (2014) show that the movement to a new steady state may span many generations, and proceed through non-monotonic paths. Equality of opportunity enhancing policies may immediately promote intergenerational mobility, but then in subsequent generations cause it to fall before the new steady-state is reached. With this in mind, it is quite reasonable to suppose that a

---

\(^2\)Clearly, this is a very simplified framework, but it can be easily extended. A child’s adult income may be influenced not just by his or her human capital and endowment, but possibly directly by parental income, which might proxy for nepotism or the role of social networks, with the children of higher income parents earning more than equally skilled children of lower income parents. See Mulligan (1997), and the application and interpretation of this formulation by Corak and Piraino (2016), who add parental income to Equation 1. This simplified model also probably encourages us to think of \( e \), and its evolution between generations, as a genetic endowment. But the literature on epigenetics cautions against making a sharp distinction between “nature and nurture,” and it might be more constructive to view \( e \) as embodying a host of inherent and environmental factors. Becker and Tomes (1986) think in these terms. Heckman and Mosso (2014) recognize this by generalizing Equation 2 to reflect the recursive process that more accurately models child development, the multi-faceted nature of skills, and the role and complimentarities of human capital investments within and across the successive stages through which development occurs.
policy like the GI Bill, introduced in the aftermath of World War II to encourage the college education of returning veterans, may still be having an effect on current measures of the American intergenerational earnings elasticity. In subsequent generations the policy would be causing the elasticity to rise, as the children of the program’s beneficiaries—who presumably are more educated and highly paid than the average—reap the benefits of having higher status parents.

All of this should make one wonder about the informative content in time series variation of broad summary measures of intergenerational mobility like the intergenerational elasticity, and encourage more detailed analysis of the underlying drivers whose comparative static properties are clear, and which can be assessed over shorter horizons.

But this does not necessarily imply that cross-country comparisons are any more informative. Becker et al. (2015), and specifically Solon (2004), add other dimensions relevant to cross-country comparisons, permitting parental altruism (that is, tastes and values), the progressivity of government spending, and substitutability between private and public investments, to also influence the intergenerational elasticity. Countries may occupy different positions on the Great Gatsby Curve for the simple reason that they make different choices reflecting different social objectives, or different perspectives on the role of private versus public investments in children. This may be one reason time-series variation is immediately looked to—changes within a country are more likely to reveal the impact of policy than they are to reflect changes in social priorities—and underscores the fact that not all cross-country comparisons are relevant. The United States is often compared to Denmark in popular discussion, but the relevance of this comparison is not clear given the contrast between the ethnic homogeneity of the Danish population, and the much larger and more diverse American, to say nothing of possibly very different perspectives on equality of outcomes versus equality of opportunity as social goals.3

Simply put, values and institutions differ, and this has longed played a role in understanding differences in the nature and size of the welfare state, particularly between continental Europe and North America (Alesina and Glaeser 2004). In a similar way, the priority societies place on equality of opportunity enhancing policies will reflect underlying values. Arguably, these are more similar across the Canadian - American border than elsewhere. In fact, public opinion polls find that Americans and Canadians define the “American Dream” in virtually the same way. The Pew Charitable Trusts conducted a number of public opinion polls asking Americans what meaning they attach to the phrase “The American Dream.” The responses have been stable through time, and a poll conducted in late January and early February 2009 was adapted and conducted in Canada in August and September of the same year (Corak 2010). Figure 1 summarizes one of the major findings by indicating the fraction of respondents in each country answering eight or higher on a ten point scale for each of the

3DeParle (2012) is an example of a balanced cross-country comparison, but which nonetheless can’t resist bringing the Danes into the picture. The frequent use of a Denmark-US comparison in public discussion probably has its origin in the emphasis placed on this pairing in a popular book (Wilkinson and Pickett 2009), and the fact that the countries sit at the two extreme poles of the Great Gatsby Curve, Denmark having the lowest Gini coefficient and lowest intergenerational elasticity, and the United States the highest among the countries depicted in Corak (2013). That said, see Landersø and Heckman (2016) for a careful analysis of intergenerational dynamics in these two countries.
possible definitions of the “American Dream” presented to them.\footnote{The poll was conducted under the sponsorship of the PEW Charitable Trusts by Greenberg Quinlan Rosner Research and Public Opinion Strategies using a sample of 2,119 American adults 18 years and older, and by EKOS Research Associates using sample of 1,035 Canadians falling in the same age group. The specific question referring to the American Dream asked to Canadians was directly adapted from the US poll and read: “Americans often talk about attaining the American Dream to describe what it means to have a good life in their country. This means different things to different people. Here are some ways some Americans have described what the American Dream means to them. On a scale of one to ten, please tell me how accurately each statement describes what you would consider the Canadian Dream to be. One would mean the statement does not describe it perfectly.”}

The point estimates are very similar, particularly so with respect to the options most closely tied to intergenerational mobility. Sixty percent of American respondents ranked being able to succeed regardless of family background eight or higher on a ten point scale, while 59 percent of Canadians did so.\footnote{The exact wording of the option presented to respondents was: “Being able to succeed regardless of the economic circumstances in which you were born.”} The percentage indicating that the statement “Your children being better off financially than you” represents the American Dream was 64 percent in the United States, and 57 percent in Canada.\footnote{While these two options relate most directly to intergenerational mobility, they do not offer clear guidance on the appropriate way to measure it statistically. The reference to children being financially better off most clearly refers to absolute mobility using an intra-family reference point, but “being able to succeed regardless of family background” leaves the reference point undefined and may be interpreted in a relative or an absolute sense.} The differences between these responses, and those to all but one of the other options listed in the figure, are not statistically significant.\footnote{The only exception is that forty percent of Americans suggest “Owning your own business” is a strong marker of the meaning of the American Dream, but at 29 percent significantly fewer Canadians.}

The findings of the PEW poll also suggest that Americans and Canadians share a preference for equality of opportunities over equality of outcomes: 71 percent of Americans and 68 percent of Canadians felt it was more important “to ensure everyone has a fair chance of improving their economic standing” than “to reduce inequality.” In addition, both Americans and Canadians see the prime determinants of social mobility as relating more to individual decisions and effort, rather than to circumstances beyond individual control, “hard work” and “having ambition” at the top of this list (Corak 2010, 15–16).

But if American and Canadians have a similar meaning of the good life, and share similar views on how to attain it, they have significantly different views on the role of collective action through public policy in helping or hindering their pursuits.

A notable difference between the two countries concerns the role of government as a means to influence economic mobility. When asked if the government does more to help or more to hurt people trying to move up the economic ladder, respondents in both countries lacked strong proclivities. However, 46 percent of Canadians feel that government does more to help than to hurt, compared to 36 percent of Americans. On the other hand, 46 percent of Americans feel government does more to hurt versus 39 percent of Canadians. The difference
Figure 1: The ‘American Dream’ means the same thing to Canadians as to Americans: Percentage of respondents in comparable public opinion polls indicating eight or higher on a ten point scale to alternative descriptions of the American Dream.
in the responses to this question was among the largest of all questions asked (Corak 2010, 17).

