Catalogue no. 89-552-MIE, no. 7

Schooling, Literacy and Individual Earnings

Lars Osberg

Statistics Canada
Statistique Canada

Human Resources Development Canada
Développement des ressources humaines Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: T. Scott Murray, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: (613) 951-9035).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our Web site.

National inquiries line	**1 800 263-1136**
National telecommunications device for the hearing impaired	**1 800 363-7629**
Depository Services Program inquiries	**1 800 700-1033**
Fax line for Depository Services Program	**1 800 889-9734**
E-mail inquiries	**infostats@statcan.ca**
Web site	**www.statcan.ca**

Ordering and subscription information

This product, Catalogue no. 89-552-XPE, is published (irregularly) as a standard printed publication at a price of CDN $10.00 per issue. The following additional shipping charges apply for delivery outside Canada:

	Single issue
United States	CDN $ 10.00
Other countries	CDN $ 10.00

This product is also available in electronic format on the Statistics Canada Internet site as Catalogue no. 89-552-XIE free of charge. To obtain an issue visit our Web site at **www.statcan.ca**, and select Products and Services.

All prices exclude sales taxes.

The printed version of this publication can be ordered by

- Phone (Canada and United States) **1 800 267-6677**
- Fax (Canada and United States) **1 877 287-4369**
- E-mail **order@statcan.ca**
- Mail Statistics Canada
Dissemination Division
Circulation Management
120 Parkdale Avenue
Ottawa, Ontario K1A 0T6
- And, in person at the Statistics Canada Regional Centre nearest you, or from authorized agents and bookstores.

When notifying us of a change in your address, please provide both old and new addresses.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1 800 263-1136.

International Adult Literacy Survey

Schooling, Literacy and Individual Earnings

Lars Osberg

Department of Economics, Dalhousie University, Halifax, Canada

The International Adult Literacy Survey (IALS) was a seven-country initiative conducted in the fall of 1994. The Canadian component of the IALS study was primarily funded by the Applied Research Branch and the National Literacy Secretariat of Human Resources Development Canada.

Published by authority of the Minister responsible for Statistics Canada

June 2000

Catalogue no. 89-552-MPE, no. 7

ISSN 1480-1566

ISBN 0-660-17956-3

Catalogue no. 89-552-MIE, no. 7

ISSN 1480-9516

ISBN 0-660-17956-3

Frequency: Irregular

Ottawa

Statistics Canada

Human Resources Development Canada

National Literacy Secretariat

The data interpretations and policy prescriptions presented in this report are those of the authors and do not necessarily reflect those of the granting agencies or reviewers.

Canadian Cataloguing in Publication Data

Osberg, Lars
Schooling, literacy and individual earnings

(International Adult Literacy Survey, ISSN 1480-1566)
Co-published by: Human Resources Development Canada
[and] National Literacy Secretariat.
Issued also in French under title: Scolarité, alphabétisme et
gains personnels.
ISBN 0-660-17956-3
CS89-552-MPE no. 7

1. Literacy – Economic aspects – Canada. 2. Education –
Economic aspects – Canada. 3. Wages – Effect of education –
Canada. I. Statistics Canada. II. Canada. Human Resources Development Canada.
III. Canada. National Literacy Secretariat. IV. Series.

LC154 O82 2000 331.2’15
C99-988045-4

Acknowledgements

I am grateful to Statistics Canada, Human Resources Development Canada and the National Center for Education Statistics, U.S. Department of Education for their funding of research on youth literacy.

I would like to thank Scott Murray of Statistics Canada for his assistance in obtaining the micro-data on which this paper is based, Mark Lewis for his exemplary work as research assistant, and the Social Policy Research Centre of the University of New South Wales for its hospitality. Kentaro Yamomoto of the Educational Testing Service provided useful comments. Jeannie Boyle and Lynn Lethbridge assisted the passage through the revision process. Errors remaining are my sole responsibility.

Note of Appreciation

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

Table of Contents

List of figures

List of tables

Introduction

How much does literacy affect earnings? More precisely, how much of the economic payoff to education can be explained by literacy skills? Although literacy is clearly one of the main outcomes of the educational system, most of the vast literature on the impact of education on earnings actually measures the impact of *inputs* into the educational system—such as years of school attendance or expenditures on teachers and other school resources. There is a great deal of public discussion about whether educational standards are rising or declining, and about how this might be affecting Canada's productivity performance. However, in general, economic studies on the benefits of education have not used direct measures of educational outcomes (such as literacy) to explain individual earnings.

This paper uses direct measures of literacy skill levels provided by the International Adult Literacy Survey (IALS) to estimate the return to literacy skills. Using a very simple human capital earnings equation and standard ordinary least squares (OLS) regression, it tests estimates of the return to literacy skills for their robustness to alternative scalings of literacy attainment.

This paper emphasizes the importance of alternative possible scalings of literacy scores, because literacy scores are inherently ordinal, not cardinal, numbers. Although tests of literacy skills can be used to assess whether one individual is "more literate" than another, the statement that one individual is "25% more literate," or "10% less literate," than another would make little sense. The size of the differences between individuals at each point in the distribution of literacy (e.g., the magnitude of the differences among those with poor literacy skills, or among those who are highly skilled) is simply impossible to measure.

This implies that the scaling of literacy scores is inherently somewhat arbitrary. However, it has become commonplace to make comparisons between the average literacy levels of different jurisdictions. These comparisons are often used to justify appeals for additional public expenditure, on the basis that a more literate work force will be more productive. Although this is likely true *to some degree*, a rigorous cost–benefit calculation requires an answer to the question "*how much* does literacy matter?".

