



Canadian Labour Market and Skills Researcher Network

Working Paper No. 98

Ageing and Literacy Skills: Evidence from Canada, Norway and the United States

David A. Green
University of British Columbia

W. Craig Riddell
University of British Columbia

April 2012

CLSRN is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) under its Strategic Knowledge Clusters Program. Research activities of CLSRN are carried out with support of Human Resources and Skills Development Canada (HRSDC). All opinions are those of the authors and do not reflect the views of HRSDC or the SSHRC.

**Ageing and Literacy Skills: Evidence from Canada, Norway and the
United States***

**David A. Green
Department of Economics
University of British Columbia**

**W. Craig Riddell
Department of Economics
University of British Columbia**

* We thank Lori Timmins and Jinwen Xu for excellent research assistance, Ian Walker and participants in the Ageing and Productivity Workshop sponsored by the WDA Forum for comments on an earlier version and the SSHRC for research support.

Abstract

We study the relationship between age and literacy skills in Canada, Norway and the U.S. – countries that represent a wide range of literacy outcomes -- using data from the 1994 and 2003 International Adult Literacy Surveys. In cross-sectional data there is a weak negative partial relationship between literacy skills and age. However, this relationship could reflect some combination of age and cohort effects. In order to identify age effects, we use the 1994 and 2003 surveys to create synthetic cohorts. Our analysis shows that the modest negative slope of the literacy-age profile in cross-sectional data arises from offsetting ageing and cohort effects. Individuals from a given birth cohort lose literacy skills after they leave school at a rate greater than indicated by cross-sectional estimates. At the same time, more recent birth cohorts have lower levels of literacy. These results suggest a pervasive tendency for literacy skills to decline over time and that these countries are doing a poorer job of educating successive generations.

All three countries show similar patterns of skill loss with age, as well as declining literacy across successive cohorts. The countries differ, however, in the part of the skill distribution where falling skills are most evident. In Canada the cross-cohort declines are especially large at the top of the skill distribution. In Norway declining skills across cohorts are more prevalent at the bottom of the distribution. In the U.S. the decline in literacy skills over time is most pronounced in the middle of the distribution.

Keywords: Human capital, Cognitive skills, Literacy, Ageing

JEL codes: I20, J14, J24

1. Introduction

Adult literacy skills are of both fundamental and instrumental importance. Sen (1999) argues that we should aim for a society in which every person has the capability to pursue any reasonable version of what they perceive to be good. To do so requires both access to at least minimal levels of resources and possession of characteristics that Sen calls “functionings”. One of the most important of these functionings is literacy. Without literacy, individuals cannot take a full and equal role in social and political discourse: they become less than equal members of society without the basic tools required to pursue their goals. Thus, in any attempt to build a better society, the distribution and generation of literacy is of fundamental importance. Literacy is also potentially important for instrumental reasons. Individuals with higher levels of literacy enjoy better employment opportunities and command higher earnings, leading to a higher level of well being (Green and Riddell, 2003; Statistics Canada and OECD, 2005). From a societal point of view, a more literate workforce may be better positioned to adjust to change and to adopt new technologies. Thus, improving literacy for individuals may have spill-over effects that benefit the economy as a whole.

In this paper, we study the relationship between age and basic literacy skills using data from the 1994 International Adult Literacy Survey (IALS) and the 2003 International Adult Literacy and Life Skills Survey (IALSS). For reasons discussed below, most of our analysis focuses on the Canadian data, but we also carry out a similar analysis for the U.S. and Norway, two countries that also participated in the 1994-1998 IALS and 2003 IALSS rounds of international comparative data collection on basic cognitive skills.

The IALS and IALSS surveys are unique in providing measures of basic literacy and numeracy skills for a representative sample of the adult population. In IALS 1994 skills were assessed in three domains: prose literacy, document literacy and quantitative literacy, while in IALSS 2003 a fourth skill – problem solving – was also assessed. These surveys combine methods of educational testing with household survey techniques, and also provide detailed individual and demographic information on respondents. A key objective is to assess skills used in daily activities – at work, in the home, and in the community. In other words, these are basic cognitive skills used in daily life.

In previous research using the 1994 IALS data we found that there is a surprisingly weak relationship between literacy skills and age, after controlling for other influences (Green and

Riddell, 2003; Ferrer, Green and Riddell, 2006). In this paper we investigate the relationship between literacy skills and age more thoroughly using the much larger sample provided by the 2003 IALSS data. Again, we find that the partial relationship between literacy skills and age is quantitatively modest and, if anything, negative in direction. The absence of a positive relationship between basic cognitive skills and age is perhaps surprising as individual earnings typically increase with age, albeit at a decreasing rate. The virtually universal finding of positively sloped age-earnings profiles is generally attributed to accumulation of human capital with work experience over the life cycle.¹ The relationship between literacy skills and age appears to not follow the pattern displayed by other forms of human capital.

However, the relationship between skills and age in cross-sectional data could reflect some combination of age and cohort effects. A 35-year old in 2003 may differ from a 25-year old in 2003 both because she is older and comes from an earlier cohort that experienced schooling and work experience in a different time period. In order to distinguish between age and cohort effects, we use the 1994 IALS and 2003 IALSS to create synthetic cohorts. For example, 26-35 year olds surveyed in 1994 and 35-44 year olds surveyed in 2003 are both representative samples of the 1959-68 birth-cohort. The analysis focuses on prose and document literacy, which are measured in a comparable way in the two surveys. We study the following 1994 age cohorts (2003 in parentheses): 26-34 (35-44); 36-45 (45-54); 46-55 (55-64); 56-65 (65-74); 65+ (74+). We also study cohorts by education.

Our investigations of literacy using 2003 IALSS data yield several noteworthy results. Literacy increases strongly (though at a decreasing rate) with years of schooling. Parental education levels also have a positive association with literacy, and parental immigrant status also plays a role. Perhaps most interestingly, we find a relatively weak negative relation between age and literacy, mirroring the finding of our earlier work with the 1994 IALS data (Green and Riddell, 2003). At first glance, these results appear to imply that individuals acquire their literacy through formal schooling and through the efforts of their parents but that their literacy levels do not develop further (and, indeed, gradually decline) upon leaving school. However, using a combination of the 1994 IALS and 2003 IALSS data, we show that the small negative slope of the profile of literacy relative to age in the cross-sectional datasets actually arises from a

¹ Other explanations include internal labour markets and incentive-based pay structures in which wages rise more rapidly with seniority than does worker productivity.

combination of offsetting ageing and cohort effects. In particular, individuals from a given birth cohort lose literacy skills in the years after they leave school at a rate that is typically greater than is indicated by cross-sectional estimates. There is less evidence of decline at the bottom of the skill distribution (where initial skill levels are low) but strong negative age effects at the top. At the same time, we find evidence that more recent birth cohorts have lower levels of literacy. This is particularly true for more highly educated individuals. Thus, a 35 year old in IALSS 2003 has approximately the same average literacy score as a 25 year old in the same survey not because that 25 year old should expect to be at the same literacy level in 10 years but because the 35 year old started from a higher literacy level at age 25 (i.e., comes from a more literate cohort) but lost some of their initial literacy skills during the time since they left school. These results suggest, on the one hand, a tendency for literacy skills to decline over time and on the other that we are doing a poorer and poorer job of educating successive generations.

The strong relationship between formal schooling and literacy is more evident when we use the sample that pools the 1994 and 2003 data and controls for cohort. The correlation of differences in literacy across cohorts with differences in education results in under-estimation in cross-sectional data of the impact of schooling on literacy.

We complement our analysis of Canadian data with that for Norway and the United States. These three countries are of interest for several reasons, including the fact that Canada generally falls between Norway and the U.S. both in terms of literacy levels and the extent of inequality in their literacy distributions. All three countries show the same pattern of literacy loss with age. Thus, whatever Norway is doing better it seems not to have to do with maintaining literacy levels after leaving school. Furthermore, both the U.S. and Norway show much the same cross-cohort pattern as Canada. The countries differ, however, in the part of the skill distribution where falling skills are most evident. In Canada the cross-cohort declines are especially large among those at the top of the skill distribution. The Norwegian data also shows declining skills across cohorts, but these are more prevalent at the bottom rather than the top of the distribution. In the U.S. the decline in literacy skills over time is most pronounced in the middle of the distribution. Thus, in all three countries there is evidence of declining literacy skills with successive generations, although there are differences in the extent of these declines among different segments of the population.

The paper is organized as follows. The next section describes our data and reports

summary statistics. Section three analyses the generation of literacy skills using the cross-sectional 2003 Canadian IALSS data. In the fourth section we employ “artificial cohorts” analysis to estimate separate age and cohort effects for Canada using the pooled 1994 and 2003 data. Section five reports a similar artificial cohorts analysis for the US and Norway. The final section concludes.

2. Data

Our data comes from the International Adult Literacy and Skills Survey (IALSS03), a fascinating survey carried out in several countries in 2003.² We also make use of the International Adult Literacy Survey (IALS94), an earlier survey of literacy skills carried out, for Canada, in 1994. The IALSS03 includes standard questions on demographics, labour force status and earnings, but it also attempts to measure literacy and related cognitive skills in four broad areas: Prose Literacy, Document Literacy, Numeracy, and Problem Solving.³ Perhaps of most importance for our discussion, the IALSS03 and the IALS94 did not attempt to just measure abilities in math and reading but tried to assess capabilities in applying skills to situations found in everyday life. Thus, the Prose questions in the surveys assess skills ranging from items such as identifying recommended dosages of aspirin from the instructions on an aspirin bottle to using “an announcement from a personnel department to answer a question that uses different phrasing from that used in the text.” The Document questions, which are intended to assess capabilities to locate and use information in various forms, range from identifying percentages in categories in a pictorial graph to assessing an average price by combining several pieces of information. The Numeracy component ranges from simple addition of pieces of information on an order form to calculating the percentage of calories coming from fat in a Big Mac based on a nutritional table. It is important for the work that follows that the Numeracy component changed substantially between the 1994 and 2003 surveys in response, in part, to concerns that the way the questions were stated in 1994 meant that math skills could not be separated from reading skills. In contrast, the Document and Prose tests have substantial overlap in the two survey years, with approximately 45% of the questions being identical across years. Statistics Canada also renormalized test results from the remaining 55% of the questions in 2003 so that the overall

² The other countries participating in this first round of the IALSS03 were Bermuda, Italy, Mexico, Norway, Switzerland and the U.S. The earlier IALS survey was carried out in over 20 countries during the period 1994 to 1998.

³ The latter is not included in the earlier IALS94.

average test scores from 2003 bore the same relationship to the overall average in 1994 as do the averages on the questions that are identical between the two years. Based on this, in the analysis that follows, we treat the Prose and Document test scores as perfectly comparable between the two survey years.

One defining feature of this data is the strong correlation among the various literacy scores. The correlation between the Prose and Document scores is particularly high (0.96 in the IALSS03), while the correlation between Document and Numeracy is .92, and the correlation between the Document and Problem Solving scores is .92. The Document and Prose scores are the ones that are common across IALS94 and IALSS03, so either could be used for the artificial cohorts analysis. Because of the high correlation between the two scores, we focus on the average of Document and Prose literacy. A principal components analysis indicated essentially one principal component placing equal weight on the individual scores. Thus the simple average score captures most of the variation in the data. Averaging the scores also reduces measurement error.

The IALS94 sample contains observations on 5660 individuals while IALSS03 is substantially larger at 23038 individuals. Both surveys cover all adults rather than being limited to those of working force age. Our goal is to focus on literacy generation in the Canadian economy and, as a result, we exclude from both samples anyone born outside of Canada, whose skills were not influenced by the Canadian educational system. We also drop observations on aboriginals, whose schooling system and skills are very distinct from those of other Canadians. The surveys cover individuals over age 16 but we restrict the analysis to those over age 25 in order to focus on those who have completed schooling. The result is samples of size 3,964 for the IALS94 and 14,666 for the IALSS03, which form the basis of our analysis of the distribution of literacy skills in Canadian society. We include both males and females throughout, dividing the analysis on gender lines in some places. Finally, we use the sample weights provided with the data in all tables and estimation; thus all summary statistics and regression estimates are representative of the adult Canadian population age 26 and above.

Table 1 reports summary statistics for our sample drawn from the 2003 Canadian IALSS data. The literacy test outcomes are scaled to fall between 0 and 500, with our average of the Prose and Document scores having a mean of 270 and a wide range. Females are slightly over-represented in the sample relative to the population but we control for gender in all of our

estimates. The average years of completed schooling (12.5) is typical of what is found in other Canadian surveys. Comparing this number to the fact that almost 60% of respondents report that their parents were high school drop-outs reflects strong progress in education across generations. What we call the Ability Proxies (the variables containing recollections of perceptions of school) exhibit real variability, implying they may be useful as covariates.