This is as much a statement about differences in beliefs about the efficacy and efficiency of public policy as it is about ideological differences on the role of the state, and has a clear echo in the public opinion research conducted by Alesina, Stantcheva, and Teso (2017). They focus on the relationship between perceptions of intergenerational mobility and preferences for redistribution in the United States and four European countries. Canada is not part of their analysis, and their purpose is not the same as ours, focusing on intergenerational mobility to understand the strength of preferences for equality of outcomes. But they underscore the point that beliefs about the role of government intervention are at the core of political polarization, and lead to different views on what to do about intergenerational mobility even if the lack of it is perceived as a problem. For example, in the PEW poll a slight majority of Americans (51 percent) feels that cutting taxes would be a “very effective” government action to improve mobility, but only 38 percent of Canadians feel the same way (Corak 2010, Figure 7). These differences imply different public policy capacities. For example, Hoynes and Stabile (2016) document the very dramatic differences in income support for lower income families in the two countries, the safety net being drawn much more tightly in Canada. Even if underlying values are the same in these two countries, perceptions on the role of public policy—and ultimately the range and design of policy—may be very different.

This relates directly to the emphasis Solon (2004) puts on the “progressivity” of public investment in human capital as a determinant of intergenerational mobility. His model also suggests that cross-country differences in the design of public policies—in addition to differences in the inheritability of endowments, the returns to human capital investment, and the capacity of family income to influence a child’s human capital—play a role in determining differences in intergenerational mobility. All of these factors may come into play in understanding Canada-US differences. But because these two countries define and value mobility in the same way, the contrast between them may help to place a sharper focus on differing beliefs about the role of public policy. A Canada-US comparison might open up a wider menu of choices in American public debate than if the comparison was just over time within the country, or for that matter to European countries whose policy agenda can be more easily dismissed as not relevant to American values.

## 3 Data and measurement

The economics literature on intergenerational income mobility has grown significantly since the early 1990s when the maturing of the Panel Study of Income Dynamics offered Solon (1992) the opportunity to estimate intergenerational income elasticities with nationally representative data spanning two generations.\(^8\) The detailed surveys of this research by Björklund and Jäntti (2011), Black and Devereux (2011), Blanden (2013), Corak (2013; 2006), Mulligan (1997),

\(^8\) Zimmerman (1992) offered a contemporaneous analysis that did not use the Panel Study of Income Dynamics.
and Solon (2002; 1999) reveal the important role that the construction and availability of new data plays in offering opportunities to both revisit longstanding issues, and imagine new possibilities. In Canada this took the form of the development of inter-generationally linked income tax data by Corak and Heisz (1999), which were used to estimate an intergenerational elasticity of about 0.2—a finding similar to Fortin and Lefebvre (1998) who use Census data—and at least half the magnitude of the best available estimate for the US (Solon 1992; Mazumder 2005a (2005b); Zimmerman 1992). The most important recent advance in the American literature has been the development and use of similar tax-based data from the Internal Revenue Service by Chetty, Hendren, Kline, and Saez (2014) and Mitnik et al. (2015), confirming intergenerational income elasticities of at least 0.4 and approaching 0.6, but also suggesting the use of a wider variety of intergenerational statistics and, as illustrated by Chetty, Hendren, Kline, and Saez (2014), sub-national analysis as an important new research possibility.

We update and re-construct the Canadian data to permit a direct comparison with the regional analysis of the United States offered by Chetty, Hendren, Kline, and Saez (2014). This involves examining recent labour market outcomes (in 2011 and 2012), but for a younger cohort born in 1980 and 1982. Our development of these new Canadian data are meant to line up as closely as possible to the American equivalent, and as such we are not in a position to examine intergenerational income mobility at the most appropriate stage in the life cycle, when the children are in their late 30s to mid 40s. As such, we follow Chetty, Hendren, Kline, and Saez (2014) in focusing on different measures of rank mobility, which tend to be stable by the time individuals reach their early 30s (Corak 2016b Table 3; Nybom and Stuhler 2015).

Canadians file income taxes as individuals, but the tax form—referred to as the T1 Form—requires identifying information about a spouse or common-law partner to be provided. We are able to construct a “family income” variable for both parents and children (in adulthood) and their spouses in the manner of Chetty, Hendren, Kline, and Saez (2014): the total (before tax) income of both partners in the household using the Canada Revenue Agency definition of total income (all market sources of income plus all government transfers).

---


10Chen, Ostrovsky, and Piraino (2017) offer an analysis of cohorts born in the mid 1960s who are followed up to 2008 in a manner that takes full account of the life-cycle biases discussed by Grawe (2006), Jenkins (1987), and Haider and Solon (2006). They estimate the father-son elasticity to be 0.32, and offer a Canada-US comparison from this perspective. Corak (2016b) uses the data on this older cohort in the same manner we do, offering regional estimates for the 266 Census Divisions defined in the 1986 Canadian Census, with the more appropriate point in the life-cycle allowing estimates of the regression to the mean model in incomes.

11More specifically, this consists of: earnings from an employer-employee relationship as indicated on the T4 form, and including commissions, interest and investment income, other employment income, other income, pension and super-annuity income, rental income, self-employment net income (from business, commissions, farming, fishing or professional activities), as well as capital gains and losses, and dividends (taxable and appropriately grossed up to reflect changes in tax treatment). It also includes: Old Age Security pension, Canada/Quebec Pension Plan benefits, and Employment Insurance benefits. In addition, for the years from 1986 onward total income is defined to include tax credits associated with the Goods and Services Tax.
Parent total income is averaged over the tax years 1996 to 2000, a potential of up to five years for each parent. If a parent’s T1 record is not found for a particular year, income is assigned a value of zero. The sum of the father’s and mother’s total income is averaged: all the income is added up and divided by 10 for an intact two-parent family, or by five if only one parent is present in the year the link is made with the child. If a parent dies, the denominator is adjusted to reflect the number of years the parent was alive. To be included in the analytical sample a parent must appear in the T1 files at least once between 1996 and 2005, depending on the cohort and the year linked. If a child is recorded as having two parents in the year linked but the parents subsequently separate or divorce, then the individual incomes of the mother and father continue to be added together to make up total income.

The intergenerational link of tax returns between parents and children requires that the child have a Social Insurance Number while living at home. Some parents obtain a Social Insurance Number for their children early in their lives, but others do not and it is usually obtained at some point later when the child enters the labour market, whether part-time as a teenager or even later. The link between parent and child Social Insurance Numbers is made by using the algorithm described in the appendix to Corak and Heisz (1999). In short, this involves the use of the T1 Family File, a Statistics Canada database of the universe of T1 Forms that uses name, address, and other administrative data associated with the delivery of child benefits, to link family members. Some of the children in these data have Social Insurance Numbers, but others are imputed, that is, they are known or presumed to be present but no identifying information is available. For this reason it is, on its own, not an appropriate data source for intergenerational research. The algorithm searches for the tax files of children who may have filed when they were between the ages of 16 and 19, while still living at home. If a link is not made, then the search is repeated for 17 to 20 year olds in the next year, and successively for a total of five years when the children are 20 to 23 years of age. In other words, this requires children who don’t otherwise have a Social Insurance Number to have obtained one through filing an income tax form while living at home. We capture from 73 percent to 85 percent of the individuals in the two birth cohorts we focus upon (based upon the estimates of these populations from the Canadian Census), those born in 1980 and 1982.