For public policy purposes, it would be useful to know which of the skills produced by the educational system actually pay off, and by how much. In trying to answer this sort of question, the standard reaction of many economists is to use some variety of regression analysis to estimate the return to one characteristic (such as years of schooling), holding constant other influences.

However, there are good reasons for caution when using measures of skill levels as explanatory variables to predict earnings. Although input measures such as years of schooling or per-pupil expenditures have natural metrics (such as years or dollars) which are cardinal numbers, there is no natural metric for literacy, or for other social or cognitive skills. Literacy tests can rank individuals, but there may be many ways to scale literacy scores that are equally plausible. Unfortunately, standard econometric techniques such as multiple regression assume that all variables are cardinal numbers. Since literacy attainment is an ordinal ranking, it is therefore important to test the robustness of empirical results to alternative possible scalings of literacy scores.

This paper explores the sensitivity of econometric results to a range of monotonic transformations[1] of literacy scores and concludes that, particularly for males, much of the economic return yielded by education is due to literacy skills—perhaps as much as 40% to 45%. While education has a higher overall rate of return for females than for males, not as much of that return is explained by literacy. Literacy may also account for a larger proportion of the impact of education on earnings among those with higher literacy skills. We conclude that literacy does affect earnings over quite a wide range of metrics. However, alternative possible scalings of literacy scores also imply that the rankings of Canadian provinces in average literacy vary with the scaling adopted, sometimes quite dramatically.[2]

Section 1

Methodology

Ordinality and inference

Literacy scores are a special case of a more general problem. Like many other public services, the public education system would like to have a measure of the achievement of its students, in order to judge the effectiveness of its operation. However, *quality* is a key dimension of education. In education, as in many other public services, there is no inherently obvious way of putting a unique number on "quality".

When assessing the average quality of public services (like education), it is often possible to determine whether one outcome is preferable to another, but it is rarely possible to assign a unique ratio to that preference. Measures of outcome quality, while often crucial to program evaluation, are also highly problematic. Since quality measures are usually ordinal rather than cardinal numbers, their size depends heavily on the scale used.

Measures of educational outcomes, such as school grades or literacy scores, provide a good example. Even though we may agree that an A is better than a B, which in turn is better than a C, we may well disagree on how to average an A and a C.

If the readers of this paper were told only that A>B>C and were then asked to compute the average of A, B and C, most would no doubt reply that averaging ordinal variables (such as A, B and C) does not make much sense. Yet it is common for professors to compute the grade point average of their students and to conclude, for example, that a student with an "A" and a "C" has a "B" average. Although the calculation (A+C)/2=B is only true if there is some prior agreement on the scaling of grades (e.g. A=3, B=2, C=1), such a scaling is inherently arbitrary. There is no meaningful way in which such a scaling implies that a student who earns an A knows "three times as much" as one who earns a C. The calculation of a grade-point average would rank some students quite differently if letter grades were weighted exponentially. For example, if A = 1,000, B = 100, C = 10, then the average of an A and a C is clearly better than B.

In the debate about whether educational standards are rising or falling, or whether standards in one jurisdiction are higher than another, it is common to calculate average test scores among students and to use these averages to establish rankings. Such rankings of average scores depend on the scales used to rank relative attainment, and on the differences across school populations in the dispersion of outcomes. For example, under the A=3, B=2, C=1 scaling, a school in which half the student body functions at an A level of competency and half scores at the C level will be ranked as equivalent to a student body which scores uniformly at the B level. However, the school with half A and half C achievement will be ranked as clearly superior under the A=1,000, B=100, C=10 scaling.[3]

Although conclusions about which school is best may well be affected by alternate scalings of grades, the ordinal nature of achievement scores may not always be important. To decide which individual student is best (or to decide which students to admit to elite graduate schools), all that matters is that a student with all A+ grades be ranked at the top by *any* monotonic transformation of achievement scores. The scaling of achievement scores will also not matter much if everyone gets much the same score. Whether the ordinal nature of measures of skill attainment "matters" or not will depend on what particular question is being asked, what alternative scalings of scores are reasonable, and the degree of dispersion of outcomes in the population.

This article attempts to discover what proportion of the economic returns from schooling can be explained by literacy skills, given that other variables (such as work experience) also affect individual earnings. To assess the relative contribution of literacy, a multivariate approach is required. This is a different issue than distributional dominance (that is, description of the relative location of the central tendency of two distributions—whether one group can be said to score higher than another), which has been the focus of much of the statistical literature on ordinal inference (see, for example, Cliff 1993).

Many authors have used regression models to predict ordinal relationships—for example, the probability that one individual is happier than another (see Maddala 1983 or McCullagh 1980). The probit, logit, proportional odds or proportional hazards models are familiar to economists and can be used to predict an ordinal relation if one can assume cardinal explanatory variables. Cliff also examines the issue of how best to use ordinal explanatory variables to predict a dominance relationship (1996:89–115).