For the analysis of ageing that follows it is useful to provide some background information on the distribution of literacy skills in the population. Figure 1 shows a kernel density plot for the average literacy score for our sample. The density is negatively skewed with a mean of 270, a median of 285 and a standard deviation of 54. The fact that the distribution is negatively skewed when ability is typically thought of as distributed according to a positively skewed distribution such as the log normal may reflect a lack of sufficiently difficult questions to differentiate among highly literate respondents. The minimum average score in the sample is 94, representing extremely low proficiency at the test tasks, while the highest is 433.⁴ Inequality in the distribution is reflected in an associated Gini coefficient value of .107 and a value for the log of the ratio of the 90th percentile to the 10th percentile of .504. To put this in perspective, the Gini for pre-tax and transfer family income from the 2001 Canadian Census is .438 and the log of the 90-10 ratio is 3.48 (Frenette et. al., 2009). Thus, the distribution of literacy is much less unequal than that of market income. This is not surprising since literacy skills are only one of the components playing a role in earnings generation.

⁴ Figure 1 is constructed using the pooled data from 1994 and 2003 and so does not exactly match the statistics in Table 1. We present the pooled data density in order to be comparable with the other figures in the paper, all of which examine changes over time between the two surveys. The plot is truncated at the 0.5th and 99.5th percentiles in order to cut out distracting long tails on either end. The plot uses Stata 9's default values for the kernel and smoothing parameters.

3. The Generation of Literacy Skills

We turn now to using IALSS03 to examine the sources of literacy. We use the log of the average literacy score as our dependent variable so our estimated coefficients can be interpreted as showing impacts in terms of percentage changes in literacy.

Before presenting the estimation results, we begin by setting out a brief, heuristic model of literacy generation. The model will help to put our estimates in context as well as providing guidance in thinking about identification issues. Consider a simple model in which individuals start out at birth endowed with two key characteristics: their ability and parental resources. By parental resources, we mean something quite broad, incorporating both parental income and parental willingness and ability to support their children's education and literacy acquisition. Pre-school children begin to acquire literacy based on these fundamental characteristics. Once they enter school, ability and parental resources interact with characteristics of the school such as teacher quality, class size and the attitudes and abilities of peers. New additions to literacy with each year of schooling are then functions of ability, parental resources, school characteristics and the literacy level at the beginning of the period. These influences may interact in complicated ways. These additions continue until the legal school leaving age. After that point until the end of high school, students make a decision each year on whether to continue in school. That decision will be a function of ability, parental resources and school characteristics, but it is also likely to be a function of literacy acquired to that point. The more literate a student is, the less onerous they are likely to find school and the more likely they are to choose to stay an extra year. Finally, after high school, whether an individual continues to go to school will be determined by a combination of their own decision to apply to continue and the decision of the college or university on whether to admit them. The latter decision will likely be a function of the student's literacy as reflected in her grades. Thus, schooling and literacy are co-determined with extra years of schooling leading to increased literacy but increased literacy also leading to more years of schooling, especially after the legal school leaving age. Indeed, once we account for expectations, the inter-relation between the two may be even tighter. Individuals who do not expect to continue with school past the legal minimum may rationally under-invest in acquiring literacy skills while they are in school.

Once individuals leave school, literacy acquisition is likely more difficult. Literacy skills may be acquired on the job if they are needed for carrying out tasks at work but otherwise further

acquisition would require active investment in non-work hours. Indeed, it seems quite possible that individuals could lose literacy skills after they leave formal schooling if those skills depreciate when they are not used.

We are interested in characterizing the components of literacy generation, especially in whether literacy declines or rises after leaving school and how this ageing process is related to individual characteristics. If literacy has a “use it or lose it” form then there may be a case for adopting policies that encourage literacy maintenance and “lifelong learning” activities. Also of interest is the relationship of literacy to parental characteristics and resources as well as the linkage between formal schooling and literacy since this is a main channel through which one could hope to influence literacy outcomes. Many of these parameters of interest reflect causal relationships that are difficult to establish definitively. We will make efforts to estimate the causal parameters where the data permit but some of what we will discuss is necessarily in the form of correlations rather than clear causal impacts.

The first column of Table 2 presents our simplest OLS regression in which the dependent variable is the log of individual literacy score and the independent variables are age, age squared, years of schooling, years of schooling squared, and a gender dummy. The coefficient on the gender dummy indicates that there are not economically or statistically significant differences in literacy between men and women conditional on age and education.⁵ All the other variables are statistically significantly different from zero at the 1% level but this does not mean their actual impacts are sizeable. The age and age-squared coefficients are highly statistically significant but together they imply that the impact of an extra year of age on literacy is actually -0.6% at age 30 and -1.0% at age 50. The one relationship that is economically substantial is that between literacy and formal schooling. One extra year of schooling, evaluated when the individual already has 12 years of education, increases literacy by 3.2%. This is very similar to what Green and Riddell (2003) calculated using the IALS94 data.

⁵ Females achieve higher scores on prose literacy than do males, and the reverse is true for document literacy. The small and insignificant coefficient in the regression for the average score reflects these offsetting coefficients in the individual score regressions.

The second column replaces the quadratic specification of Age with a less parametric specification using 10-year age categories. The quadratic specification actually yields a slightly better fit. The implications of the estimates are very similar. Literacy declines very gradually with age – for example by less than 3% for those 46-55 years of age relative to the omitted category (26-35 years of age). However, among individuals over age 65 literacy skills are substantially lower, controlling for gender and education.

As we discussed earlier, literacy and years of schooling are likely to be jointly determined. In that case, OLS estimates are likely to be biased. Although most attention focuses on the estimated impact of schooling on literacy, other coefficients (including those on age) could also be biased. We attempt to address this in two ways. First, biases may arise because of a correlation between literacy and schooling arising from unobserved variables that are correlated with both education and literacy. One important set of variables that is often not available consists of parental and family background characteristics. Here we fortunately are able to control for a variety of family background factors including educational attainment, occupation and immigrant status of the respondent's mother and father. Other potential unobserved factors include individual characteristics such as innate ability and motivation. If high ability people do not view it as particularly costly to either acquire literacy or go to school then we could observe a positive coefficient on schooling because years of schooling is proxying for ability rather than as a reflection of a causal impact of schooling on literacy. This problem can be addressed if we have a measure of ability since once we control for ability, any relationship between schooling and literacy cannot be due to an omitted ability term. Note, though, that many studies that try to control for ability (in, for example, earnings regressions) actually use scores on tests much like our literacy tests. What we would require is a test score from a very young age - before the process we are trying to study really begins. Since we don't have that, we instead try to proxy for ability using two variables that are plausibly related to it.

In the third column of Table 2 we add variables on parental education and parental immigrant status. Introducing these variables has virtually no impact on the gender coefficient, but it does have the expected consequence of reducing the coefficient

on years of schooling, albeit by a small amount (<5%). Including them also leads to an increase in the age coefficient. Given that the coefficient on the age-squared variable also becomes more negative, the net effect of age remains quite small. The parental education variables are jointly highly significantly different from zero but, perhaps surprising, the effect is found mainly at low levels of parental education. Having a parent (either mother or father) who is a high school drop out decreases average literacy by about 3%. However, parental education beyond high school graduation has relatively modest further impacts on literacy. Interestingly, not knowing a parent's education level (which is the case for approximately 9% of the sample) has a strong effect, being associated with approximately 5% lower literacy. While we included this variable in order to allow us keep the observations with missing information on parental education, it seems possible it represents something real. For example, children who do not know a parent's education likely did not have a close relationship with that parent. Thus, the estimated coefficient may reflect the extent to which literacy is generated through direct parental involvement. Finally, having a mother or father who is an immigrant has a mild association with literacy (increasing literacy by 1%-2%). We also tested specifications in which we included a set of parental occupation dummy variables but these were never jointly statistically significant.⁶ We also find that a dummy variable representing whether the individual's mother was working when the individual was 16 does not have a statistically significant effect. Overall, the results point to a surprisingly weak association between literacy and parental background. Only schooling seems to have a substantial impact on literacy generation.

In the fourth regression, we add the ability proxies. Both are based on the respondent's experiences while in secondary school. The first is a dummy variable equalling one if the person agreed or strongly agreed with the statement that they got good grades in math when they were in school and the second is a dummy variable equalling one if the respondent agreed or strongly agreed with the statement that teachers often went too fast and the person often got lost. Either of these could plausibly be seen as proxies for innate ability. Both of these variables enter significantly, with people who

⁶ In particular, a test of the hypothesis that the set of father's occupation dummy variables jointly had zero effects has an associated P-value of .13. The same test for mother's occupation has a P-value of .79.

claimed to have gotten good grades in math having 1.5% higher literacy and those who thought teachers went too fast having 2.5% lower literacy. However, including these variables has almost no impact on the age coefficients. Their inclusion does reduce the estimated impact of schooling on literacy, but the magnitude of the decline is modest (< 10%).

An alternative approach to identifying a causal effect is to adopt an instrumental variable strategy. We use the minimum school leaving age as an instrument for schooling.⁷ Changes in compulsory schooling laws have been shown to have significant effects on educational attainment, and have been a commonly used instrument for education in many studies (see, for example, Acemoglu and Angrist, 2000; Lochner and Moretti, 2004; Milligan, Moretti and Oreopoulos, 2004; Harmon and Walker, 2005; and Oreopoulos, 2006a, 2006b). In Canada, education falls under provincial jurisdiction and compulsory schooling laws vary across provinces as well as over time within provinces. Based on detailed information on compulsory schooling in Canadian provinces we created three indicator variables for the school leaving age: <15, 15 and 16 (see Table 1 for mean values in our sample). Identification of the causal impacts of education on literacy skills assumes that variations in the minimum school leaving age over time and across provinces induced changes in years of schooling that are unrelated to unobserved factors such as ability or motivation.

Although we have a suitable instrument for years of schooling, we were unable to obtain a separate instrument for schooling squared. A standard approach to instrumenting for a squared term would be to use higher order terms in the instrument as instruments for the squared term. However, since our instrument consists of a set of dummy variables, the higher order terms equal the dummy variable itself and we are left without a separate instrument for the squared term. We therefore restrict the IV analysis to those with 16 years of school or less, the sub-sample for which the partial relationship between log literacy and years of education is approximately linear. This is also the range of schooling within which compulsory schooling laws are likely to be binding.

The fifth column of Table 2 reports OLS estimates for this restricted sample. The

⁷ We compiled data on the school leaving age from provincial statutes and the date of proclamation or ascension from provincial regulations. Our data are similar to those compiled by Oreopoulos (2006a).

education coefficient is highly significant and indicates that literacy increases by approximately 3.4% with each year of schooling up to 16 years of schooling.⁸ This estimate is similar in size to that previously reported as the impact of an added year of education on literacy for an individual with 12 years of schooling based on the quadratic specification. The coefficients on the age-related variables are very similar to the comparable estimates for the full sample (see column 3). The first stage regression, in which years of schooling is the dependent variable, is reported in the Appendix. As expected, parental education is strongly positively related to years of schooling. Having parents who are immigrants also corresponds to having more years of schooling. The school leaving age variables are jointly highly statistically significant, with a first-stage F-statistic of almost 60.⁹ The estimated school leaving age coefficients indicate that changes in compulsory schooling exert sizeable impacts on educational attainment.

The results from two-stage least squares estimation using school leaving age as an instrument is given in column 6. Several differences between the OLS and IV estimates are evident. The coefficient on education is much larger (approximately double) in the IV specification, a result that has been found in many other studies based on an IV strategy (Card, 1999). The gender coefficient is now negative, significantly different from zero, but still small in size. Perhaps most important for the purposes of this study, the age coefficients estimated with IV are somewhat smaller in size than their OLS counterparts. Thus OLS estimation was not causing a misleading view regarding the small impact of age on literacy skills. The family background variables are also generally smaller in size than their OLS counterparts.

Two key findings follow from the estimates reported in table 2. First, if the assumptions underlying our instrument are correct, the estimates indicate that education has a strong causal effect on literacy and that schooling is the dominant determinant of literacy. The second conclusion is that our attempts to address potential bias due to the endogeneity of schooling do not alter the finding that literacy gradually declines with age,

⁸ Adding the years of schooling squared term to this specification results in a coefficient that is very small and not significantly different from zero.

⁹ The test statistic corresponding to the joint hypothesis that the province of residence dummies and their interactions with age are significantly different from zero is distributed as F(22,13821) and has an associated p-value of 0.009.

beginning as early as the mid- to late 20s, but at a very modest rate.

4. Age and Cohort Effects

One of the most striking results from the Table 2 regressions is that literacy skills decline (albeit at a slow rate) with age, beginning as early as the late 20s . Taken at face value, this suggests that individuals do not enhance these skills after they finish schooling. Essentially, literacy is acquired principally at school and then gradually declines thereafter. This finding is reinforced if we replace age and age squared by experience and experience squared in the regressions in Table 2.¹⁰ The coefficients on schooling, schooling squared and the female dummy are virtually identical to those presented in Table 2.

Interpreting the coefficient on age in a cross-sectional regression requires some care, however. As has been extensively studied in the literature on immigrant earnings, differentials between two age groups in a cross section could reflect a variety of possible combinations of true age and cohort (or generational) effects. Thus, while we are tempted to view the literacy level of 35 year olds in the IALS03 as a reflection of the literacy level the 25 year olds are likely to be at in 10 years time, we need to bear in mind that the 35 year olds come from an older generation and their observed literacy reflects a combination of any generational differential as well as any ageing effect. Only if there are no systematic differences across cohorts does the cross-sectional literacy-age profile reflect the true impact of ageing on literacy.