The analytical sample we create from these data is aligned with the core sample in Chetty, Hendren, Kline, and Saez (2014). Their analysis uses children born between, inclusively, 1980 and 1982, whose parents are identified and have a positive mean income between 1996 and 2000. As such, we should stress that our analytical sample is intended not to be the “best” possible data for the analysis of intergenerational mobility in Canada, but rather the “best” possible data for a comparison with the United States.

and Harmonized Sales Tax; from 1992, Net Federal supplements, Social Assistance payments and Workers’ Compensation payments; from 1996, Guaranteed Income Supplements; and from 1998, limited partnership income and Registered Retirement Savings Plan income.

What I’m hearing from Marie is that this may need to be reworded to more accurately reflect what was actually done. Is that right?

These linkages were all performed by officials at Statistics Canada, who also derived and attached weights to the file based upon Census counts stratified by region, gender, and family income. Our analysis uses the weighted data.
Some considerations by Statistics Canada during construction of the data leads us to use only the 1980 and 1982 birth cohorts, the 1981 birth cohort not being made available. The adult income of these children is defined in the same way as for parents on the basis of the average total income during 2011 and 2012, including any spousal income if a spouse is present.\textsuperscript{14} Children not filing a T1 Form in both 2011 and 2012 are dropped from the analytical file. All dollar amounts are adjusted to 2012 Canadian dollars using the national Consumer Price Index, and then to US dollars using the Purchasing Power Parity rate for 2012 produced by the Organisation for Economic Co-operation and Development (1.28 Canadian dollars per US dollar).

A child’s geographic location is based upon the postal code provided in the T1 form in the year the child is linked to his or her parents. This is consistent with Chetty, Hendren, Kline, and Saez (2014), who use the 1996 location for 96 percent of their sample. The postal code is converted to Census geography codes using Statistics Canada’s Postal Code Conversion File.\textsuperscript{15} The analysis is based on the Census Division, which roughly corresponds to a municipality, the sub-provincial level of government used to deliver provincial and municipal services. Not all provinces have such a level of government, and in these cases Statistics Canada in consultation with provincial counterparts defines an equivalent geographic area so that the entire country is covered. In the 1996 Census there are 288 Census Divisions. Conceptually, this geographic unit is somewhat narrower than the Commuting Zones used in the American data, but not necessarily so practically, as some Census Divisions cover a significant geographic area.\textsuperscript{16} This implies that our sub-national analysis involves, together with the 740 Commuting Zones in the US data, over 1,000 regions.

We also use the one-in-five Census micro-data for the 1996 Census, the data from the so-called “Long Form” that roughly 20 percent of Canadians are required to complete, to develop community profiles for each of the Census Divisions.\textsuperscript{17} The appendix lists and defines all of these variables, which are meant, to the extent possible, to replicate the set of variables constructed from the US Census by Chetty, Hendren, Kline, and Saez (2014).\textsuperscript{18}

\textsuperscript{14}The original text at this point says: “We only use married spouses even if a non trivial fraction of children in 2011 and 2012 report having a common-law partner, to be close to Chetty’s et al.’s definition, who do not have information on common-law spouses (filling separately).” I’m surprised to hear this. I thought we talked this through and agreed that this was an important difference between Canada and the United States that should not be ignored, but incorporated into the analysis. The treatment of common-law in Canada, and the benefits that they have rights to is very different, and very different at the lower end of the income distribution. I also can’t imagine what this choice has done to our sample in Québec. Also the original text simply says “non trivial fraction” without making clear just how big this is, and in particular how it varies across regions. I think we really need to rethink this.

\textsuperscript{15}Ask Yuri what version was used, and place in this footnote.

\textsuperscript{16}Page 4 of the original text is somewhat confusing to me in a couple of instances. In particular, the page starts by saying that we are using a 1996 cohort for the 1980 birth cohort, and a 2001 cohort for the 1982 birth cohort. It repeats this at the bottom of the section 1.1 when it states that there “are 288 Census divisions in the 1996 and 2001 Census geographies.” I don’t understand why the 1980 cohort is said to come from the 1996 cohort, and then the 1982 birth cohort comes from a cohort five years later in 2001. What is this about?

\textsuperscript{17}Page 5 of the original text says “thus covering 4% of the Canadian population.” I don’t understand why we are saying that the long form data covers 4% of the population?

\textsuperscript{18}We need more clarity in the original text on these variables. For example, the variables
Theory offers only a partial guide on the choice of statistics to measure intergenerational mobility, and to some degree research is increasingly informed not just by the limitations and opportunities of available data, but also by public policy discussion. Becker and Tomes (1986, 1979) focus empirical attention on the regression to the mean model of incomes:

\[ y_{i,t} = \alpha + \beta y_{i,t-1} + \epsilon_i, \]

where once again \( y_{i,t} \) represents the natural logarithm of permanent income of a member of family \( i \) in generation \( t \), and \( \alpha \) and \( \beta \) are parameters usually estimated by least squares that measure absolute and relative income mobility. As stressed, only under very specific circumstances is \( \beta \) a structural parameter, and it should more accurately be understood to be a broad summary indicator of intergenerational income mobility reflecting both the correlation of standardized incomes and differences in the variance of incomes between generations. This said, even as a descriptive statistic it does not capture all dimensions of the process, rank mobility not being explicitly measured, and the linearity assumption implying that the rate of mobility is the same across the entire parental income distribution. As such, we are agnostic as to the appropriate measure of mobility. Like Chetty, Hendren, Kline, and Saez (2014) we downplay the intergenerational elasticity in large measure because of the potential for life-cycle biases.\(^{19}\)

Our focus is on certain cells of the quintile transition matrix, in particular the chances that a child born to bottom quintile parents will rise to the top of the quintile—so-called rags to riches mobility—and the chances that he or she will grow up to in turn be a bottom quintile adult—the intergenerational cycle of low income. These are referred to, respectively, as \( P_{1,5} = \Pr\{Y_t \in \text{top} | Y_{t-1} \in \text{bottom}\} \) and \( P_{1,1} = \Pr\{Y_t \in \text{bottom} | Y_{t-1} \in \text{bottom}\} \), which we calculate for the country as a whole and for each sub-national unit. But more generally, we also focus on summary indicators of rank mobility derived from a rank-rank regression:

\[ R_{i,j,t} = a_j + b_j R_{i,j,t-1} + \epsilon_{i,j}, \]

where \( R_{i,j,t} \) refers to the percentile rank in the national income distribution of an individual in family \( i \), belonging to generation \( t \), from region \( j \) when the analysis is sub-national. To be clear, we are following Chetty, Hendren, Kline, and Saez (2014) in defining a child’s geography on the basis of where he or she lived as a teenager, but our measure of permanent income is determined without regard to where the child may be living as an adult, a little more than a decade later. Geographic mobility is embodied in this analysis, and children and parents—regardless of where they live—are placed in the national income distributions to determine their rank, not in the local income distribution. The parameters of this model are estimated using least squares, with \( a_j \) offering a measure of absolute mobility—the expected rank of a child raised by bottom percentile parents—and \( b_j \) a measure of relative rank mobility—the increase in a child’s rank for every percentile increase in the parents’ rank.

Canadian incomes are placed both in the Canadian income distribution, and also in the American income distribution. The sub-national analysis is based entirely on the latter ranking. In this case the assigned ranks refer to the percentile rank in the United States, and the transition matrices derived from this ranking are not strictly transition matrices as the are described as being “fractions” but it is never stated fractions of what? Also we might delete all of this if we don’t make substantive use of these variables, and correlations in a comparative way.