However, the issue here is the use of an ordinal variate (literacy) as one of several potential explanatory variables. We want to know the relative size of its impact in explaining a cardinal variable (earnings), controlling for the influence of other variables (for example, work experience). The standard reflex of most empirical economists would be to enter an individual's literacy score as an additional right-hand-side variable in a multiple regression whose dependent variable is individual earnings. Unfortunately, the use of a particular scaling of an ordinal variate is open to the criticism that alternative scalings, which stretch or compress the distribution of measured scores at different points in their distribution, may alter the sign and/or statistical significance of an ordinal explanatory variable by altering the relative size of negative or positive deviations from the median.[4] Grether (1974) establishes the necessary and sufficient conditions for the sign of sample correlation coefficients to be invariant with respect to order-preserving transformations of one or both variables. However, the size of multiple regression coefficients is a much harder issue, analytically. It is not a solution to use only rank-order information in multiple regression, because doing so implicitly imposes its own scaling—that the difference between the top score and the next best score is of the same magnitude as the difference between any other pair of adjacent rankings (Kim 1975, 1978 and O'Brien 1982).

In general, if an independent variable X has an ordinal scale, the only type of statement one can make about a functional relationship U(X), which will also always be valid for any monotonic transformation of X, is that U(X) is a constant (Kim 1990:26-29). To get around this very negative general conclusion, one common approach has been to assume that an ordinal variable follows a specific distribution, and thereby limit the range of possible monotonic transformations (Cliff 1996:89).[5] However, since the whole point of making international comparisons of literacy (such as IALS) is to examine whether nations differ in their distribution of literacy levels, imposing the assumption of a common distribution would be counter-productive.

Assuming that any monotonic transformation of literacy scores is reasonable would, however, also be counter-productive. The class of order-preserving transformations is extremely broad and it is not surprising that little can be said in general about multivariate relationships in which an independent variable could undergo an arbitrary monotonic transformation.[6] The focus of this paper is, however, quite specific—the impact of literacy on earnings—and our knowledge of literacy may allow us to exclude some possible monotonic transformations of literacy scores as unreasonable.[7] If the set of reasonable monotonic transformations can be limited, it may be possible to assess empirically the robustness of the impact of literacy on earnings. However, assessing the range of reasonable transformations of IALS scores requires some discussion of how those scores are generated.

Section 2

Findings

The International Adult Literacy Survey

Most modern literacy studies, including the International Adult Literacy Survey (IALS), reject the historic dichotomy that labels individuals as literate or illiterate, in favour of the concept of a continuum of literacy skills in the population. It has become apparent that although historical notions of literacy may have been entirely appropriate, such standards are inadequate in a modern society. Today we define literacy as the ability to understand and use written information and we recognize that information requirements change with economic and social development. In the 19th century, the ability to sign one's name and thereby assent to legal contracts was considered a reasonable criterion of literacy—but in the 21st century "computer literacy" is required for many jobs.

For many years, schooling credentials, or years of attendance, were used as a proxy definition of literacy, largely because such measures are so easy to obtain. However, such a definition clearly cannot measure the adequacy of the school system in producing literacy skills, or the ability of individuals to acquire literacy skills without formal education. The IALS methodology was therefore developed to assess the functional literacy of individuals—the ability to use written information in real-life situations.

IALS breaks literacy down into three categories: prose literacy, document literacy and numeracy. Respondents were presented with realistic situations, of graduated difficulty, to assess their competency in each category. Numeracy was, for example, assessed by how well a person could balance a chequebook, calculate a tip, complete an order form or determine interest from a loan ad. Document literacy was defined as the ability to locate and use information from documents such as job applications, payroll forms, transportation schedules, maps, tables or graphs. Prose literacy was defined as the ability to understand and use information from texts such as editorials, news stories, poems and fiction.

In IALS, an individual score between 0 and 500 was assigned for each category, and scores were grouped into five broad levels of literacy for tabular analysis.[8] As Kirsch notes, however: "While the literacy scales make it possible to compare the prose, document and quantitative skills of different populations and to study the relationships between literacy skills and various factors, *the scale scores by themselves carry little or no meaning*" (Kirsch, 1995:27, emphasis added; see also OECD and Human Resource Development Canada, 1997:85).

In IALS, document, prose and quantitative literacy each have five separate scores (labelled Plausible Values 1, 2, 3, 4, and 5). As well, users of the data will note that many respondents were assigned scores outside the range of difficulty of the questions actually used in the survey. (For example, although the easiest question in the quantitative literacy segment was rated 225, and the hardest was rated 408, 26.0% of Canadian respondents were assigned a score of less than 225 and 0.5% were assigned a score over 408.)[9] To understand these issues, some familiarity with item response theory and its methods of imputation is essential.

Solving real-life problems using written information takes time. If a questionnaire testing functional literacy is to be completed within a reasonable total time, only a limited number of questions are possible. In IALS, there were a battery of 33 potential test items in quantitative literacy, 34 in document literacy and 34 in prose comprehension. These questions were divided into seven blocks. Each respondent answered three blocks (that is, no more than 15 questions in each major category of literacy).[10]

In a real world test situation, only a limited number of test items can be asked of each individual, and respondents do not always answer all questions asked. It is also not inherently obvious when designing the test which items respondents will find to be the most difficult (e.g., finding the best bus connection from a bus schedule, or the best buy on a radio from a Consumer Reports evaluation). Since each “real world” task takes significant time to answer, it is costly in terms of questionnaire administration to add additional test items. However, information on variables which can reasonably be expected to be correlated with literacy (such as years of schooling, or reading habits) can easily be obtained for each respondent, with very little response burden. Because it is relatively cheap to get background characteristics, and relatively expensive to include real world tasks, “item response theory” uses *all* this information to estimate literacy proficiency. This means that literacy scores embody actual responses to real world tasks *and* imputation based on background variables.