A more complete investigation of cohort and ageing effects requires the use either of true panel data or of at least two cross-sectional datasets constructed in such a way that we can follow “synthetic” cohorts through time. We make use of the IALS94 and the IALSS03 for this purpose. More specifically, in the public use version of the IALS94, we can observe a set of 10-year wide age groups for the respondents (i.e., ages 26-35, 36-45, 46-55, 56-65, and 65+). Since we have a continuous age variable in the IALSS03, we can construct age groups that correspond to the age people in these initial groups would be 9 years after IALS94 (i.e., 35-44, 45-54, 55-64, 65-74, and 74+). We refer to the people

¹⁰ Experience is defined according to the standard Mincer expression (i.e., experience = age - years of schooling - 6).

who were age 26-35 in 1994 and 35-44 in 2003 as “Cohort 1” and number the remaining cohorts in ascending order from there. Since both surveys provide representative samples of the adult population, it follows that each provides an unbiased estimate of the literacy distribution for the cohort at two different points in time and we can follow the progress of a given cohort over time. In part of our examination, we will further break the cohorts down by education. Linking across surveys would not provide unbiased estimates of the progress of cohort-education groups in this case if the educational composition of the groups being studied changes over time. Thus, we cannot say that the average literacy of 35-44 year old high school graduates in IALSS03 provides a consistent estimate of what happened to the 26-35 year old high school graduates we observe in IALS94 over the intervening 9 years if some of the initial high school graduates would be expected to have increased their education level. In that case, the average literacy in 2003 would reflect a combination of ageing effects and the systematic selection of people out of the high school graduates group. To avoid this problem, we continue to focus on individuals over age 25, after which point changes in formal education across our broad categories are rare. It is also worth noting that the oldest cohort (made up of those over age 65 in 1994) is special in that its composition is likely to change over time because of deaths. As a result, while we include it in our analysis, we do not place much weight on results related to that cohort in our conclusions.

We begin with plots of the densities of the average of Document and Prose literacy from the IALS94 and IALSS03 broken down by age, education and cohort groups. Thus, in Figure 2, we present the density plots for individuals who are age 26 to 35 in 1994 and individuals who are the same age in 2003. These densities correspond to the literacy for the youngest cohort we examine in 2003 and for the cohort just before them at the same age (observed in 1994). The density for the younger cohort (observed in 2003) has noticeably less spread, with both the left and right tails being thinner. In other words, the younger cohort experiences an improvement in literacy at the low end of the distribution but a worsening at the top end. This pattern is reflected in a 10th percentile of 228 and a 90th percentile of 358 in 1994 compared to values of 249 and 350, respectively, in 2003. The reductions in inequality in both tails nearly cancel out, as reflected in a median value of 300 in 1994 as compared to 302 in 2003. Thus, one might conclude

from a measure of central tendency that little has changed, whereas the tails reflect more substantial differences.

In Figure 3, we present the same density comparison for individuals aged 36 to 45. The outcome is much the same as in Figure 2 except that the relative improvement in the lower end of the distribution is not as great and neither is the relative decrease at the top. In contrast, Figure 4 shows that for the 46 to 55 year olds, the results in 2003 are better across the distribution. This is important, in part, because observing different relative changes between 1994 and 2003 for different groups suggests that we are witnessing something real rather than just a difference in the tests in the two years. If all plots for all groups showed improvements at the bottom and declines at the top between 1994 and 2003 then the simplest explanation for the patterns would be that the test changed in such a way that it generated better results on the easy questions that will constitute most of the scores at the bottom but worse results on the harder tests that will define the shape of the top of the distribution. Finally, Figure 5 also shows that the 56 to 65 age group made improvements over time, though the 1994 and 2003 densities are similar at the very top end.

We next examine the changes between the two years broken down by education group. Figure 6 plots the averaged literacy densities in 1994 and 2003 for people with less than a high school diploma. The improvement at the bottom end of the overall distribution is also evident at the bottom of this distribution and is quite large. The 10th percentile for this drop-out group rises from 139 in 1994 to 155 in 2003. At the other end, the differences are smaller but in the same direction as we have seen in the earlier figures: the 90th percentile declines from 300 to 292 between 1994 and 2003. Of course, the 90th percentile of their distribution is relatively low - being around the 70th percentile of the overall distribution. For the high school graduates, depicted in Figure 7, the 1994 and 2003 densities are very similar, with the 2003 density being shifted slightly to the left. Both the Non-university Post-Secondary and the University groups show, in Figures 8 and 9 respectively, declines in the upper part of their distributions, with the former group also showing some declines lower in the distribution as well. Both the Some Post-Secondary and the University groups register substantial declines in their 90th percentiles between the two years (from 361 to 338 for the Post Secondary and from 373 to 364 for

the University group). Thus, in terms of within education group movements, the improvements at the bottom of the overall distribution are really only clearly evident in the high school drop-outs distribution while the declines at the top are mainly observed in the distributions for the top two groups.

The changes observed for the overall distributions may, in part, also stem from changes in the educational composition of the population between the two years. The replacement of older, less educated cohorts with younger, more educated ones shifts the educational distribution. Thus, the proportion of the population with a university degree rises from 17% in 1994 to 21% in 2003. This improvement undoubtedly plays a role in the improvement in the lower tail of the overall distribution. It is difficult to see, however, how this change could cause the declines at the top of the distribution. If a person of the same underlying ability was a high school graduate in the oldest cohort but a university graduate in the newest cohort then if we were to replace the former with the latter, one would expect the overall distribution to either improve (if university education increases literacy) or stay the same (if it does nothing). Even the change at the top of the distribution for university graduates is unlikely to be related to an expansion of the university-educated pool. While the additional university graduates (those who would not have attended university in an earlier generation) are likely to be of lower ability than those who would have been university graduates in both periods, this factor would affect the lower not the upper part of the university-educated distribution.

We now turn to the analysis of ageing effects. Figures 10 through 13 present density plots for each year for specific birth cohorts, as defined earlier. Note that this contrasts with Figures 2 through 5 that examine people of the same age in the two surveys. Figures 2-5 make comparisons across cohorts (at the same age) while the Figure 10–13 comparisons follow individual cohorts through time. Figure 10 shows the plots for the cohort that was age 26 to 35 in 1994 and 35 to 44 in 2003. The plot indicates little difference at the low end of the distribution for the two years but a clear worsening at the top end. In fact, the 10th percentile of the literacy distribution increases from 228 to 234 for this cohort while the 90th percentile falls from 358 to 342. The cohort that was age 36 to 45 in 1994, plotted in Figure 11, shows a very similar pattern. In contrast, the densities for the cohort that was 46 to 55 show, if anything, an improvement from 1994 to 2003

(Figure 12). The 10th percentile for that cohort increased from 193 to 204 and the 90th percentile decreased slightly from 333 to 331 between the survey years. Again, this is important for considerations of the comparability of data since it shows that this cohort answers the full range of questions essentially as well (or slightly better) in 2003 as in 1994. The differences in patterns across cohorts would be very hard to rationalize simply as being due to more difficult top level questions being asked in 2003 than 1994. A more likely explanation of the patterns is that literacy skills deteriorate after leaving school but that the rate of deterioration declines with age. For the oldest cohort (who are 56 to 65 in 1994 and 65 to 74 in 2003) there is a marked leftward shift in the entire distribution.

Finally, Figures 14 through 16 show the effects of ageing for specific education/cohort groups. Here, we define wider cohorts in order to ensure enough observations. Figure 14 shows the plots for those with less than a high school education. There is clear evidence of deterioration over time in literacy skills at all levels above about the 20th percentile. For the high school graduates, in Figure 15, the deterioration shows up at a similar place. The university graduates (shown in Figure 16) show mixed changes up to about the median but, again, clear deterioration above that point. Thus, there is evidence of deterioration of higher-level literacy skills for all education groups.

To pursue these ageing and cohort effects further, we estimate a series of regressions in which we pool the IALS94 and the IALSS03 data and add cohort dummies to our standard specification. The public use version of IALS94 does not contain a continuous age variable so we adjust our specification by using age categories instead of age and age squared variables. This also facilitates comparisons with US and Norway, countries for which the data on age are only available in categories. We report the results from this specification using simple OLS estimation in column 1 of Table 3.. As in the earlier specification, the schooling variables enter strongly and significantly. The schooling effect is substantially larger than we witnessed in earlier OLS estimation based on cross-section data. This is evident once we include the cohort dummy variables. Essentially, the schooling variable in the earlier specifications was partly picking up the cohort effects shown here: older cohorts have both less schooling and higher literacy, leading to an under-estimate of the true impact of schooling. The parental education variables exhibit the same patterns as in earlier estimation, with low as well as

unknown/not reported parental education having a negative impact but the remaining variables having small and insignificant coefficients.¹¹

The age and cohort group effects in the pooled OLS results show interesting patterns mirroring those seen in the figures. Recall from Table 2 that when we estimate our standard specification with dummies for age categories, we find a cross-sectional age profile with a small negative slope for much of the age range. In particular, the first two dummy variables (corresponding to the 36-45 and 46-55 age groups) have coefficients that indicate declines in literacy skills of 1.9% and 2.8% respectively relative to the base category (the 26-35 age group). The older 56-65 year age group shows a larger negative effect – a decline of 5.8%. These estimates fit closely with the patterns using age and age squared variables. Once we include the cohort dummy variables, however, the age effects indicate a steeper downward sloping profile, with the 46-55 year olds having 3.9% lower average literacy and the 56-65 year olds having 7.2% lower average literacy than the base group. At the same time, all the cohort dummies apart from the last one have positive effects that increase with the cohort (although that for the cohort that was 56-65 in 1994 is not statistically significant). The cohort that was 46-55 in 1994, for example, has average literacy levels that are 2.8% higher than those for the base cohort, which was 26-35 in 1994. The very oldest cohort has literacy levels that are approximately the same as the youngest cohort. Overall, the implication from these results is that the small negative slope of the literacy-age profile (at least up to age 65) arises from a combination of literacy that declines with age at a greater rate than is suggested by the cross-sectional estimates and lower average literacy for more recent cohorts.

We investigate these phenomena further using quantile regressions in the remaining columns of Table 3. Several points stand out from these regressions. First, the impact of schooling declines across quantiles: years of school have a greater impact on literacy in lower quantiles than upper quantiles. In examining this evidence, it is important to take account of the second order terms since the schooling effect declines faster with additional years of school at the lowest quantile while it is closer to linear at the top. This means that schooling has a much larger effect in shifting the bottom than the

¹¹ Parental immigrant status is not observed in the US and Norway data. We therefore exclude these variables in the Canadian pooled specification for comparability purposes. Excluding parental immigrant status in the estimation in Table 3 has very little impact on the results.

top of the distribution at low years of schooling but about equal effects at the top and bottom of the distribution around 16 years of schooling. Second, the declining literacy with age effect is larger at the top than at the bottom of the distribution. Thus, as in the figures, the bottom of the distribution for various age groups is relatively similar but older age groups have much lower 90th percentiles. According to the table, for example, the 56-65 year old group has a 90th percentile that is 11% lower than the 90th percentile for the base (26-35 year old) group, controlling for education and parental education. Third, while the cohort effects are evident across the distribution, they are much stronger at the top end. Thus, the cohort who was 56 to 65 in 1994 (i.e., those born between 1938 and 1947) has a 90th percentile that is 9% higher than that for the base group (who were born between 1968 and 1977). The intriguing question is whether this reflects a decline in effectiveness of the school system. The fact that the figures show some improvements in literacy across cohorts at the bottom end of the distribution while the regression results (with the exception of cohort 4) show no significant change is likely due to the fact that the latter control for education and parental education; newer cohorts have more education, as do their parents. Taken together these results may suggest that schools are doing a poorer job of imparting literacy at any given level but that there have been real benefits to the fact that successive generations have attained greater schooling.

In Table 4, we present quantile regressions for those with a high school or less education and those with a university education in order to take a closer look at education-related effects. For the high school or less educated group, the ageing effects are virtually non-existent at the lowest quantiles, which perhaps reflects that there is little loss with age of very basic literacy. Also, there is evidence that the lowest quantiles for this education group have actually increased across cohorts. At the top end of the distribution, though, the results are much like the overall results in Table 3: strong declines with age combined with declines across recent cohorts. Thus, it seems possible that changes to public and high school education have been equalizing across cohorts, with improved literacy for those at the bottom being offset by not providing as good literacy training for those at the top. For the university educated, the declines with age are evident across the whole distribution. There is also evidence of declines across cohorts but these are not statistically well defined. This could be simply a sample size problem or

it could suggest that the real difficulties in literacy across generations are associated with other post-secondary education rather than with university education.