\(^{19}\)This said, we do characterize our communities by the average parental income, an important correlate of the expected adult income of children.
Table 1: Selected percentiles of the parent and child income distributions in Canada and the United States: US (2012) dollars

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Parents</th>
<th></th>
<th>Children</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Canada</td>
<td>United States</td>
<td>Canada</td>
<td>United States</td>
</tr>
<tr>
<td>1</td>
<td>1,810</td>
<td>1,700</td>
<td>960</td>
<td>-43,800</td>
</tr>
<tr>
<td>5</td>
<td>9,070</td>
<td>9,200</td>
<td>5,730</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>13,450</td>
<td>15,000</td>
<td>9,650</td>
<td>2,300</td>
</tr>
<tr>
<td>20</td>
<td>22,590</td>
<td>24,900</td>
<td>17,860</td>
<td>11,000</td>
</tr>
<tr>
<td>50</td>
<td>52,890</td>
<td>59,500</td>
<td>39,980</td>
<td>34,600</td>
</tr>
<tr>
<td>80</td>
<td>89,750</td>
<td>107,900</td>
<td>80,140</td>
<td>74,400</td>
</tr>
<tr>
<td>90</td>
<td>113,730</td>
<td>144,500</td>
<td>108,270</td>
<td>99,900</td>
</tr>
<tr>
<td>95</td>
<td>140,000</td>
<td>194,300</td>
<td>132,000</td>
<td>125,300</td>
</tr>
<tr>
<td>99</td>
<td>246,760</td>
<td>420,100</td>
<td>183,920</td>
<td>193,300</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>1,408,800</td>
<td>0</td>
<td>408,400</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations, Chetty et al (2014) online tables.

rows and columns are not constrained to add to one. The online appendix to Chetty, Hendren, Kline, and Saez (2014) lists the national marginal income distributions by percentile.20

The construction of our analytical file differs from Chetty, Hendren, Kline, and Saez (2014) in the treatment of very low incomes, an explicit decision made to reflect important substantive differences between the countries. Table 1 shows average parent and child incomes for selected percentiles of the income distributions in each country. These percentiles are based upon the national cut-offs. The average income in the bottom percentile of the child’s income distribution in the analytical file Chetty, Hendren, Kline, and Saez (2014) use is -$43,800, and all percentiles up to the sixth have an average income of zero. At first look this may appear odd as the objective is to estimate permanent income, and it is hard to rationalize how anyone can have a negative or even zero permanent income. But Chetty, Hendren, Kline, and Saez (2014) motivate the decision to keep these very low incomes because of high incarceration rates among some groups of the US population in this cohort, excluding these groups would likely induce a more severe sample-selection bias at the low end of the income distribution.21

This is not a consideration in Canada, and previous research with these data suggests that very low market incomes are more likely to reflect measurement error. In particular, there are notable spikes in market incomes at one and two dollars (Grawe 2006). Parent-child pairs in which the average total child income is less than $500 US are dropped from the

20Specifically, we use Table 2 retrieved at http://www.equality-of-opportunity.org/index.php/data, which offers the average parent and child incomes for each percentile of the respective income distributions, rounded to the nearest $100. We define the percentile cutoffs to be the midpoint between two means.
21Correspondence with Raj Chetty.
Canadian data.\textsuperscript{22} Research based on survey data for the 1990s finds that when Canadian children are placed in the US income distribution they are much less likely to be in the bottom decile of the US income distribution than their American counterparts, reflecting differences in the polarization of labour markets, the structure of families, and the more generous Canadian system of income support.\textsuperscript{23}

Table 1 also shows that while Canadian parents tend to have lower average income within each percentile, other than the very bottom percentile, this is markedly so higher up in the income distribution. To be in the top 20 percent, and certainly to be in the top five percent, and the top one or two percent, implies having a good deal more income in the United States than in Canada. Children in the United States are raised in a context of greater top-end income inequality. At the same time, the adult incomes of Canadian children are higher at each percentile than in the United States, in part reflecting the very different outcomes at the lowest end. This said, top incomes are also higher in the United States.\textsuperscript{24}

4 Between country comparisons

Canada is characterized by more intergenerational mobility than the United States, whether in terms of intergenerational income mobility, intergenerational rank mobility, or in terms of particularly policy relevant elements of the intergenerational transition matrix, like rags to riches movement and intergenerational cycles of low income. The Canadian data we construct and use probably allows a more accurate comparison with the United States than any country-wide comparison that has been made in the literature to date: tax-based administrative data, used to define similar measures of income, and coming close to covering the total population of similarly defined birth cohorts. We confirm the general notion in the literature that views Canada as being significantly more intergenerationally mobile, but in motivating the remainder of our analysis, note that this finding may be less relevant than previously thought.\textsuperscript{25}

Least squares estimates of relative income and rank mobility, the estimates of the

\textsuperscript{22}This is the preferred decision rule in the literature, but some of our analysis is also conducted with an alternative decision rule, assigning an income of $500 US to all observations with less than that amount, and another with a cut-off of $1,000 US. The analytical sample is constructed to also require the average parent income from 1996 to 2000 to be $500 US or higher. There are 9,337 observations, or 1.67 percent of the sample, excluded on this basis.

\textsuperscript{23}To be clear, however, this result does not account for significant in-kind support in the United States. This limitation also applies to analyses based on income tax data. See Corak, Curtis, and Phipps (2011).

\textsuperscript{24}This should be interpreted loosely as these data make no account for household size, and appropriate adjustments through an equivalence scale. But the general point does accord with Corak, Curtis, and Phipps (2011) who do make such adjustments, though with survey data, which may not as accurately capture very low and very high incomes.

\textsuperscript{25}The one dimension of this literature to which we do not speak involves rank mobility measured in terms of exceeding or falling below the specific rank of a parent. This is examined in Corak, Lindquist, and Mazumder (2014) who find that when a son’s percentile rank is compared to his father’s rank, Canadians are distinguished from Americans in that having a top ranking parent is more likely to imply downward mobility in this relative sense. This study is based upon percentile ranks defined on the basis of national income distributions.
Table 2: Least squares estimates of relative intergenerational income and rank mobility, Canada and United States

<table>
<thead>
<tr>
<th>Income definitions</th>
<th>Parent</th>
<th>United States</th>
<th>Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intergenerational income elasticities</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1  Excluding incomes below $500</td>
<td>Logarithm of family income</td>
<td>0.344</td>
<td>0.252</td>
</tr>
<tr>
<td>2  Recoding incomes below $500 to $500</td>
<td>Logarithm of family income</td>
<td>0.300</td>
<td></td>
</tr>
<tr>
<td>3  Recoding zero incomes to $1</td>
<td>Logarithm of family income</td>
<td>0.618</td>
<td>0.381</td>
</tr>
<tr>
<td>4  Recoding incomes below $1,000 to $1,000</td>
<td>Logarithm of family income</td>
<td>0.413</td>
<td>0.289</td>
</tr>
<tr>
<td><strong>Relative rank mobility</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5  Rank of family income in US</td>
<td>Rank of family income in US</td>
<td>0.341</td>
<td>0.220</td>
</tr>
<tr>
<td>6  Rank of family income in US</td>
<td>Rank of family income in US 1999 to 2003</td>
<td>0.339</td>
<td>0.212</td>
</tr>
<tr>
<td>7  Rank of family income in US</td>
<td>Rank of top parent income in US</td>
<td>0.312</td>
<td>0.191</td>
</tr>
<tr>
<td>8  Rank of individual income in US</td>
<td>Rank of family income in US</td>
<td>0.287</td>
<td>0.250</td>
</tr>
<tr>
<td>9  Rank of individual income in US excluding incomes below $500</td>
<td>Rank of family income in US</td>
<td>0.282</td>
<td>0.201</td>
</tr>
</tbody>
</table>

Note: All estimates are statistically significant, with the maximum standard error being 0.004.

parameters $\beta$ and $b$, are presented in Table 2 for a number of different sample selection rules in a way that follows some of the findings in Chetty, Hendren, Kline, and Saez (2014 Table I), from which the estimates for the United States are drawn. The first panel offers the intergenerational elasticities, and the second the rank-rank slope estimate, where ranks are determined by the within-country percentile distribution.