Item response theory (IRT) allows one to see test scores according to probability. Kirsch (1996) draws an analogy to the probability with which a high jumper would clear a bar set at a given level. Occasionally, an athlete of given competency will fail at a lower height and occasionally succeed above their normal level of achievement. If one sets a specific probability of success (80%), the literacy level of an individual can be defined as “the point at which individuals with that proficiency score have a given (80%) probability of responding correctly . . . .This means that individuals estimated to have the particular skill score will consistently perform tasks—with an 80% probability—like those at that point on the scale. It also means that they will have a greater than 80% chance of performing tasks that are lower than their estimated proficiency on the scale” (Kirsch 1996:86).

Estimating the probability that an individual with given characteristics will successfully complete a particular question is done using logistic regression. Mislevy and Beaton (1992:135) argue that “the essential idea of IRT is that observed item responses are driven by an unobservable proficiency variable.” They also point out that in using all the information available (for example, an individual’s responses to test items and such characteristics as age or education) “the distribution of point estimates that would be preferred for making inferences about individuals can depart substantially from the distribution of the underlying variable. The marginal procedures that possess superior properties for population level analysis . . . possess rather paradoxical characteristics from the point of view of individual measurement” (1992:159).

Item response theory is not a straightforward method of scoring (such as adding up correct responses and calculating the percentage correct). In fact, “it cannot be emphasized too strongly that plausible values are not test scores for individuals in the usual sense . . . which are, in some sense, optimal for each respondent . . . . Plausible values are constructed explicitly to provide consistent estimates of population effects, even though they are not generally unbiased estimates of the proficiencies of the individuals with whom they are associated” (Yamamoto and Kirsch 1997:180).

To estimate the probability that an individual will complete a task of given difficulty, the individual does not have to be observed performing that task. Indeed, because of the spiral sampling design, no IALS respondent was asked to respond to more than 45 questions, and some individuals actually responded to as few as five. Proficiency scores were imputed using background variables and observed responses. The background variables “included sex, ethnicity, language of interview, respondent education, parental education, occupation and reading practices, among others” (Yamamoto and Kirsch 1997:181). The specific plausible value for literacy attainment assigned to an individual depends on the critical probability level required for task success, the vector of conditioning characteristics used to estimate the probability of successful completion, and the specific statistical technique used in estimating that probability.[11] Literacy values are then assigned to individuals—including literacy levels both greater than, and lower than, the difficulty level represented by any of the individual questions.

This paper lays some stress on the issue of imputed values greater or less than the difficulty level of any question actually asked[12] because these scores are at the extremes of the distribution of literacy proficiency, and for some policy issues, it is the extremes which matter. Advocates for remedial literacy programs emphasize the exclusion of people with very low literacy from employment and normal written communication and social discourse—their emphasis on "exclusion" suggests that very low literacy attainment can have qualitatively different, and very significant, employment and social impacts, compared to mid-range differences (e.g., in reading speed). On the other hand, advocates of greater educational streaming and of elite programs for gifted children stress the importance of scientific and literacy excellence and of breakthrough discoveries—as indicated by the number of Nobel Prize winners or high-tech billionaires.

However, people at the extremes of literacy are rare[13] and it is difficult to design questions to accurately test for very high, or very low, skill levels. Literacy tests of the general population are best able to assess variation in the middle range of literacy skills, but this may not be the relevant range for some public policy debates. Section 3 of this paper asks whether the method of calculation of literacy scores "matters" or not, for assessment of the role literacy plays in earnings attainment *for the labour force as a whole*. However, in general, whether or not the scoring of literacy "matters" depends heavily on why we want to know about literacy.

Section 3.1

Rankings of average literacy

Item response theory provides estimates of population means, and when the media compare literacy levels in different jurisdictions, the most commonly used statistic is the rank order of average scores. Although the ranking of average outcomes may seem like a fairly dependable statistic, how robust are such rankings to the rescaling of alternative plausible values?

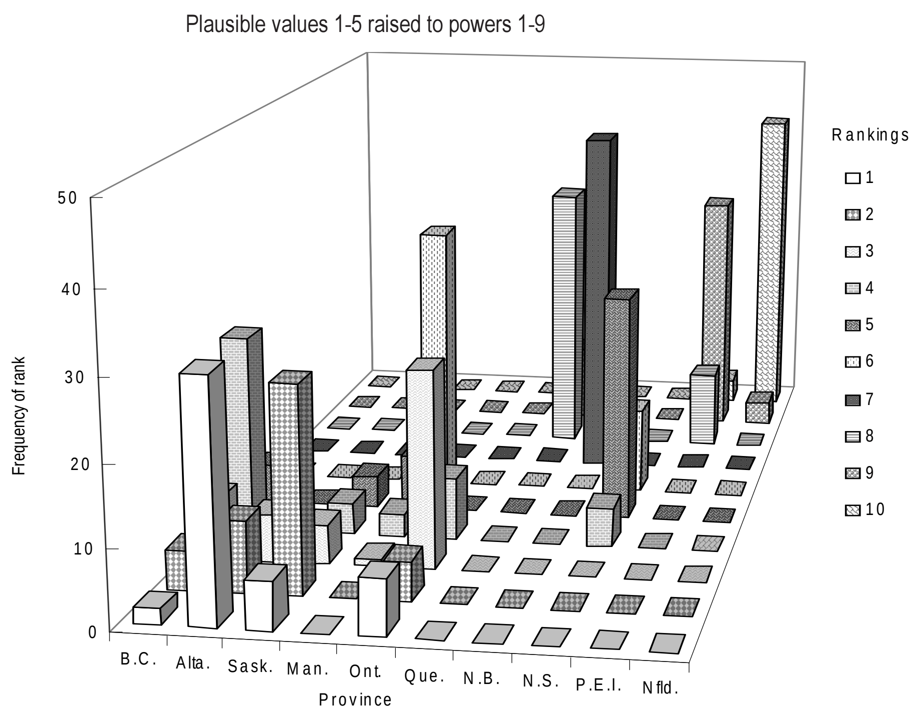
Figure 1 presents the frequency with which the different Canadian provinces occupy specific rankings, when average literacy $(\overline{L})$ is calculated according to each of the five plausible values, raised to the power 1....9.