The cross-cohort patterns of declining outcomes at the top of the test score distribution and improvements at the bottom fit broadly with results from the PISA evaluations – though over a different time frame. PISA is an OECD initiative in which 15 year olds in a range of countries are tested on their abilities in reading, math and science. For Canada, between 2000 and 2009, the 10th percentile of the reading score distribution declined by a statistically insignificant 1% while the 90th percentile declined by 2.3% (significant at the 5% level) (OECD, 2010)). Thus, even over this short time frame, there is some evidence of declines in the upper tail of the distribution. In addition, the average reading scores for children whose parents were high school drop-outs was essentially unchanged between 2000 and 2009 while the average score for children whose parents had a BA or higher education declined by approximately 5%. Given correlations in ability and outcomes across generations, this again fits with declines in literacy at the top end of the distribution over extended periods of time. The implication is that Canada’s socio-economic gradient in student outcomes is declining over time but for the wrong reason: because of declines at the top end rather than gains at the bottom.

5. Ageing and Skills in Norway and the United States

Do the ageing patterns observed in Canada also hold in other countries, or are they unique to Canadian society? We use Norway and the U.S. to address this question.¹² We chose these countries principally because in each case we have access to consistent data from both rounds of the IALS, allowing us to make the same kind of age and cohort comparisons as carried out for Canada. However, there are distinct advantages associated with these choices. Norway provides an interesting benchmark because the Nordic countries tend to perform well internationally in literacy assessments –providing something like the “gold standard” in terms of what is attainable.¹³ This is the case both for the high mean level of literacy in the adult population, and for the low degree of inequality in the distribution of literacy skills. The U.S. is interesting in part because of

¹² It is worth re-iterating that the IALSS surveys were explicitly designed to allow direct comparability of literacy levels across countries and over time.

¹³ See Statistics Canada and OECD (2005) for international comparisons based on IALSS 2003. Similar cross-country patterns are evident in the 1994-1998 IALS data.

its large and technologically advanced economy but also because it typically has lower average skills and greater variability in literacy outcomes than Canada. Canada thus sits between Norway, which has both superior literacy levels and less inequality in literacy, and the U.S., which has generally lower literacy levels and higher literacy inequality. As a group, the three countries thus provide a wide range of cognitive skill outcomes. The Norwegian and U.S. samples are much smaller than those for Canada with 2522 in our usable 2003 sample for Norway and 1486 in our 2003 sample for the U.S. Thus the precision of the estimates for these two countries is likely to be much lower than was the case for Canada. To save on space we focus on the regression estimates for the two countries. It is worth noting, though, that in our samples the distribution of the average literacy scores for Norway has both a higher mean (297) and lower standard deviation (45) than the Canadian values (mean = 280, standard deviation = 48), while the US distribution has a mean of 274 and a standard deviation of 48. The Scandinavian model generates higher literacy and less inequality in literacy than the Canadian and American systems.

Table 5 reports the estimates (OLS plus quantile regressions) for the average of document and prose literacy in Norway. The results are similar to those for Canada. In particular, we again see a strongly declining pattern with age across the distribution. Indeed, the rate at which skills fall with age is significantly greater in Norway than in Canada. The main difference is that while this decline is stronger at higher quantiles in Canada, the opposite is true in Norway. The differences across quantiles are also more pronounced in Norway. Similarly, while both countries show declines across cohorts on average, the effects are stronger at the lower conditional quantiles for Norway but for the middle and higher conditional quantiles for Canada.

Table 6 presents the same type of analysis for the U.S. samples. The regression estimates show similar patterns to those seen for the Canadian and Norwegian data. In particular, cohort differences exhibit behaviour as in Canada and Norway, with the most recent cohorts having the lowest skill levels. U.S. cohort differences are most prominent in the middle of the literacy skills distribution. In addition, the estimated age coefficients show strong declines with age that are, again, most substantial in the middle of the skills distribution.

The strong influence of education on literacy is evident in all three countries, having similar magnitudes in Norway and the U.S. and larger estimated impacts in the Canadian data. In all three countries formal schooling exerts the greatest influence at the bottom of the skills distribution. Parental education is also associated with the respondent's literacy skills in the U.S. and Canada, with the largest estimated impacts being those in the U.S. However, in Norway the evidence suggests that father's education has no influence while that of the mother is modest relative to North America.

Overall, comparisons to both Norway and the U.S. indicate that the ageing pattern we identified for Canada is also present in countries with quite different institutions. In all three countries, literacy declines strongly with age for any given cohort. The countries are also similar in the importance of formal schooling for generating literacy and in the fact that schooling plays a particularly strong role in raising minimum literacy levels. Remarkably, these three countries are characterized by similar cross-cohort patterns of movements in average literacy. However, Norway differs from Canada and the US in that the cohort movements happen mostly at low quantiles in Norway but are stronger at higher quantiles in the US and Canada. The fact that the ageing patterns are similar across the three economies suggests that the impact of post-schooling institutions is similar in all three. That is, none of the countries has established a superior system in terms of maintaining post-schooling literacy levels. Cohort effects, on the other hand, are related to "permanent" differences associated with people who were born and went through schooling, work and other life experiences at different times. Differences in cohort patterns are thus reflections of institutions that have persistent effects on literacy, with differences in the efficacy of formal schooling being the most likely candidate. Under this interpretation, Norway has been successful relative to the U.S. and Canada in generating both higher overall literacy levels and less literacy inequality. However, Norway has not been able to maintain high levels of literacy generation over successive cohorts, especially at the bottom of the conditional skills distribution. At the same time, despite lower literacy levels overall, Canada and the U.S. have also not been able to maintain the skills of successive generations, especially in the middle and top of the skills distribution.

6. Conclusion

The IALS and IALSS surveys are unique in providing measures of basic literacy skills

for representative samples of the adult population in participating countries. In this paper, we use these data to investigate the relationship between literacy skills and ageing in Canada, Norway and the United States. Our cross-sectional analysis of the factors that influence literacy skills uses the 2003 Canadian IALSS survey that has a large sample size and rich information on respondents. This analysis concludes that formal education is the primary driver of adult literacy skills. The characteristics of the respondents' parents – such as education and immigrant status – have significant effects on the respondents' education, but a direct impact on literacy that is relatively modest in size. Perhaps surprising is the finding that literacy skills do not increase with age even over relatively early phases of the life cycle. Rather, there is a weak negative relationship between literacy skills and age beginning in the mid- to late 20s. Taken at face value, these results suggest that literacy skills are primarily determined by formal schooling, and then slowly but gradually decline.

To investigate these issues further, we take advantage of the fact that the 1994 and 2003 surveys provide representative samples of the adult population at two points in time in these three countries, which allows us to separately identify cohort and ageing effects. Doing so indicates that prose and document literacy skills decline with age after completing formal schooling in all three countries. In Canada and the U.S. there is less evidence of a decline with age at the bottom of the literacy skill distribution, but strong negative age effects in the middle and at the top of the distribution. Norway also exhibits strongly declining literacy skills with age, but these are more evident at the bottom of the distribution.

We also find that in all three countries successive birth cohorts have had poorer literacy outcomes. In Canada and the U.S. these occur in the middle and at the top of the skill distribution. Norway also exhibits declining literacy across cohorts, but these declines are more evident at the bottom of the distribution. The weak negative relationship between literacy skills and age found using cross-sectional data in these three countries appears to result from offsetting age and cohort effects. Once we control for cohort effects, the decline in literacy with age is much more pronounced. The results for these three countries – that represent a wide range of literacy outcomes – suggest that declining literacy with age is a pervasive phenomenon.

References

- Acemoglu, Daron, Joshua Angrist. 2000. "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws." *NBER Macroeconomics Annual 2000*: 9-59.
- Card, D. "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. IIIa. Amsterdam: Elsevier, 1999.
- Ferrer, A., D. Green and W.C. Riddell. "The effect of literacy on immigrant earnings" *Journal of Human Resources* 41 (Spring 2006) 380-410.
- Frenette, M., D.A. Green and K. Milligan. "The tale of the tails: Revisiting recent trends in Canadian after-tax income inequality using Census data," *Canadian Journal of Economics*, Vol. 40, No. 3, August 2007, pp. 734-764.
- Green, D.A. and Riddell, W.C. "Literacy and Earnings: An Investigation of the Interaction of Cognitive and Unobserved Skills in Earnings Generation," *Labour Economics* 10 (April 2003) 165-84.
- Harmon, Colm and Ian Walker. "Estimates of the Economic Return to Schooling for the United Kingdom" *American Economic Review* 85 (December 1995) 1278-1286.
- Lochner, Lance, and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Reports." *American Economic Review* 94:155-189.
- Milligan, Kevin, Enrico Moretti and Philip Oreopoulos. 2004. "Does education improve citizenship: Evidence from the U.S. and U.K." *Journal of Public Economics* 88 (August): 1667-1695.
- OECD (2010). "PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000 (Volume V) <http://dx.doi.org/10.1787/978264091580-en>.
- Oreopoulos, Philip. 2006a. "The compelling effects of compulsory schooling: Evidence From Canada" *Canadian Journal of Economics* 39 (February) 22-52.
- Oreopoulos, Philip. 2006b. "Average Treatment Effects of Education when Compulsory School Laws Really Matter." *American Economic Review* 96 (1): 152-175.
- Sen, A. (1999). *Development as Freedom*. New York: Anchor Books.
- Statistics Canada and Organisation for Economic Co-operation and Development. *Learning a Living*. Ottawa and Paris, 2005.

Table 1: Summary Statistics- Canada

Variables	Mean	Std. Dev.	Min	Max
DEPENDENT VARIABLES				
Prose Literacy Score	271.37	53.26	83.80	435.80
Document Literacy Score	267.69	55.28	82.80	430.20
Average of Prose and Document	269.53	53.79	93.60	433.00
EXPLANATORY VARIABLES				
Female	0.550	0.498	0	1
Years of Schooling	12.505	3.717	0	33
Age of Respondent				
Age	50.38	14.49	26	100
Age: 26-35	0.163	0.369	0	1
Age: 36-45	0.247	0.431	0	1
Age: 46-55	0.253	0.435	0	1
Age: 56-65	0.170	0.376	0	1
Age: 66+	0.167	0.373	0	1
Mother's Education				
Less than High School	0.561	0.496	0	1
High School	0.193	0.395	0	1
Some Post Secondary	0.108	0.311	0	1
University Degree	0.048	0.213	0	1
None Reported	0.089	0.285	0	1
Father's Education				
Less than High School	0.606	0.489	0	1
High School	0.135	0.342	0	1
Some Post Secondary	0.088	0.284	0	1
University Degree	0.071	0.257	0	1
None Reported	0.099	0.299	0	1
Parental Immigration				
Immigrant Mother	0.097	0.297	0	1
Immigrant Father	0.114	0.317	0	1
Ability Proxies				
Good Math Grades	0.732	0.443	0	1
Teachers Too Fast	0.281	0.449	0	1
School Leaving Age				
<15 years	0.180	0.384	0	1
15 years	0.338	0.473	0	1
16 years	0.482	0.500	0	1

Table 2: Average Prose and Document Literacy Regressions- Canada

VARIABLE	OLS 1	OLS 2	OLS 3	OLS 4	OLS 5	IV 1
Female	-0.0015 [0.0027]	-0.0019 [0.0027]	-0.0003 [0.0026]	0.0022 [0.0026]	0.0003 [0.0030]	-0.0112*** [0.0041]
Years of Schooling	0.0610*** [0.0017]	0.0619*** [0.0017]	0.0579*** [0.0017]	0.0532*** [0.0020]	0.0336*** [0.0006]	0.0679*** [0.0067]
Years of Schooling Squared	-0.0012*** [0.0001]	-0.0012*** [0.0001]	-0.0012*** [0.0001]	-0.0011*** [0.0001]		
Age of Respondent						
Age	0.0067*** [0.0006]		0.0094*** [0.0006]	0.0095*** [0.0006]	0.0098*** [0.0007]	0.0079*** [0.0008]
Age Squared	-0.0099*** [0.0005]		-0.0120*** [0.0005]	-0.0122*** [0.0005]	-0.0125*** [0.0006]	-0.0087*** [0.0010]
Age: 36-45		-0.0193*** [0.0039]				
Age: 46-55		-0.0277*** [0.0041]				
Age: 56-65		-0.0582*** [0.0047]				
Age: 66+		-0.1770*** [0.0048]				
Mother's Education						
Less than High School			-0.0328*** [0.0036]	-0.0347*** [0.0036]	-0.0368*** [0.0041]	-0.0124* [0.0066]
Some Post Secondary			-0.004 [0.0047]	-0.005 [0.0046]	-0.0071 [0.0056]	-0.0203*** [0.0069]
University Degree			0.0169** [0.0066]	0.0162** [0.0066]	0 [0.0085]	-0.0106 [0.0099]
None Reported			-0.0536*** [0.0061]	-0.0555*** [0.0061]	-0.0602*** [0.0066]	-0.0169 [0.0113]
Father's Education						
Less than High School			-0.0252*** [0.0039]	-0.0243*** [0.0039]	-0.0236*** [0.0045]	0.0084 [0.0080]
Some Post Secondary			0.0089* [0.0053]	0.0097* [0.0053]	0.0124** [0.0062]	0.0077 [0.0071]
University Degree			0.0161*** [0.0056]	0.0173*** [0.0056]	0.0228*** [0.0073]	0.0102 [0.0086]
None Reported			-0.0455*** [0.0061]	-0.0432*** [0.0061]	-0.0378*** [0.0066]	0.01 [0.0120]
Parental Immigration						
Immigrant Mother			0.0098** [0.0045]	0.0097** [0.0045]	0.0056 [0.0052]	-0.0120* [0.0068]
Immigrant Father			0.0173*** [0.0042]	0.0181*** [0.0042]	0.0267*** [0.0049]	0.0165*** [0.0059]
Ability Proxies						
Good Math Grades				0.0154*** [0.0034]		
Teachers Too Fast				-0.0248*** [0.0035]		
Constant	4.9607*** [0.0186]	5.0708*** [0.0122]	4.9617*** [0.0189]	4.9958*** [0.0209]	5.0608*** [0.0189]	4.6240*** [0.0879]
Observations	12,573	12,573	12,573	12,573	10,728	10,728
R-squared	0.51	0.50	0.53	0.53	0.49	