The Canadian intergenerational income elasticity is lower. All the coefficients, and all the differences between the countries, are statistically significant, the largest standard error in the table being 0.0038 (that for the Canadian estimate in row 3), with most being a little lower than 0.002 for Canada, and in the range of 0.0003 to 0.0004 for the United States. The Canadian estimates of the elasticities are similar to those in the existing literature, but not directly comparable to them: our income definition is unique, the samples include both men and women, and are, at least for a comparison with the best current estimates in Chen, Ostrovsky, and Piraino (2017), early in the life cycle.\footnote{The exception is a comparison to Corak (2016b), who uses the same administrative data source and income definitions, but with an older cohort of Canadians born in the mid 1960s, whose outcomes are measured over a five year period ending in 2008. Interestingly, he reports a Canada-wide elasticity that, at 0.20, is actually a bit lower than the lowest estimate offered in the first panel of Table 2. Both of these sets of estimates are much lower than those offered by Chen, Ostrovsky, and Piraino (2017), but these authors}

The estimates of relative rank mobility in Table 2 imply that a child raised by top percentile parents in the United states will rank about 30 to almost 35 percentiles higher in the income distribution than a bottom percentile child, but in Canada this difference would, at 20 to 25 percentiles, be a full decile lower.\footnote{Figure 2 contrasts rank-rank mobility between the two countries in a way that permits a}
The slope is flatter in Canada, and the intercept is higher, the combination implying that Canadian children will rank higher than their American counterparts until about the top quintile of parent income is reached. The expected rank of a Canadian child raised by bottom quintile parents is about the median, but for an American closer to the 40th percentile. To reach a similar point on the US income ladder an American child would have to have parents who ranked as high as the 45th percentile. The “middle class” is within easier rich for low income Canadian children, than it is for low income Americans.

This said, the information in Figure 2 also implies that American children raised by top quintile parents are more likely to rank higher than their Canadian counterparts, though the rank-rank relationship tends to show more non-linearity in Canada in a way that top percentile Canadians can expect to see their children attain the same rank as top percentile

also suggest that their results should be understood to be about half as great as comparable US estimates.

Corak (2016b Table 3) reports a rank-rank slope for Canada of 0.242 when child outcomes are measured at 35 to 48 years of age, and 0.240 when measured for the same cohort at 31 and 32 years of age. This is exactly in the range of our estimates, suggesting that percentile ranks are established by the time children reach the early 30s—as Chetty, Hendren, Kline, and Saez (2014) speculate with US data, and as Nybom and Stuhler (2015) show with Swedish data—and also, when compared to our findings, suggesting that the rank-rank slope has been stable over time, or more specifically stable between cohorts born in the mid 1960s and those born in the early 1980s.

Would Charles prepare a spreadsheet with the data used in this figure so that I can create it in R, and it will have a cleaner look and feel. Also I don’t understand why the slope estimates offered in the figure don’t line up exactly with an estimate that is in the previous table. We should also report all the regression results, that include also the estimate of the intercept.
Table 3: **Quintile transition matrices: Canadians in the Canadian income distribution, Canadians in the US income distribution, and Americans in the US income distribution**

<table>
<thead>
<tr>
<th>Child’s quintile</th>
<th>Parent’s quintile</th>
<th>Bottom</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>Top</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>United States</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bottom</td>
<td>33.7</td>
<td>24.2</td>
<td>17.8</td>
<td>13.4</td>
<td>10.9</td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>28.0</td>
<td>24.2</td>
<td>19.8</td>
<td>16.0</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>18.4</td>
<td>21.7</td>
<td>22.1</td>
<td>20.9</td>
<td>17.0</td>
<td></td>
</tr>
<tr>
<td>Fourth</td>
<td>12.3</td>
<td>17.6</td>
<td>22.0</td>
<td>24.4</td>
<td>23.6</td>
<td></td>
</tr>
<tr>
<td>Top</td>
<td>7.5</td>
<td>12.3</td>
<td>18.3</td>
<td>25.4</td>
<td>36.5</td>
<td></td>
</tr>
</tbody>
</table>

| **Canada in the Canadian income distribution** |                  |        |        |       |        |     |
| Bottom | 30.0 | 20.9 | 17.3 | 14.8 | 13.5 |     |
| Second | 23.0 | 23.0 | 20.7 | 17.8 | 14.4 |     |
| Third  | 18.3 | 21.4 | 21.9 | 21.0 | 17.7 |     |
| Fourth | 15.1 | 18.8 | 21.4 | 23.4 | 22.7 |     |
| Top    | 13.6 | 15.9 | 18.6 | 23.1 | 31.7 |     |

| **Canada in the US income distribution** |                  |        |        |       |        |     |
| Bottom | 17.0 | 9.2  | 7.2   | 6.1   | 5.8   |     |
| Second | 21.9 | 17.5 | 14.5 | 12.1  | 10.3  |     |
| Third  | 22.1 | 23.2 | 21.5 | 19.2  | 16.8  |     |
| Fourth | 23.1 | 28.2 | 30.0 | 29.5  | 27.6  |     |
| Top    | 15.9 | 21.9 | 26.8 | 33.0  | 39.4  |     |

Source: Authors’ calculations, and Chetty et al. (2014, Table II)

Americans. There also seems to be more pronounced non-linearities at the bottom of the Canadian income distribution, certainly over the course of the bottom decile of the parent distribution and possibly up to the bottom quintile. To some degree the linear rank-rank model is a less accurate summary indicator of rank mobility in Canada than Chetty, Hendren, Kline, and Saez (2014) suggest it is for the United States.

Table 3 offers another perspective on these movements by offering estimates of the quintile transition matrices for the two countries, including two estimates for Canada, the first, in Panel 2, based on the Canadian quintile cut-offs, and another, in Panel 3, based upon the ranks that would be ascribed to Canadian parents and children if they were placed in their respective American income distributions. It should be noted, by first focusing on Panels 1 and 2, that for a large segment of the population in these two countries—for families from the 20th to the 80th percentile—there is a good deal of intergenerational mobility. In fact, in these middle quintiles family income seems very loosely related to child outcomes, the quintile transition probabilities hovering a bit above and a bit below 0.20. Intergenerational
mobility may contribute to middle class anxiety, parents not being able to greatly influence
their child’s station in life. There certainly are distinct gradients in this range, but these are
notably sharper for children from bottom quintile and top quintile parents.