$$\overline{L}_{jK} = \left(\sum_{i=1}^{N} L_{iK}^{J} \right) / N$$

K = 1....5 plausible values

j = 1....9

Figure 1 Frequency of provincial mean total literacy rank



There are 45 separate possible rankings from this procedure, and although a clear general tendency can be observed, there are also a significant number of reversals (for example, British Columbia is most often ranked fourth, but occasionally first or fifth).

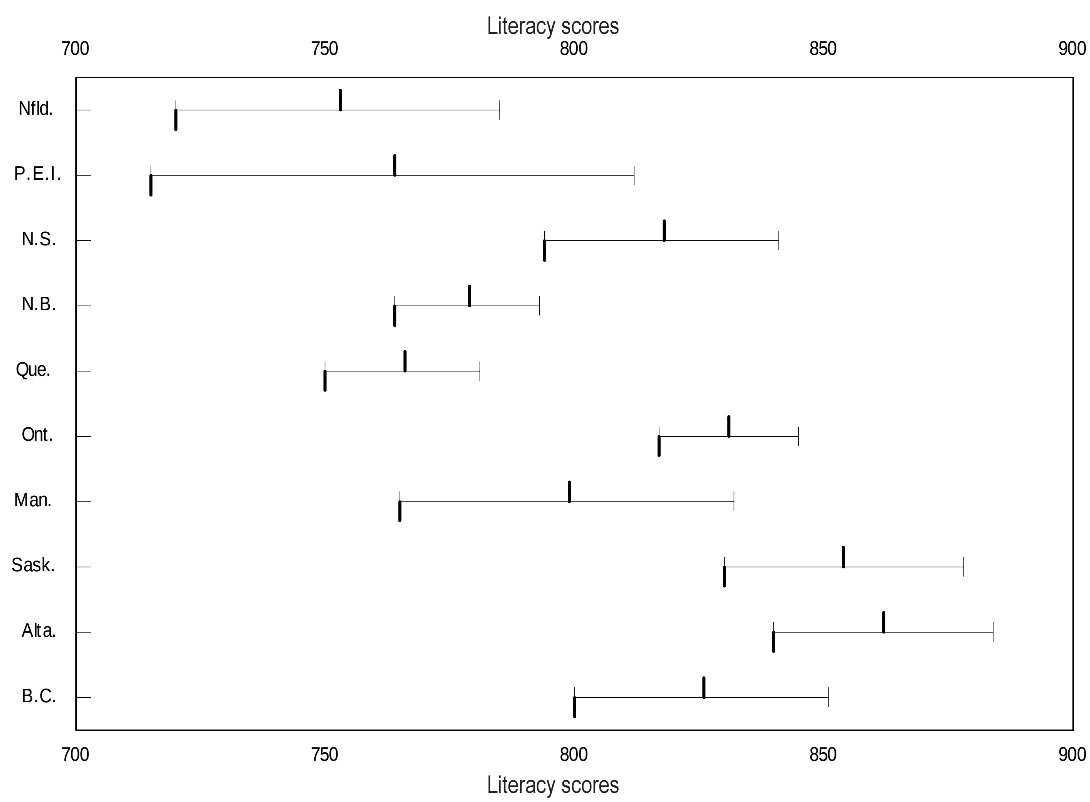
Figure 2 shows the range of estimates of average literacy obtained from these 45 alternative scalings of total literacy. Clearly, there is some important information in these averages—the east-to-west gradient in average literacy in the total population stands out. However, it is also clear that relative rankings within the east and the west are sensitive to the scaling of individual scores.

Average literacy rankings for the entire population are, however, very much a lagging indicator. People who left school 40 or even 50 years ago are mingled with those who have recently left school, and it takes decades for the impact of educational policy changes to appear in the overall average. Figure 3, therefore, presents the rankings of average literacy levels of those under 30 years old. It is notable that among Canada's youth there is not much indication of a west-to-east gradient.[14]

The moral is that great caution should be used in interpreting average literacy attainment. Based on comparison of the simple population averages of "Plausible Value 1," there might, for example, be a temptation to say that average literacy levels are lowest in Atlantic Canada.[15] However, Figure 3 indicates that when a range of monotonic transformations of plausible values is considered, Nova Scotia is frequently ranked as having the highest average literacy level in the country among those under 30 years of age. In general, it would appear desirable to be sure that rankings of average proficiency are not just artefacts of the scaling of individual scores, before policy conclusions are drawn.