Note: Dependent variable is log of the average prose and document literacy scores. Standard errors in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Pooled Quantile Regressions with Cohort Effects- Canada

VARIABLE	OLS	10th Quantile	Median	90th Quantile
Female	0.0107** [0.0053]	0.0124* [0.0073]	0.0014 [0.0026]	0.0026 [0.0037]
Years of Schooling	0.0860*** [0.0069]	0.1219*** [0.0047]	0.0883*** [0.0017]	0.0472*** [0.0030]
Years of Schooling Squared	-0.0021*** [0.0002]	-0.0030*** [0.0002]	-0.0022*** [0.0001]	-0.0010*** [0.0001]
Age of Respondent				
Age: 36-45	-0.0190* [0.0111]	-0.0323** [0.0162]	-0.0260*** [0.0055]	-0.0392*** [0.0085]
Age: 46-55	-0.0394*** [0.0141]	-0.0251 [0.0213]	-0.0553*** [0.0076]	-0.0763*** [0.0108]
Age: 56-65	-0.0709*** [0.0199]	-0.0824*** [0.0272]	-0.1059*** [0.0104]	-0.1088*** [0.0142]
Age: 66+	-0.1587*** [0.0319]	-0.1941*** [0.0344]	-0.1774*** [0.0137]	-0.2413*** [0.0162]
Mother's Education				
Less than High School	-0.0325*** [0.0066]	-0.0350*** [0.0102]	-0.0233*** [0.0035]	-0.0241*** [0.0057]
Some Post Secondary	-0.0079 [0.0085]	-0.0187 [0.0150]	-0.0043 [0.0052]	0.0101 [0.0065]
University Degree	-0.0202 [0.0204]	-0.0129 [0.0237]	-0.0028 [0.0083]	0.011 [0.0075]
None Reported	-0.0515*** [0.0112]	-0.0429*** [0.0141]	-0.0466*** [0.0061]	-0.0235*** [0.0084]
Father's Education				
Less than High School	-0.0219*** [0.0081]	-0.0347*** [0.0124]	-0.0277*** [0.0040]	-0.0193*** [0.0065]
Some Post Secondary	0.0165 [0.0110]	0.0101 [0.0177]	0.0072 [0.0058]	0.0062 [0.0087]
University Degree	0.0138 [0.0119]	0.0183 [0.0184]	0.0028 [0.0066]	0.0118 [0.0078]
None Reported	-0.0323*** [0.0114]	-0.0439*** [0.0154]	-0.0409*** [0.0059]	-0.0444*** [0.0086]
Cohort				
Cohort 2	0.0185** [0.0093]	0.0083 [0.0143]	0.0254*** [0.0053]	0.0294*** [0.0071]
Cohort 3	0.0272* [0.0160]	0.0291 [0.0212]	0.0470*** [0.0088]	0.0462*** [0.0112]
Cohort 4	0.0406 [0.0282]	0.0622** [0.0279]	0.0604*** [0.0120]	0.0963*** [0.0138]
Cohort 5	-0.0169 [0.0316]	-0.0381 [0.0341]	0.0044 [0.0134]	0.0786*** [0.0159]
Constant	4.9387*** [0.0558]	4.4733*** [0.0429]	4.9640*** [0.0127]	5.4281*** [0.0231]
Observations	14,899	14,899	14,899	14,899

Note: Dependent variable is log of the average prose and document literacy scores. Standard errors in brackets. Cohort 1 (the omitted category) consists of people who were age 26 to 35 in 1994; Cohort 2 consists of people who were 36 to 45 in 1994; Cohort 3 consists of people who were 46 to 55 in 1994; Cohort 4 consists of people who were 56 to 65 in 1994; and cohort 5 consists of people who were over 65 in 1994. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Pooled Quantile Regressions with Cohort Effects, by Education Group- Canada

VARIABLE	High School or Less			University Degree		
	10th Quantile	Median	90th Quantile	10th Quantile	Median	90th Quantile
Female	0.0776*** [0.0181]	0.0115*** [0.0036]	0.0043 [0.0067]	-0.0027 [0.0106]	0.0126*** [0.0039]	0.0008 [0.0007]
Age of Respondent						
Age: 36-45	-0.015 [0.0385]	-0.0115 [0.0077]	-0.0256* [0.0153]	-0.0570*** [0.0209]	-0.0071 [0.0090]	0.0072*** [0.0013]
Age: 46-55	-0.0271 [0.0489]	-0.0342*** [0.0107]	-0.0842*** [0.0226]	-0.0738** [0.0326]	-0.0602*** [0.0119]	-0.0276*** [0.0018]
Age: 56-65	-0.0732 [0.0637]	-0.0569*** [0.0141]	-0.1106*** [0.0260]	-0.2091*** [0.0395]	-0.1193*** [0.0147]	-0.0391*** [0.0024]
Age: 66+	-0.0556 [0.0978]	-0.1691*** [0.0185]	-0.2025*** [0.0307]	-0.3625*** [0.0487]	-0.1905*** [0.0183]	-0.0692*** [0.0037]
Mother's Education						
Less than High School	-0.1341*** [0.0263]	-0.0714*** [0.0054]	-0.0529*** [0.0091]	-0.0235* [0.0140]	-0.0240*** [0.0046]	0.0132*** [0.0011]
Some Post Secondary	-0.0168 [0.0368]	-0.0042 [0.0093]	0.0103 [0.0151]	0.0119 [0.0152]	0.0230*** [0.0059]	0.0427*** [0.0010]
University Degree	-0.1739 [0.1229]	-0.0044 [0.0160]	0.0103 [0.0215]	0.0228 [0.0200]	0.0199** [0.0083]	0.0439*** [0.0010]
None Reported	-0.2001*** [0.0309]	-0.0828*** [0.0072]	-0.0487*** [0.0127]	-0.1369*** [0.0210]	-0.0387*** [0.0111]	0.0245*** [0.0030]
Father's Education						
Less than High School	-0.0769** [0.0326]	-0.0613*** [0.0061]	-0.0167 [0.0105]	-0.0312** [0.0153]	-0.0063 [0.0051]	-0.0210*** [0.0010]
Some Post Secondary	0.0294 [0.0594]	0.0214** [0.0100]	-0.0024 [0.0150]	-0.0059 [0.0160]	0.0058 [0.0071]	-0.0036*** [0.0013]
University Degree	0.0725 [0.0687]	0.0034 [0.0130]	-0.0019 [0.0186]	-0.0123 [0.0156]	0.0002 [0.0066]	0.0185*** [0.0010]
None Reported	-0.0863** [0.0354]	-0.0772*** [0.0078]	-0.0380*** [0.0131]	-0.0542 [0.0399]	-0.0429*** [0.0110]	0.0476*** [0.0024]
Cohort						
Cohort 2	0.0169 [0.0310]	0.0309*** [0.0075]	0.0540*** [0.0180]	0.0184 [0.0243]	0.0089 [0.0081]	0.0150*** [0.0012]
Cohort 3	-0.0376 [0.0473]	-0.0033 [0.0118]	0.0449** [0.0218]	0.0945*** [0.0304]	0.0455*** [0.0113]	0.0200*** [0.0021]
Cohort 4	-0.2103** [0.0872]	-0.0398** [0.0164]	0.0326 [0.0270]	0.1634*** [0.0416]	0.0824*** [0.0144]	0.0208*** [0.0033]
Cohort 5	-0.3225*** [0.0946]	-0.1069*** [0.0181]	0.0133 [0.0299]	0.1739*** [0.0467]	0.0716*** [0.0181]	-0.0728*** [0.0040]
Constant	5.4820*** [0.0456]	5.7031*** [0.0078]	5.8408*** [0.0163]	5.7160*** [0.0187]	5.8018*** [0.0091]	5.8792*** [0.0014]
Observations	9,368	9,368	9,368	2,168	2,168	2,168

Note: Dependent variable is log of the average prose and document literacy scores. Standard errors in brackets. Cohort 1 (the omitted category) consists of people who were age 26 to 35 in 1994; Cohort 2 consists of people who were 36 to 45 in 1994; Cohort 3 consists of people who were 46 to 55 in 1994; Cohort 4 consists of people who were 56 to 65 in 1994; and cohort 5 consists of people who were over 65 in 1994. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Average Prose and Document Literacy Regressions- Norway

VARIABLE	OLS	10th Quantile	Median	90th Quantile
Female	-0.0072 [0.0057]	-0.0281** [0.0117]	0.0012 [0.0053]	-0.0024 [0.0071]
Years of Schooling	0.0612*** [0.0068]	0.0861*** [0.0114]	0.0608*** [0.0049]	0.0395*** [0.0075]
Years of Schooling Squared	-0.0015*** [0.0002]	-0.0022*** [0.0004]	-0.0015*** [0.0002]	-0.0009*** [0.0003]
Age of Respondent				
Age: 35-44	-0.0154 [0.0115]	-0.0388 [0.0239]	-0.0143 [0.0103]	-0.0012 [0.0140]
Age: 45-54	-0.0741*** [0.0159]	-0.1346*** [0.0342]	-0.0633*** [0.0149]	-0.0321 [0.0202]
Age: 55-65	-0.1201*** [0.0193]	-0.1731*** [0.0411]	-0.1021*** [0.0184]	-0.0817*** [0.0246]
Mother's Education				
Less than High School	-0.0077 [0.0068]	-0.017 [0.0153]	-0.0093 [0.0066]	0.0087 [0.0088]
PS, University, or More	-0.0034 [0.0132]	-0.0601** [0.0260]	0.0081 [0.0100]	0.0171 [0.0136]
None Reported	-0.0430** [0.0184]	-0.0829** [0.0341]	-0.0448*** [0.0166]	-0.0074 [0.0250]
Father's Education				
Less than High School	-0.01 [0.0070]	-0.0072 [0.0146]	-0.0065 [0.0063]	-0.0121 [0.0086]
PS, University, or More	0.0118 [0.0089]	0.0104 [0.0200]	0.0133 [0.0082]	0.0116 [0.0105]
None Reported	-0.0007 [0.0145]	0.0345 [0.0244]	0.0089 [0.0143]	-0.0087 [0.0207]
Cohort				
Cohort 2	0.0086 [0.0120]	0.0471* [0.0249]	0.0023 [0.0105]	-0.0098 [0.0141]
Cohort 3	0.0426** [0.0167]	0.1185*** [0.0345]	0.0299** [0.0150]	-0.002 [0.0204]
Cohort 4	0.0628*** [0.0200]	0.1566*** [0.0418]	0.0414** [0.0185]	0.0115 [0.0254]
Cohort 5	0.0331 [0.0238]	0.1106** [0.0474]	0.007 [0.0219]	0.0045 [0.0294]
Constant	5.1816*** [0.0481]	4.8114*** [0.0801]	5.2102*** [0.0357]	5.4932*** [0.0548]
Observations	3,247	3,247	3,247	3,247

Note: Dependent variable is log of the average prose and document literacy scores. Standard errors in brackets. Cohort 1 (omitted category) consists of people who were from 16 to 25 in 1998; Cohort 2 consists of people who were from 26 to 35 in 1998; Cohort 3 consists of people who were from 36 to 45 in 1998; Cohort 4 consists of people who were from 46 to 55 in 1998; and Cohort 5 consists of people who were from 56 to 65 in 1998. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Average Prose and Document Literacy Regressions- US