The children of top quintile parents in the United States have an almost 37 percent
chance of staying in the top quintile as adults, and face only about a one-in-ten chance of
falling to the bottom fifth of the income distribution. In Canada, this tilt is also present but
not as extreme: just less than a one-third chance of remaining in the top, and not quite a 14
percent chance of falling to the bottom.\footnote{Nybom and Stuhler (2015) and O’Neill, Sweetman, and Van de gaer (2007) note that measurement
error in both parent and child incomes influences transition matrices in a non-classical way, leading to an
overstatement of mobility in middle parts of the distribution, and an understatement at the extremes. While
our correction for measurement error in child incomes is imperfect—an averaging of only two years of child
incomes—we don’t expect this to play out significantly for quintile transition matrices, but for this reason
avoid a detailed look at the extremes of the percentile transition matrices.}

A similar pattern is displayed at the other extreme of the income distribution. In Canada,
the mobility of bottom quintile children is in magnitude almost exactly the inverse of top
quintile children: a 30 percent chance of an intergenerational cycle of bottom income, and not
quite a 14 percent chance of rags to riches movement. In the United States, bottom quintile
mobility is more challenging, these children facing a 34 percent chance of staying in the
bottom, and only a 7.5 percent chance of rising to the very top, the lowest probability among
all those listed in Panels 1 and 2. The mobility of low income Canadian children is even
greater when they and their parents are ranked according to the US income distribution.\footnote{Why don’t the rows of the transition matrix in panel 2 add up to 100, and can you explain
whey they don’t add up to 100 in panel 3?}

The estimates in Panel 3 show that they are almost half as likely to stay in the bottom as
their American counterparts (17 percent versus 33.7 percent), and a bit more than twice as
likely to rise to the top quintile (15.9 percent versus 7.5 percent).

All three dimensions of this cross-country comparison, but in particular this last result
related to the prospects of low income children, put the so-called “American Dream” in sharp
relief, and offer stronger evidence to support the growing perception in the literature that
this Dream—if it is defined in terms of income mobility and the opportunity to, in terms of
Figure 1, “succeed regardless of family background”—is more of a reality in Canada than
the United States. But the sub-national research strongly and thoroughly documented by
Chetty, Hendren, Kline, and Saez (2014) presents a challenge to this conclusion: the degree
of intergenerational mobility varies significantly within the United States, with some regions
showing much more mobility than that recorded even for the most mobile countries. Between
country comparisons have less relevance unless they are also accompanied by within country
comparisons.

5 Between and within country comparisons

There is less variation in some of these markers of intergenerational mobility across the
288 Canadian Census Divisions, than across the 740 American Commuting Zones. But
the two countries are not sharply distinguished by the border between them. The border demarcates the most populous region of Canada—the Quebec City-Windsor corridor—from the Northeastern seaboard and the Great Lakes region of the United States. In part this reflects a greater degree of labour market inequality in these parts of the United States. But the other reason the national statistics suggest Canada is more mobile has to do with the fact that there is a high concentration of low mobility in the Southern United States. While certain regions of Northern Canada share this limited mobility, they make up a much smaller share of the population, and do not influence the national statistics to the same degree.

Figure 3: The intergenerational cycle of low income: Bottom to bottom quintile transition probabilities in Canada and the United States show distinct patterns on either side of the national border.

Figure 3 maps Canada and the United States according to the value of $P_{1,1}$, defined by the quintile cut-offs of the American income distribution. The 1,000 regions of the two countries are categorized into six groups, with the most mobile cut-off placed at 0.2, and the upper cut-off at 0.4. There is a clear border effect, the possible exception extending from the southeastern parts of Saskatchewan and neighbouring North Dakota through the areas around Manitoba and Minnesota, and into the adjoining areas of northwestern Ontario, Wisconsin, and Michigan. But elsewhere the national border is clearly distinguished, and particularly between the lower Great Lakes regions all the way through to the eastern seaboard. Most of the regions in the eastern part of the United States are characterized by a transition probability of 0.3 or higher, and while Canada does have some pockets having a greater than 0.4 chance of intergenerational poverty, they have a share in the Canadian population that
amounts to **XX percent**, compared to similar regions having a population share of **YY percent** in the United States.

Rags to riches mobility, $P_{1.5}$, is not as sharply distinguished between the two countries. Figure 4 maps this probability. Southern Ontario—the region north of lake Ontario and Lake Erie—displays a greater chance of bottom to top quintile mobility than adjacent regions in Michigan, Ohio, and New York State, most being categorized in the 0.10 to 0.15 range as opposed to 0.05 to 0.10. However, the areas of Quebec adjacent to New York State, Vermont, and New Hampshire display an even lower probability. Regions with very low chances of escaping low income and rising to the very top quintile are concentrated in the American south, where in many Commuting Zones the probability is less than five percent. Again, there are pockets of Census Divisions with similarly low chances in Canada, but they are fewer in number and population share.\(^\text{31}\)

![Figure 4: Rags to riches mobility: Bottom to top quintile transition probabilities are not sharply distinguished on either side of the national border](image-url)

These comparisons are hard to generalize for at least two reasons. First, these are only two possible measures of intergenerational mobility. As important as they are for a policy-relevant group, they do not embody all dimensions of mobility. Second, while the maps are striking illustrations leaving strong impressions, they are based upon arbitrary—though intuitively appealing—groupings. For these reasons we focus on an alternative descriptive approach that allows us to simultaneously use a total of six available indicators. We cluster

---

\(^{31}\)Offer these numbers and shares in a footnote here.
the 1,000 regions with K-means, an unsupervised machine learning algorithm. To be clear, this is not an exercise in prediction, but rather one of clustering units into a predetermined number of groups according to a metric of all available mobility indicators. We chose the number of groupings according to our purpose. This is an exercise in description, intended to serve a communication purpose and set directions for more detailed analysis. It permits us to be agnostic about the choice of indicator in order to highlight which regions of Canada and the United States are more similar to each other than to other parts of the country. In this way we assess more rigorously whether, and to what extent, the national border falls out as the aftermath of the categorization.

Briefly, K-means partitions observations into groups which minimize the within-cluster variation, defined as the sum of all the pairwise squared Euclidean distances. The optimization problem involves:

$$\min_{C_1,\ldots,C_K} \left\{ \sum_{k=1}^{K} \frac{1}{|C_k|} \sum_{i \in C_k} \sum_{j=1}^{p} (x_{ij} - \bar{x}_{ij})^2 \right\},$$  \hspace{1cm} (4)

where $K$ is the number of pre-defined clusters, $C_k$ denotes a cluster, $i$ and $j$ represent particular observations within a cluster, and $x_{ij}$ is one of the $p$ features of the data (Hastie, Tibshirani, and Friedman (2009); James et al. (2013)). We use six such features: relative rank mobility as measured by the least squares estimate of the rank-rank slope ($b_{ij}$), absolute mobility as measured by the predicted rank at the 25th percentile of the parental income distribution ($\bar{r}_{25,c}$), the mean and the median of the parental income in the region, and the two quintile transition probabilities highlighted in Figures 3 and 4, $P_{1,1}$ and $P_{1,5}$. It is in this sense that we claim to be agnostic as to the appropriate statistic measuring intergenerational mobility. The algorithm solving equation 4 is described in James et al. (2013) and Hastie, Tibshirani, and Friedman (2009), and we use R package XXXX to obtain our results.