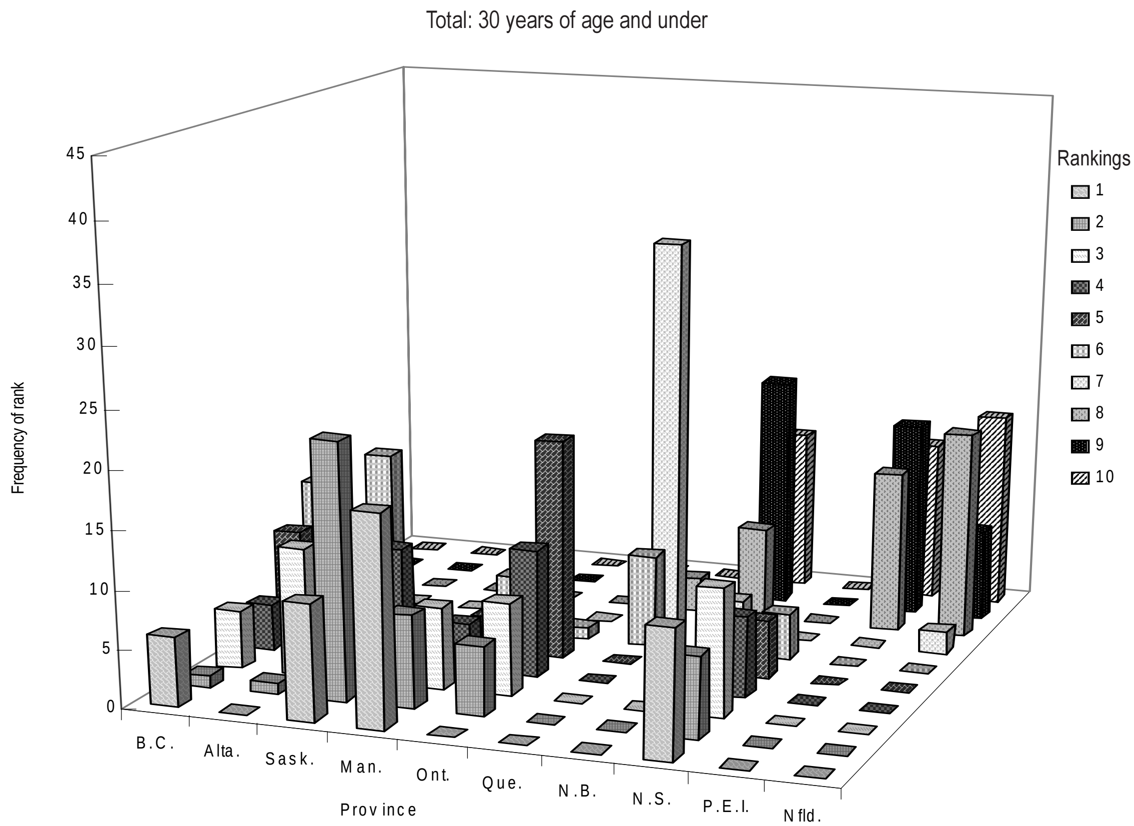
Figure 2 Mean total literacy

Showing all scores within two standard deviations of the mean



Total literacy is the sum of prose, document, and quantitative scores.
Range shown includes possible means for all five 'plausible values.'

Figure 3 Frequency of provincial mean total literacy rank, 30 years of age and under



Section 3.2

Literacy and earnings

Literacy is one of the major objectives of the educational system, and years of schooling have long been found to be a good predictor of an individual's earnings. How much of the individual return to education can be explained by an individual's level of literacy?

To assess this issue, one must choose a plausible measure of literacy. IALS data presents five plausible values for literacy scores. However, no actual questions were used with difficulty levels below 188 (prose), 182 (document) or 225 (quantitative), or above 377 (prose), 408 (document) or 408 (quantitative). Scores assigned above or below the skill levels actually tested are necessarily the result of imputation—and one might wonder whether such scores should be truncated, and how. We have argued that literacy scores are ordinal numbers, and alternative monotonic transformations are quite defensible (such as taking the natural logarithm of the score, or raising it to the power 1 to 9). It could also be argued that the information content of ordinal numbers lies in the rank assigned to individuals, so that the literacy score of an individual should be replaced by their sample literacy rank.

All these issues (truncation, power transformations, rank information) are measurement choices, but choices are not limited to the appropriate measure of literacy. Measuring education in years imposes an assumed linear impact of education, while the use of dummy variables for educational level would allow educational effects to be non-linear, and perhaps be better able to pick up the "sheep-skin" or credential effect of having actually completed a degree program. Some would also argue that estimation of literacy and education effects on earnings should be restricted to full-time, full-year workers (in order to have some control, implicitly, for hours of labour supply). However, if education or literacy enables individuals to gain access to employment, it could also be argued that the sample should be all workers.

Does any of this matter, and do we get essentially the same estimated impact of literacy, whichever measurement choice we make?

The strategy of this paper is to take the very simplest human capital earnings equation,[16] add alternative possible measures of literacy, and experiment with alternative specifications of the sample and measures of education. The result is a large number of estimated regressions (see Tables 1, 2 and 3).

Table 1 shows that literacy appears to account for about 30% of the returns of education for males who work full time throughout the year. Whichever way the literacy score is stretched for the full-time, full-year work force, it is always statistically significant. Only by strongly accentuating the relative importance of differentials at the top end of the literacy scale (by raising the literacy score to the 9th power), can one reduce the estimated impact of literacy on the return to schooling[17] to about one-sixth.