VARIABLE	OLS	10th Quantile	Median	90th Quantile
Female	0.0266*** [0.0084]	0.0343** [0.0157]	0.012 [0.0082]	0.0033 [0.0116]
Years of Schooling	0.0553*** [0.0160]	0.1381*** [0.0099]	0.0788*** [0.0069]	0.0217 [0.0145]
Years of Schooling Squared	-0.0010* [0.0005]	-0.0033*** [0.0004]	-0.0018*** [0.0002]	-0.0002 [0.0005]
Age of Respondent				
Age: 35-44	-0.025 [0.0154]	-0.0299 [0.0298]	-0.0398** [0.0156]	0.0021 [0.0198]
Age: 45-54	-0.0497** [0.0238]	-0.026 [0.0438]	-0.0897*** [0.0221]	-0.0417 [0.0300]
Age: 55-65	-0.1083*** [0.0292]	-0.061 [0.0534]	-0.1581*** [0.0279]	-0.1157*** [0.0383]
Mother's Education				
Less than High School	-0.0668*** [0.0117]	-0.0805*** [0.0187]	-0.0535*** [0.0112]	-0.0452*** [0.0164]
PS, University, or More	0.0047 [0.0118]	0.0109 [0.0242]	-0.0023 [0.0128]	0.0156 [0.0189]
None Reported	-0.1051*** [0.0244]	-0.1215*** [0.0304]	-0.0928*** [0.0187]	-0.0502* [0.0263]
Father's Education				
Less than High School	-0.0238** [0.0113]	0.0132 [0.0195]	-0.0226** [0.0113]	-0.0290* [0.0169]
PS, University, or More	0.0172* [0.0104]	0.0499** [0.0214]	0.0102 [0.0120]	-0.0007 [0.0168]
None Reported	-0.0963*** [0.0180]	-0.1005*** [0.0268]	-0.0799*** [0.0159]	-0.0904*** [0.0211]
Cohort				
Cohort 2	0.0308** [0.0153]	0.0489 [0.0310]	0.0407** [0.0160]	0.0216 [0.0201]
Cohort 3	0.0423* [0.0231]	0.0205 [0.0430]	0.0844*** [0.0222]	0.0436 [0.0294]
Cohort 4	0.0811*** [0.0283]	0.0672 [0.0527]	0.1212*** [0.0278]	0.0880** [0.0381]
Constant	5.0621*** [0.1180]	4.1484*** [0.0749]	4.9276*** [0.0506]	5.5485*** [0.1068]
Observations	1,953	1,953	1,953	1,953

Note: Dependent variable is log of the average prose and document literacy scores. Standard errors in brackets. Cohort 1 (omitted category) consists of people who were from 16 to 25 in 1994; Cohort 2 consists of people who were from 26 to 35 in 1994; Cohort 3 consists of people who were from 36 to 45 in 1994; Cohort 4 consists of people who were from 46 to 55 in 1994. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A1: First Stage Results for Years of Schooling

VARIABLE	OLS
Female	0.3380*** [0.0463]
Age of Respondent	
Age	0.0443*** [0.0103]
Age Squared	-0.0846*** [0.0096]
Mother's Education	
Less than High School	-0.6764*** [0.0633]
Some Post Secondary	0.4039*** [0.0877]
University Degree	0.3427*** [0.1325]
None Reported	-1.2574*** [0.1016]
Father's Education	
Less than High School	-0.9459*** [0.0687]
Some Post Secondary	0.1273 [0.0967]
University Degree	0.3586*** [0.1130]
None Reported	-1.4126*** [0.1021]
Parental Immigration	
Immigrant Mother	0.4797*** [0.0811]
Immigrant Father	0.2323*** [0.0764]
Excluded Instruments	
SLA=15 years old	0.8533*** [0.0807]
SLA=16 years old	0.8268*** [0.0845]
Constant	12.0155*** [0.2796]
Observations	10,728
R-squared	0.29
F-Statistic	58.47

Note: Dependent variable is years of schooling. Standard errors in brackets.
Omitted School Leaving Age is 14 years old. * significant at 10%; ** significant at 5%; *** significant at 1%