Figure 5 maps the American Commuting Zones and the Canadian Census Divisions when they are forced to belong to only two clusters. In other words, with this map we are seeking to determine if the K-means algorithm would classify these 1,000 into two groups according to the international border. It clearly would not: the vast majority of Canadian Census Divisions containing almost the entire population have more in common with a large swath of the American south, than with some parts of western Canada and large parts of the United States running from the eastern seaboard, through the mid-west, to the west coast.

The first panel of Table 4 offers the total number of regions and populations of these two clusters, and also average unweighted values of the six mobility measures. The most populous cluster includes over 240 million children, but it does not cover the most populous regions of

---

32 Change this to the intercept of the rank-rank regression: absolute mobility as measured by the least squares estimate of the intercept of the rank-rank regression $a_j$.

33 My understanding is that these variables are all standardized. Is that right?

34 This is a “top-down” approach to clustering requiring the number of clusters to be pre-determined. We also undertook our analysis using agglomerative (bottom-up) clustering, which does not require the number of clusters to be pre-specified, and can be represented as a dendrogram offering an analytical way of determining the appropriate number of clusters. We need to do this and to make a suggestion on the appropriate number of clusters.
Figure 5: The Canada-United States border would not be chosen by a machine learning algorithm minimizing within-cluster variance of six indicators of intergenerational mobility
Table 4: Summary statistics of intergenerational mobility measures, for clusters of Canadian Census Divisions and American Community Zones as determined by K-means

<table>
<thead>
<tr>
<th>Cluster identifier of regions of children</th>
<th>Population of children (thousands)</th>
<th>Rank mobility absolute $a$</th>
<th>Rank mobility relative $b$</th>
<th>Transition probability $P_{15}$</th>
<th>Transition probability $P_{11}$</th>
<th>Parent Income Average</th>
<th>Parent Income Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>562</td>
<td>242,860</td>
<td>40.9</td>
<td>0.347</td>
<td>7.5</td>
<td>33.9</td>
<td>88,770</td>
</tr>
<tr>
<td>2</td>
<td>423</td>
<td>67,824</td>
<td>45.5</td>
<td>0.236</td>
<td>11.8</td>
<td>23.8</td>
<td>72,747</td>
</tr>
</tbody>
</table>

2. Four clusters

<table>
<thead>
<tr>
<th>Cluster identifier of regions of children</th>
<th>Population of children (thousands)</th>
<th>Rank mobility absolute $a$</th>
<th>Rank mobility relative $b$</th>
<th>Transition probability $P_{15}$</th>
<th>Transition probability $P_{11}$</th>
<th>Parent Income Average</th>
<th>Parent Income Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>170</td>
<td>9,449</td>
<td>52.5</td>
<td>0.243</td>
<td>19.2</td>
<td>18.2</td>
<td>69,467</td>
</tr>
<tr>
<td>2</td>
<td>256</td>
<td>33,006</td>
<td>43.2</td>
<td>0.236</td>
<td>9.7</td>
<td>21.2</td>
<td>53,006</td>
</tr>
<tr>
<td>3</td>
<td>276</td>
<td>212,169</td>
<td>42.0</td>
<td>0.328</td>
<td>8.4</td>
<td>33.1</td>
<td>96,891</td>
</tr>
<tr>
<td>4</td>
<td>283</td>
<td>56,060</td>
<td>38.9</td>
<td>0.369</td>
<td>6.0</td>
<td>34.9</td>
<td>62,959</td>
</tr>
</tbody>
</table>

Note: Population refers to cluster totals, and other table entries are unweighted means.

Canada, which are included in the cluster that is more mobile according to the absolute and relative rank indicators, and the two transition probabilities. This smaller cluster is also a cluster of regions with on average lower and more equal parent incomes, as indicated by the average and median incomes.

Obviously, this is not our preferred number of clusters, and should not be understood to accurately capture the variance in the data. But that, in part, is the point. When sub-national indicators of mobility are available for analysis they do not naturally lead to the conclusion that a simple between country analysis is an appropriate way to view intergenerational mobility. We conduct this analysis repeatedly with up to eight clusters being pre-specified, and judge that a four-cluster mapping is the best way to communicate the regional variation in mobility for these 1,000 regions. This is mapped in Figure 6, and shows that the clustering continues to occur on either side of the international border between the most populous regions of the two countries, with a clear line being drawn along the Great Lakes and Saint Lawrence River regions, one that continues to distinguish the New England States from the Maritime Provinces. The border is also apparent in the very west of the two countries. One cluster—that labelled 2 in the Figure—lies largely on the Canadian side of the border, and another—that labelled 3—lies uniquely on the American. But the two other clusters cross the border, one running through the mid west and into Saskatchewan and most of Alberta, another grouping the Southern United States with Northern parts of Canada and some regions in the Maritimes.

The second panel of Table 4 summarizes the mobility measures in this four-cluster map. Cluster 3 contains by far the largest fraction of the total number of children, and covers significant portions of the northeast, mid-west, and western coastal areas of the United States.
Figure 6: A four cluster mapping shows that some regions lie largely on either side of the Canada-United States border but that others are not confined to one country.
It is distinguished from Cluster 2, which covers large parts of Central and Eastern Canada and also British Columbia. The contrast between these two regions is the sharpest cross-national difference. Mobility, in all the dimensions we measure, is higher in the Canadian dominant cluster, and parental income and top-end income inequality much lower. Although absolute rank mobility is about the same in both clusters, relative rank mobility is much lower in Cluster 2, with a top percentile parent having children who can be expected to rank about 24 percentiles higher than a bottom percentile parent, a decile lower than in Cluster 3. The probability of an intergenerational cycle of low income is on average 21.2 percent in Cluster 2, and fully one-third in Cluster 3.

The two clusters that span the border are areas of both higher and lower mobility. Cluster 1 running through the western provinces of Alberta and Saskatchewan into the midwest American states is, as would be imagined from the information in Chetty, Hendren, Kline, and Saez (2014), the highest mobility area of these two countries, with absolute rank mobility as high as 52.5 percent, with a transition probability of rags to riches movement of almost 20 percent, and that for intergenerational low income of less than 20 percent.\textsuperscript{35} Average parent income is relatively high, but closer to median income than in any of the other regions. These are communities of relatively more equality, but not of very low overall income.