Table 1 Male full-time, full-year regressions

	OLS regression – dependent variable for all regressions in Ln(Earnings)						
	Base regression	Regressions including literacy					
Edyrs	0.043** (0.006)	0.030** (0.007)	0.031** (0.007)	0.031** (0.007)	0.033** (0.007)	0.030** (0.007)	0.036** (0.007)
Exp	0.075** (0.007)	0.075** (0.007)	0.075** (0.007)	0.075** (0.007)	0.075** (0.007)	0.076** (0.007)	0.076** (0.007)
Exp^2	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)
Constant	9.018** (0.116)	8.676** (0.171)	8.658** (0.155)	6.141** (0.785)	8.965** (0.117)	9.002** (0.115)	9.059** (0.116)
Litot1		0.001** (0.000)					
Litot1 ($\times 10^3$) - Truncated			0.585** (0.168)				
Lnlitot1				0.449** (0.121)			
Litrnk1 ($\times 10^4$)					0.517** (0.169)		
$(\text{Litot1})^3$ ($\times 10^9$)						0.242** (0.057)	
$(\text{Litot1})^9$ ($\times 10^{28}$)							0.418** (0.108)
Adjusted R^2	0.196	0.212	0.208	0.210	0.205	0.214	0.211

Notes: Values in parentheses are standard errors.
** = significant at the 5% level
Edyrs = Years of formal education (beginning with grade 1)
Exp = Age – Years of Education – 5
Litot1 = Total literacy using plausible value 1 = Prose1 + Document1 + Quantitative1
Litot1Truncated = Total literacy, measured with truncated literacy distributions and plausible value 1: all scores above maximum-valued question were replaced with the maximum question value and all scores below mimimum valued question were given the minimum value.
Lnlitot1 = Natural log of Litot1
Litrnk1 = Rank in total sample ordered by Litot1
($\times 10^?$) = Value of coefficient and standard error have been multiplied for reporting by 10 raised to the given power

Table 2 examines whether measuring education by credentials obtained alters the conclusion that literacy skills explain a significant fraction of the return yielded by education, for males working full time throughout the year. In some cases, the impact of literacy skills appears greater. Comparing columns 5 and 6 with column 4, or columns 8 and 9 with column 7, it appears that including an explicit control for measured literacy skills reduces by 40% to 45% the estimated return to a university education. Although the impact of including measured literacy on the disadvantage associated with very low education is less (a change from –.412 to –.342 or –.306, or a 16% to 26% decline). Table 2 still indicates that much of the measured return yielded by education is due to literacy skills.

Table 3 provides a cautionary note. For female full-time and part-time workers, literacy scores are always statistically insignificant. Indeed, including literacy scores in an earnings regression for women sometimes produces implausible results. It is, for example, possible to find a rescaling of individual literacy scores such that the estimated return of education, net of literacy, rises when literacy scores are included as an explanatory variable. Clearly, the impact of literacy on women's earnings is different than the impact of literacy on men's earnings.

Over the range of measurement choices considered, literacy skills generally explain a significant part of the returns from education—but not always. This suggests that one could turn the question around and ask: what is the maximum fraction of the return of education that can be explained by the inclusion of measured literacy skills? Figures 4 and 5 illustrate the range of estimated returns to years of education obtained when various measurement choices about literacy scoring are made, compared with the baseline estimate of the return to years of education obtained when literacy scores are not considered.

We can be fairly certain that increased literacy is only part of the reason why an education pays off. Figures 4 and 5 illustrate the uncertainty about how much of the financial return of years of education can be explained by the possession of literacy skills—but whatever rescaling of literacy scores is done, it is very hard to push the contribution of literacy above 40% to 45% of the returns yielded by education.

Of course, educators have always aimed at teaching more than literacy skills. Factual knowledge, reasoning and social skills have also long been goals of education. These other outcomes are not captured in measured literacy. It has also long been argued (for example, by Arrow 1973 and Spence 1973, 1974) that education serves as a credential that signals underlying native ability. This paper cannot assess which of the human capital or credentialist non-literacy functions of education are of greatest importance, but does provide some indication of the potential range of literacy impacts.

Table 2 **Regressions including education dummy variables**

	Male – full-time, full-year regressions								
	No education dummy variables			Education dummy variables			Dummy variables for postsec. only		
	Base	With literacy	With $(\text{Lit.})^3$	Base	With literacy	With $(\text{Lit.})^3$	Base	With literacy	With $(\text{Lit.})^3$
Edyrs	0.043** (0.006)	0.030** (0.007)	0.030** (0.007)						
Exp	0.075** (0.007)	0.075** (0.007)	0.076** (0.007)	0.078** (0.007)	0.078** (0.007)	0.078** (0.007)	0.079** (0.007)	0.078** (0.007)	0.078** (0.007)
Exp^2	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)	- 0.001** (0.000)
Constant	9.018** (0.116)	8.676** (0.171)	9.002** (0.115)	9.607** (0.075)	9.134** (0.154)	9.444** (0.089)	9.549** (0.076)	8.954** (0.147)	9.364** (0.085)
Litot1		0.001** (0.000)			0.001** (0.000)			0.001** (0.000)	
$(\text{Litot1})^3$ $(\times 10^9)$			0.242** (0.057)			0.230** (0.060)			0.275** (0.060)
Dummy=1 if University				0.227** (0.053)	0.138** (0.058)	0.124** (0.059)	0.305** (0.051)	0.174** (0.057)	0.170** (0.058)
Dummy=1 if Postsecondary				- 0.162** (0.055)	- 0.188** (0.056)	- 0.187** (0.055)	- 0.076** (0.053)	- 0.128** (0.054)	- 0.118** (0.053)
Dummy= 1 if Some high school				- 0.229** (0.061)	- 0.202** (0.061)	- 0.209** (0.061)			
Dummy= 1 if Primary				- 0.412** (0.100)	- 0.306** (0.104)	- 0.342** (0.101)			
Adjusted R^2	0.196	0.212	0.454	0.221	0.232	0.235	0.196	0.218	0.251