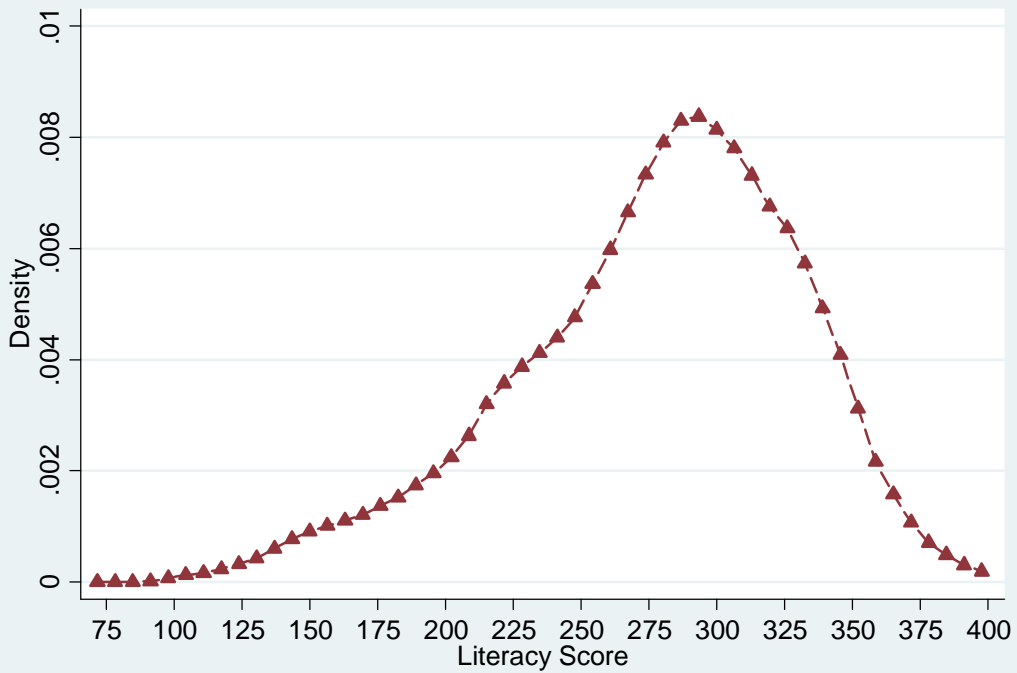


Figure 1: Average Literacy, 2003

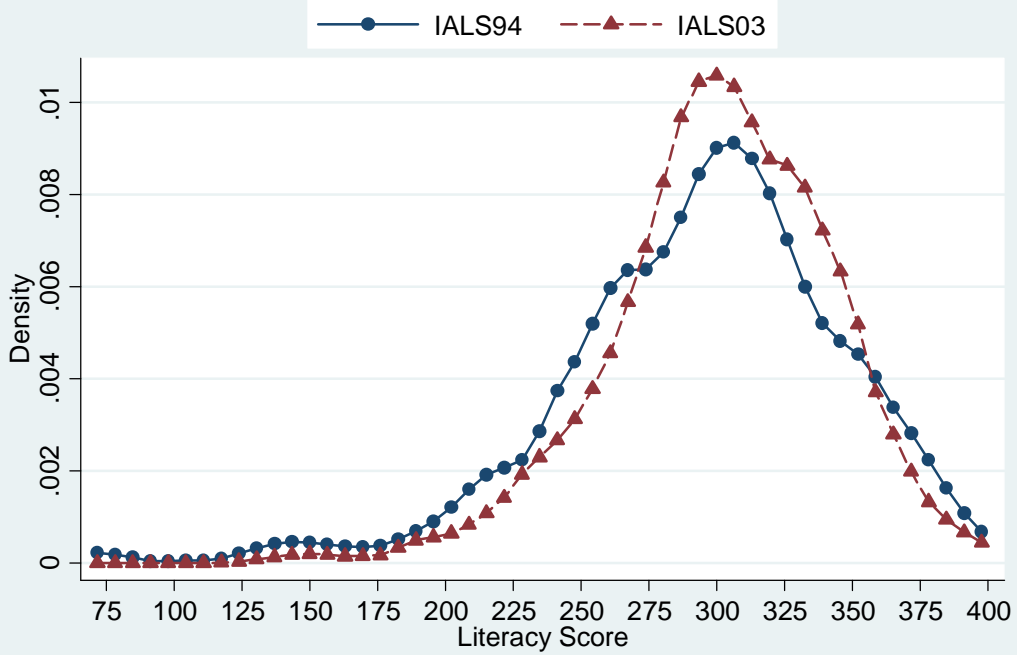


Figure 2: Average Literacy, Age 26-35

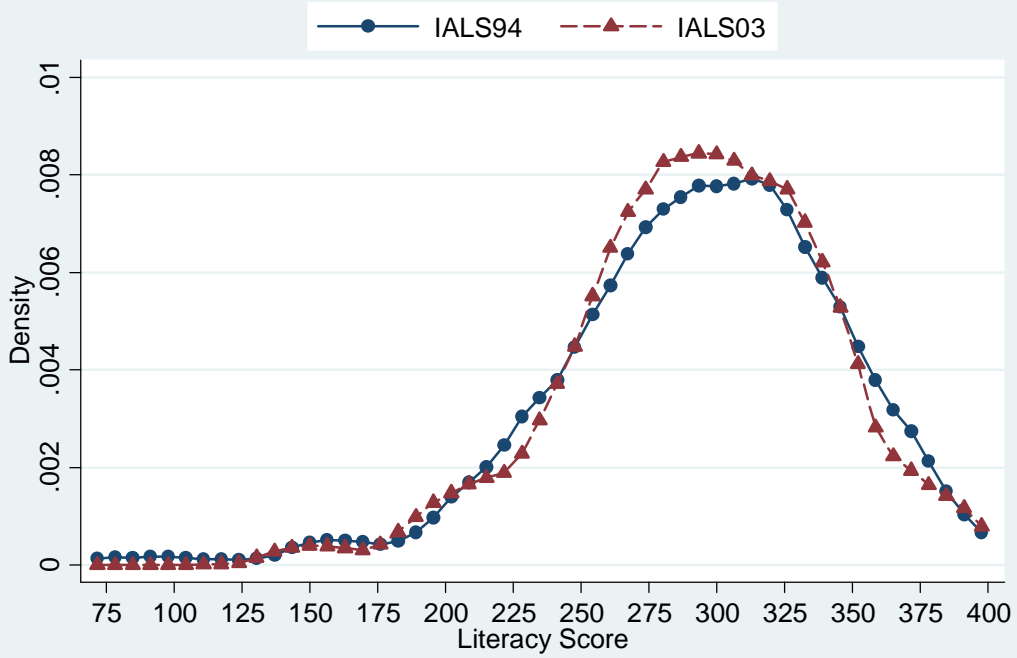


Figure 3: Average Literacy, Age 36-45

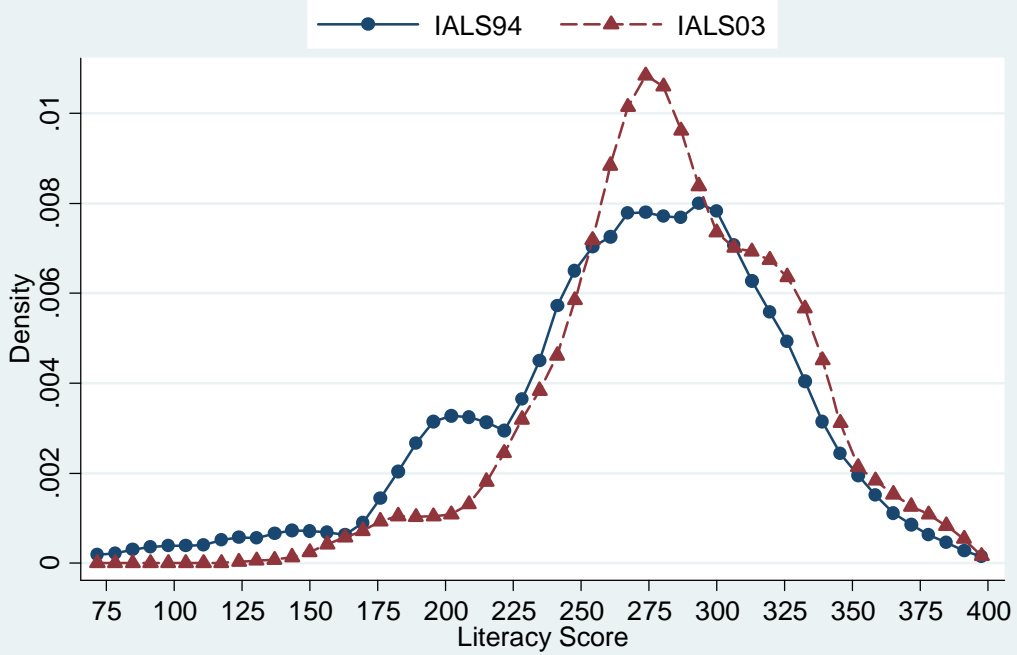


Figure 4: Average Literacy, Age 46-55

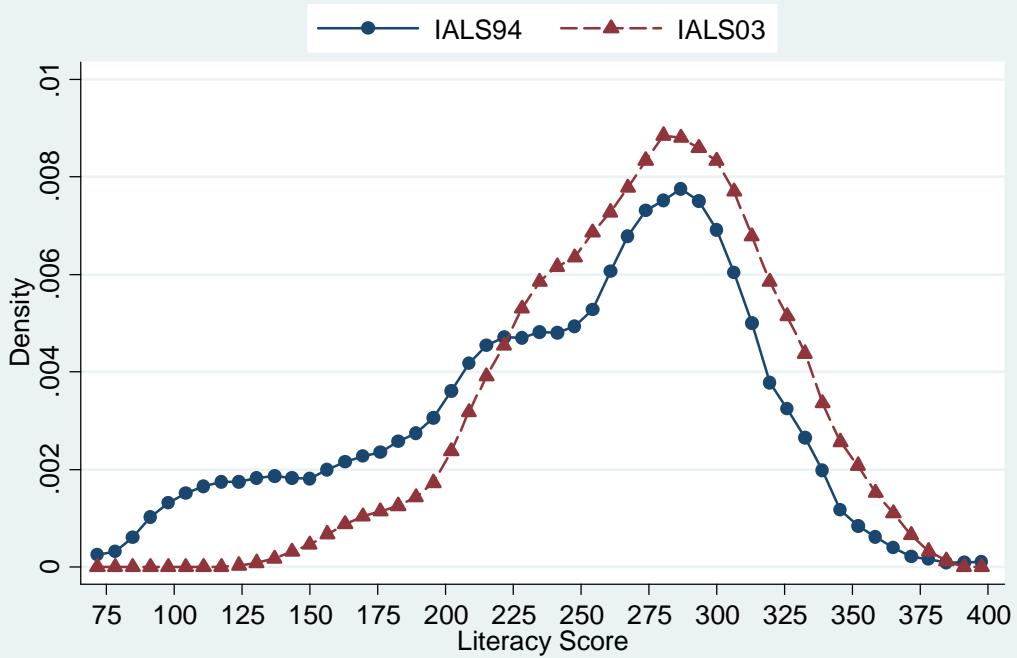


Figure 5: Average Literacy, Age 56-65

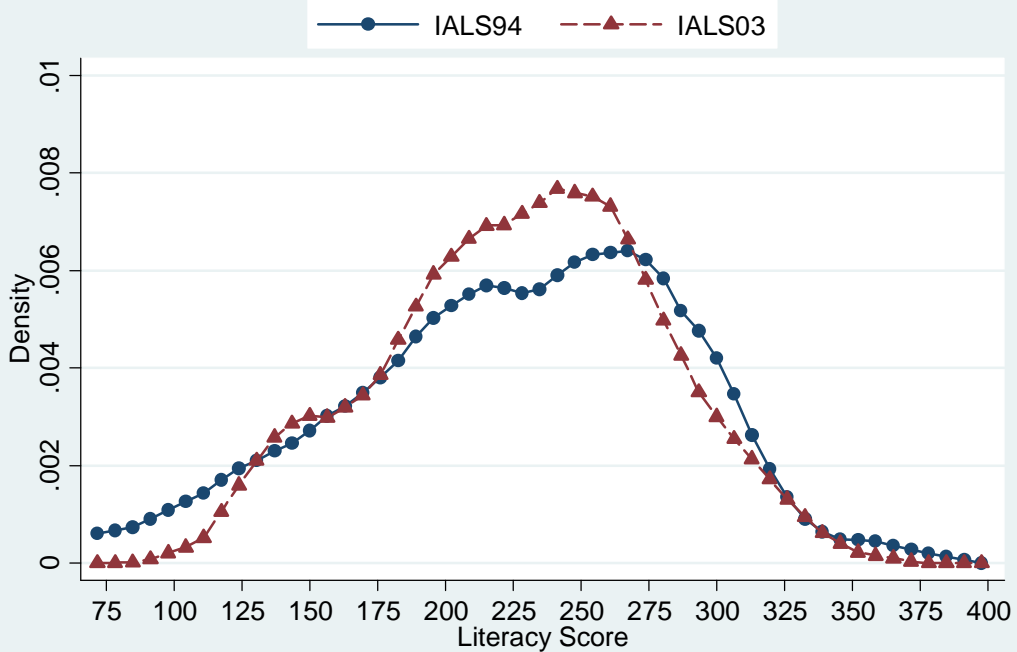


Figure 6: Average Literacy, Less than HS

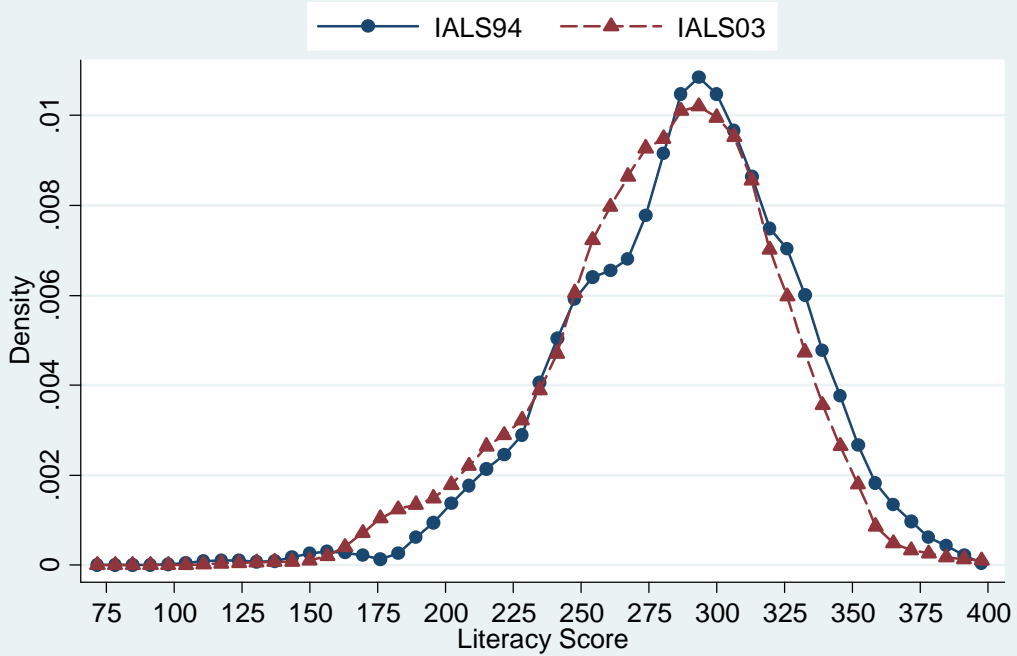


Figure 7: Average Literacy, High School

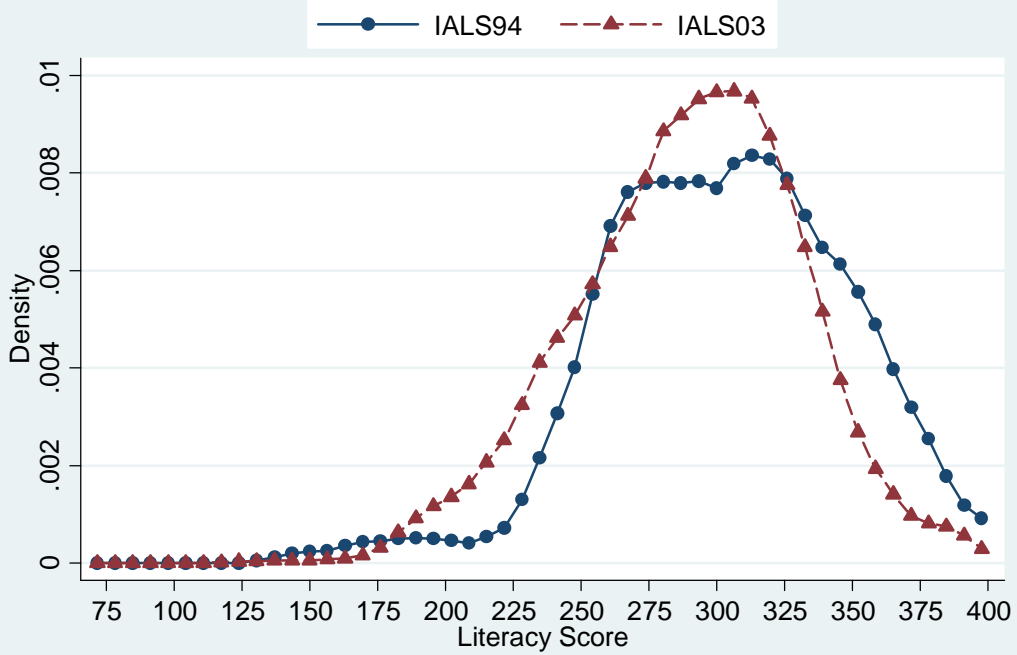


Figure 8: Average Literacy, College

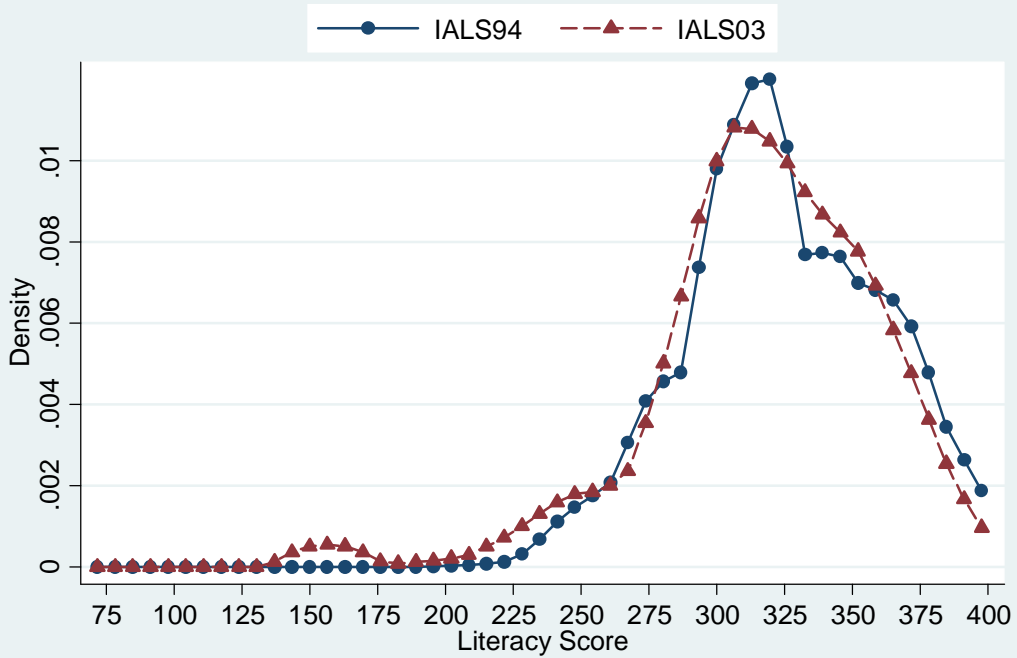