But this high mobility region covers only about 9.5 million children, and is significantly less populous than the low mobility cluster, labelled as Cluster 4 in Figure 6 and Table 4, with a population six times higher. This region is most notably distinguished from the others by a very low rags to riches movement, and a high chance of intergenerational low income. The chances that a child born to bottom quintile parents will also be a bottom quintile adult average, at almost 35 percent, about twice the rate of the high mobility cluster. Parent average and median incomes are the lowest, yet median income is $18,000 less than the average, indicating a higher gap and more top end inequality than in at least Clusters 1 and 2. Escaping low income is a challenge not only in the southern United States, but also

\textsuperscript{35}Interestingly, this cluster also includes significant parts of the province of Newfoundland and Labrador, traditionally considered a low income–low mobility region. This finding, in part, highlights the fact that our measure of permanent income for children is less than perfect. Measurement error in child incomes is often not considered to imply an attenuation bias in the estimation of the intergenerational income elasticity, leading to a loss of efficiency without implying bias. But this is not the case with rank-based statistics, which require both the right and left-hand side variables in the single-variable least squares regression to be corrected for measurement error. See, among others Nybom and Stuhler (2015). Incomes in Newfoundland and Labrador, along with parts of Northeastern Alberta and Southern Saskatchewan were strongly influenced by the boom in oil and potash prices, which led to significant increases in median incomes during the first decade of the 2000s (Corak 2016a, Figure 15). We average child incomes over only a two-year period during the height of this boom, and as such do not take the period of lower prices after 2014 into account. In this sense we are likely overstating the permanent income of children in these regions. This would also apply to regions of the United States experiencing a commodity-boom at the same time. It is interesting to note that Corak (2016b) averages child incomes over a five year period, and Newfoundland and Labrador do not appear as a particularly high mobility region. But these results also reflect a clear lesson from theory: that relatively more equal labour markets characterized by a lower return to schooling relate to more generational mobility. The commodity boom should not be understood as the sole driver of mobility in this cluster, which also includes communities that have other advantages. As theory and Chetty, Hendren, Kline, and Saez (2014) stress, these relate to family structure, community resources associated with social capital, and an ethic of geographic mobility.
northern parts of Canada.

These two countries share regions of both high and low mobility, but are distinguished by the fact that on average—where the great bulk of the population lives—most Americans live in regions of less mobility. This is true in a relative sense, with parental income ranks being more strongly related to child ranks, and the changes of escaping low income being lower, in the United States. It is also true in an absolute sense, with children on average living in regions characterized by higher incomes and more inequality.

What drives the cross-national differences in mobility indicators is the fact that the regions of common experience carry a much lower weight in the overall population. In particular, the low mobility regions of the United States, for the most part the South, weigh more heavily in determining country-wide mobility than they do in Canada. This is made explicit in Table 5. This table presents the average values of each of the six mobility measures separately for each country. The first row of each panel offering the country-wide average, with the next four rows offering an series of alternative calculations in which successively one of the four clusters is dropped from the unweighted and weighted calculations presented in the two columns.36

6 Robustness and alternative specifications

We need to consider what to put here reflecting a conversation on some of the important issues in how the data have been developed. Here are some thoughts on the most important outstanding issues that have come to my mind.

I have some major concerns with the construction of the data, that I should have brought up a lot earlier, but I thought I had understood things and had the impression they were not a problem. Obviously I did not understand things correctly.

1. The linkage procedure for the two cohorts is very odd and questionable and leads me to suggest that our core analysis should be based just upon the 1980 cohort.

I am taking my understanding from Table A3 in the original conference version of the paper. This is peculiar. If I understand correctly, the 1980 cohort is linked from ages 16 to 20. This is all very good, and what I originally understood. But the linkage only begins at age 19 for the 1982 cohort, and continues to age 23. Is this correct? If it is I think we should back off from using the 1982 cohort. The whole point of doing the linkage over a series of years is to avoid a sample selection issue associated with the requirement that some children have to file a tax return while still living at home. This implies that there could 36

---

36 This table needs to be populated with the appropriate statistics. Again if Charles can send a spreadsheet in which each row is a region—a Census Division for Canada and a CZ for the United States—and if the columns offer the six statistics as well as addition information for child populations and an indicator for the cluster, I can make the calculations and put them in a spreadsheet that can be read in by R markdown and automatically populate the table. Or he can offer me a spreadsheet that has these calculations in the same row and column format of the table.
be a bias in that children from lower income families have a tendency to leave home earlier, and the linkage can’t be carried forward to far in a child’s age because this also captures the fact that we are including children who have a tendency to not leave home, which may also signal the role of unobservables in determining labour market outcomes. The balance we originally sought was to do the linkage starting as soon as we possibly can—at 16—but not going beyond the teen years. Panel B in this table clearly shows that this is violated, and that the 1982 cohort is captured at a much later stage in the life cycle. It is good that the bulk of the sample is captured at age 19, and a bit more at age 20.

I suggest that we redo the analysis using only the 1980 cohort, or possibly the 1980 cohort with only those 1982 cohort individuals who were linked at ages 19 or 20.

I don’t think we can rationalize different selection rules for the two cohorts, and report the results as they currently stand.

2. The exclusion of common-law spouses.

It is hard to defend the exclusion of common-law spouses from the Canadian data. I remember we talked this through, and perhaps I don’t appreciate or remember all of the considerations that went into this decision. But I honestly thought we decided to keep them in the Canadian data.

It is hard to rationalize their exclusion in the way this draft of the paper explains it. The draft argues that we exclude them in order to make our data comparable to the US data. I disagree with this because it is not an exclusion that is based on demographic differences, like say the age or gender choices, but rather it reflects an important behavioral outcome that is part of the intergenerational process. It also reflects different institutional and policy decisions, particularly the design of the tax system and income support policies.

I am also struck by the very large fraction of the Canadian sample that is excluded. I wonder what has happened to the representativeness of our sample in Quebec. Canadians living common law in Canada may still be in strong committed family relationships and the fact that this is not the case in the United States reflects substantive differences between the countries—the strength of the family, the income support to families, the design of eligibility for support, the workings of the labour market that put families under more stress when there is greater inequality—that will drive the mobility process. By excluding common law families we are not making the samples comparable, we are making them more different.

We need to talk this through. I have likely forgotten many aspects of the conversation we did have. But that said, I honestly thought that we had concluded that conversation with the decision that these individuals and their spouses would remain in the analytical sample.

3. The quintile transition matrix for Canada and for Canada placed in the United States.

There is something wrong here. The rows and columns of the Canadian transition matrix as depicted in what is now table 3 do not add up to 100. They are close, but not close in the
sense of a decimal point or so ... they are full percentage points off. Check the US transition matrix. It is spot on. So just how were the Canadian numbers calculated?

Also the the second panel of the table, there is a problem. I really don’t know how this was calculated. We say in the text that the adding up constraint does not apply. But I don’t understand this when it comes to the columns. Presumably all the children of a group of parents, in whatever quintile of the US income distribution in which they are placed, have to end up somewhere. So why don’t the columns add up to 100?

Maybe I’m missing something here, but it does mean in the least that this has to be explained more clearly. A lot rides on these calculations so we really need to double check that we understand what we are intending to do, and that we are doing it appropriately and accurately. Again, maybe we are. But I don’t fully understand.

7 Conclusion

to be determined and written.

8 Appendix

to be determined and written.

9 References


Corak, Miles, and Andrew Heisz. 1998. “How to Get Ahead in Life: Some Correlates of Intergenerational Income Mobility in Canada.” In *Labour Markets, Social Institutions, and


Table 5: Mobility measures for selected Clusters, weighted and unweighted comparisons

<table>
<thead>
<tr>
<th></th>
<th>Unweighted Canada</th>
<th>Unweighted United States</th>
<th>Weighted Canada</th>
<th>Weighted United States</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Absolute rank mobility, ( a )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2. Relative rank mobility, ( b )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3. Rags to riches mobility, ( P_{15} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4. Intergenerational low income, ( P_{11} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5. Average parent income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. Median parent income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Excluding Cluster 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Regions are grouped according to the results of K-means clustering with four clusters.