Figure 9: Average Literacy, University

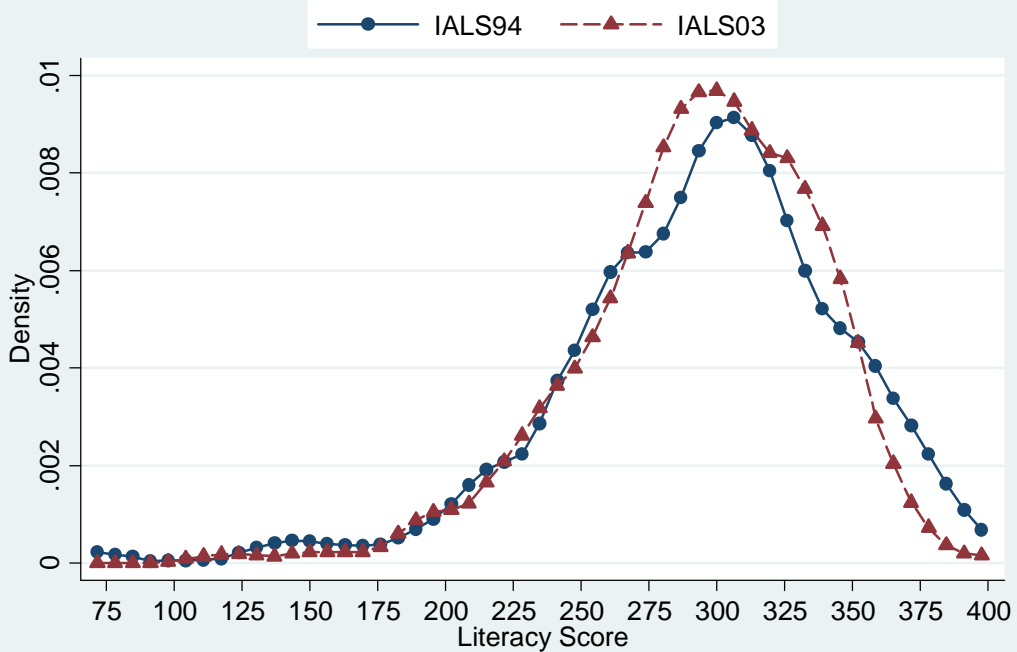


Figure 10: Average Literacy, 26-35 in 1994

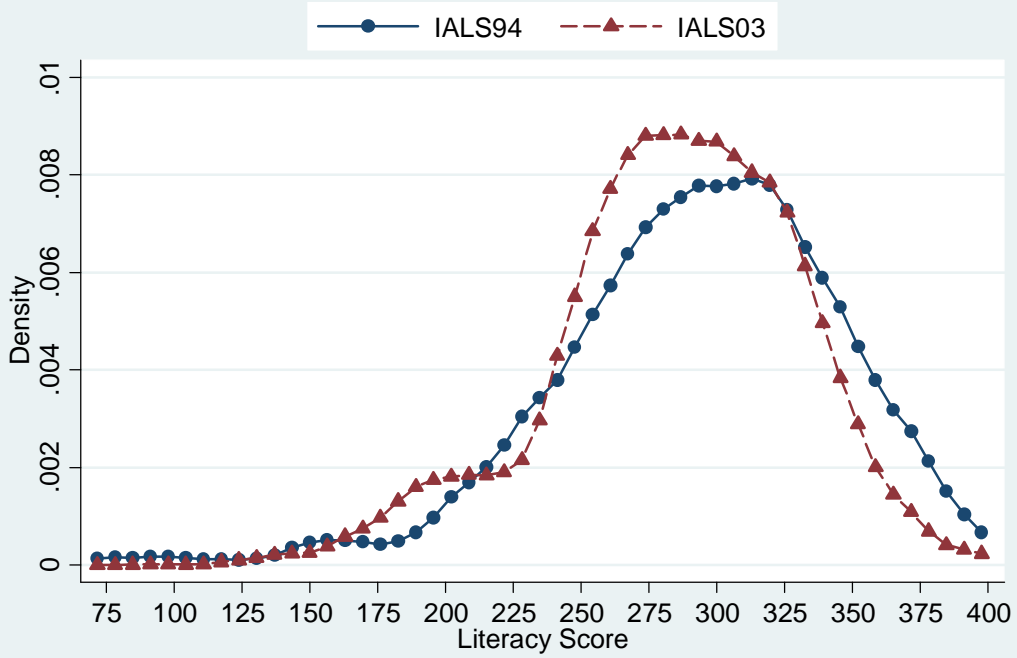


Figure 11: Average Literacy, 36-45 in 1994

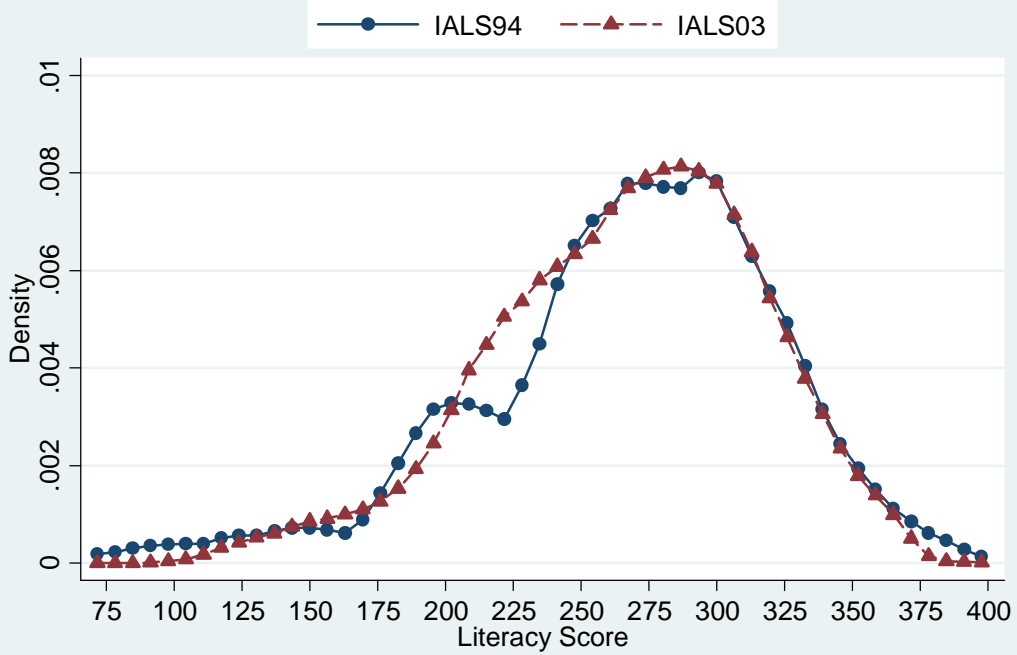


Figure 12: Average Literacy, 46-55 in 1994

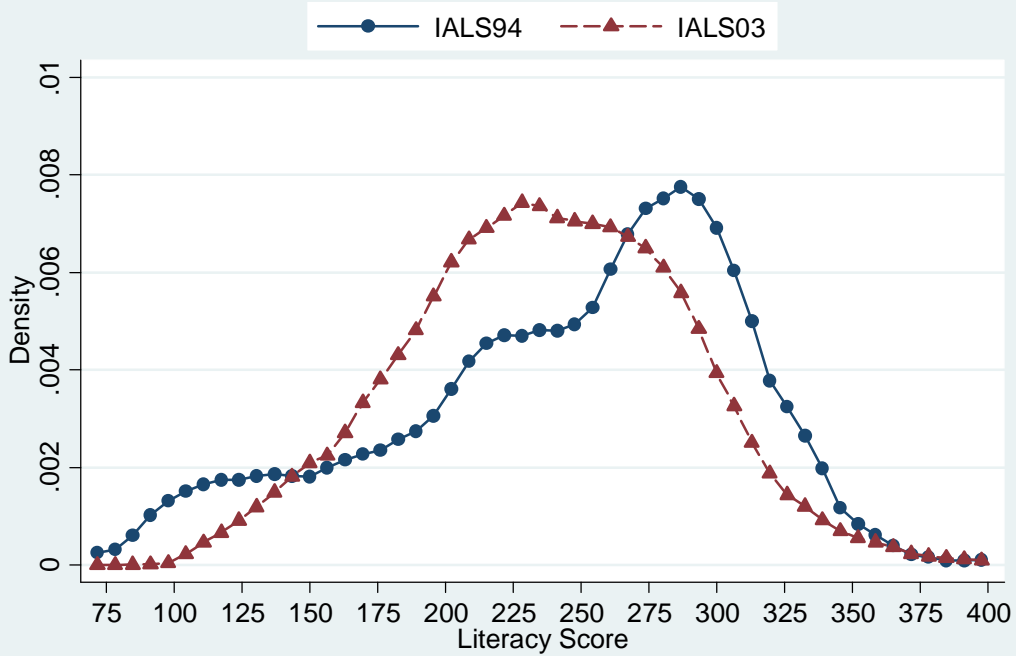


Figure 13: Average Literacy, 56-65 in 1994

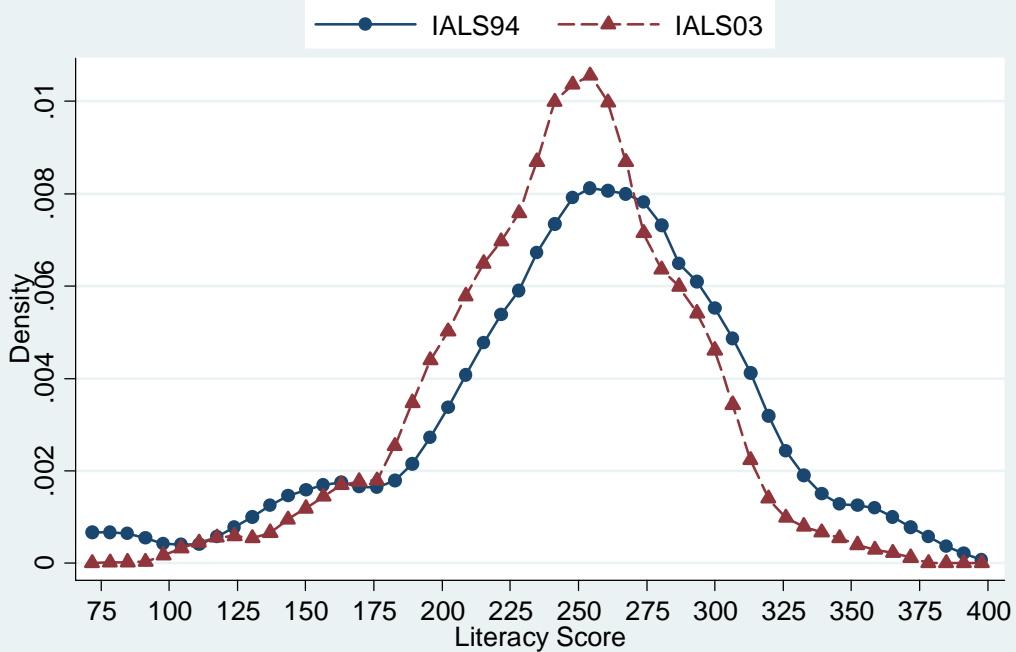


Figure 14: Average Literacy, Less than HS, 26-45 in 1994

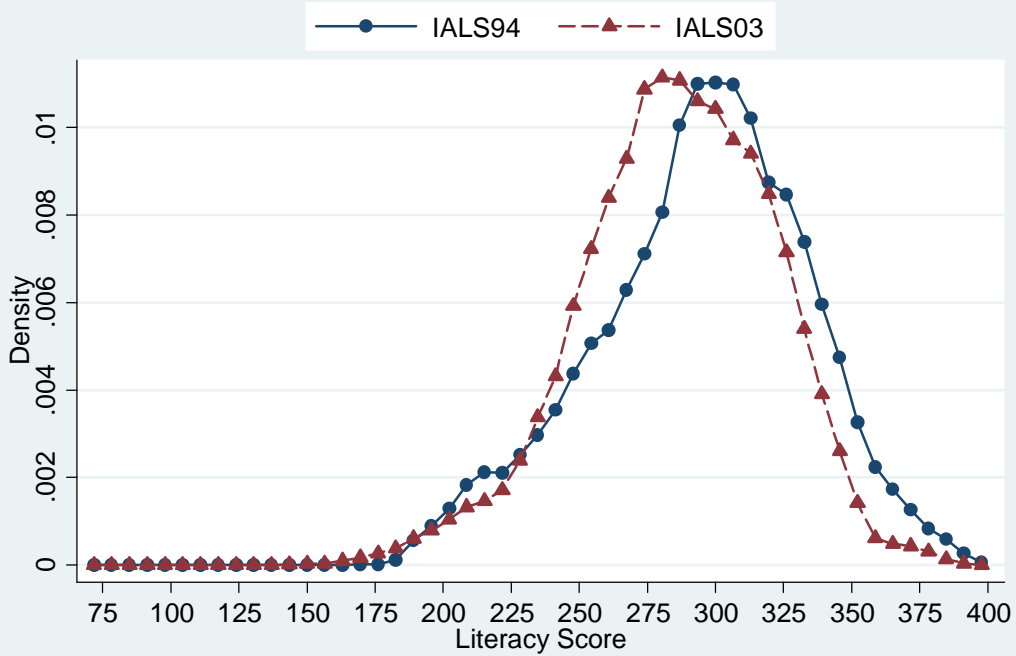


Figure 15: Average Literacy, High School, 26-45 in 1994

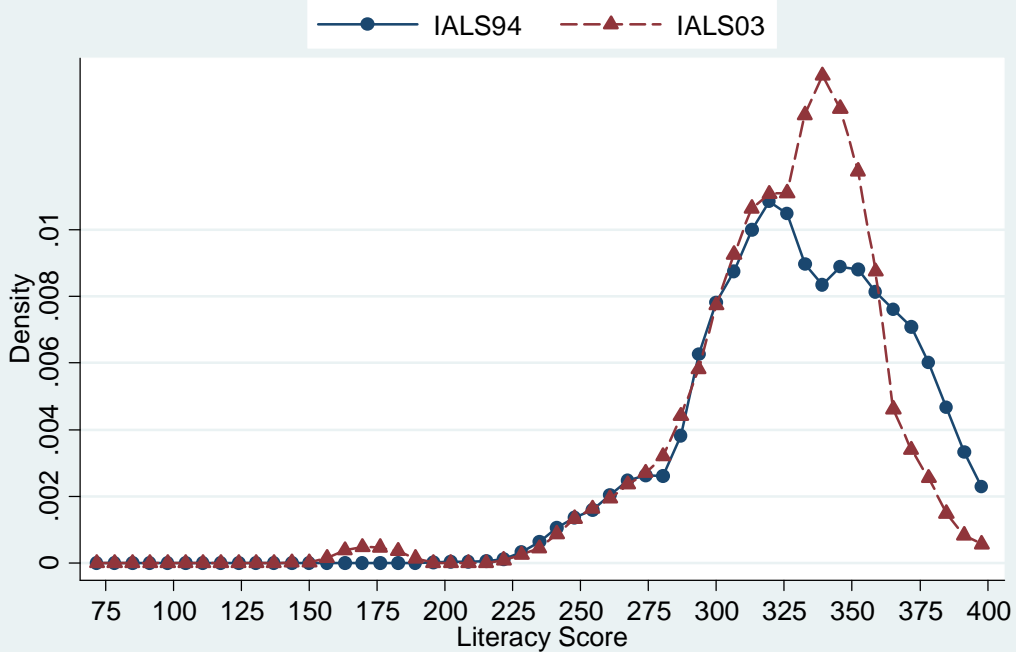


Figure 16: Average Literacy, University, 26-45 in 1994