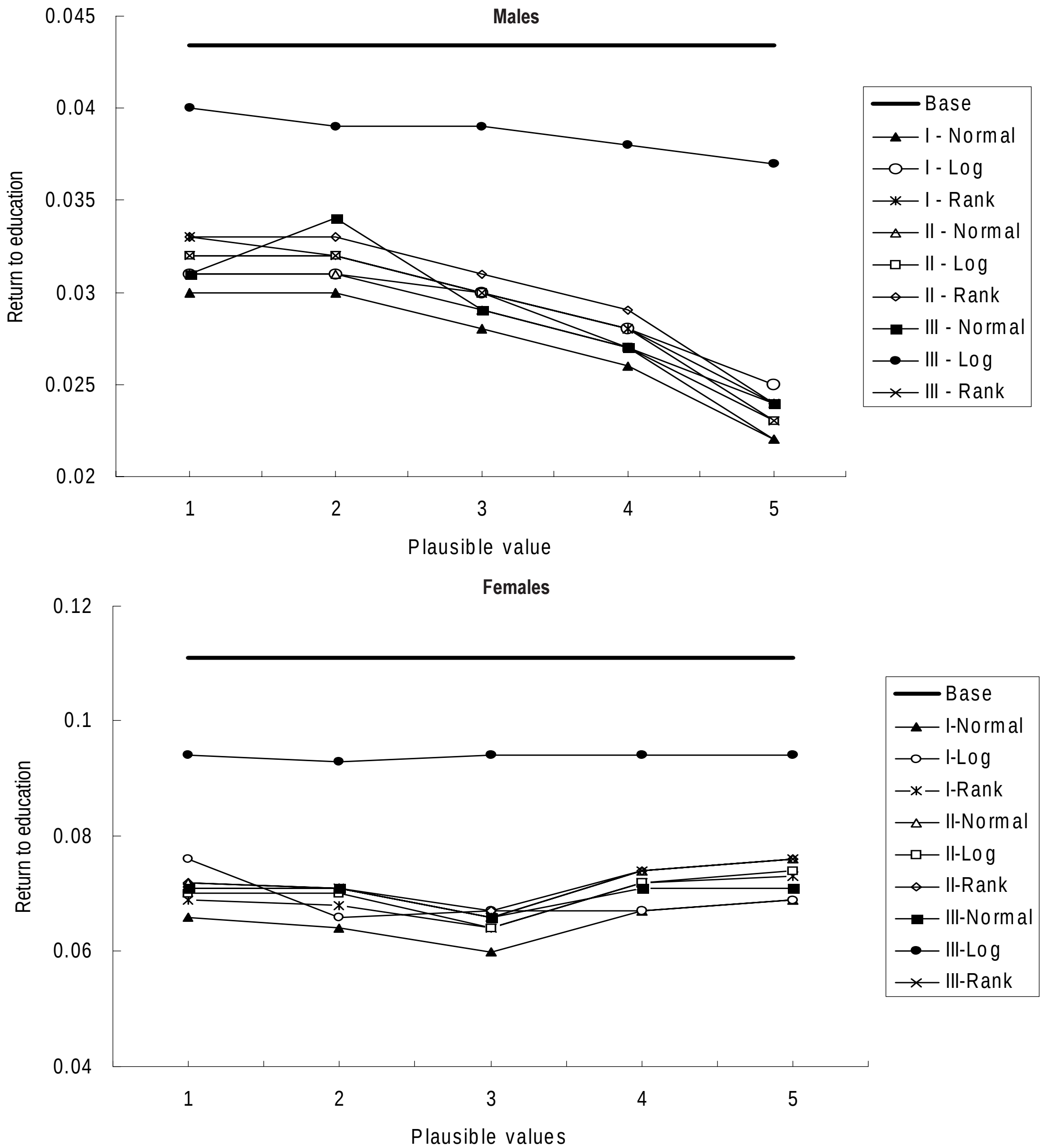
Notes: Values in parentheses are standard errors.
** = significant at the 5% level
Edyrs = Years of formal education (beginning with grade 1)
Exp=Age – Years of Education – 5
Litot1=Total literacy using plausible value 1 = Prose1 + Document1 + Quantitative
$(\times 10^3)$ = Value of coefficient and standard error have been multiplied for reporting by 1000
Education dummy variables correspond to highest level of schooling completed as follows:
University = Completed university
Postsecondary = Completed non-university post-secondary education
Some high school = Completed some secondary education
Primary = Includes both individuals who did and did not complete primary education;
The base case is completed secondary education; Individuals whose highest level of schooling completed was not stated or not definable were excluded from these regressions.

Table 3 **Female full-time and part-time regressions (includes all females who reported positive earnings)**

	OLS regressions – dependent variable for all regressions is Ln(Earnings)						
	Base reg'n	Regressions including literacy					
Edyrs	0.090** (0.008)	0.086 ** (0.010)	0.091** (0.010)	0.083** (0.010)	0.087** (0.010)	0.091** (0.010)	0.094** (0.009)
Exp	0.076** (0.008)	0.076 ** (0.008)	0.076** (0.008)	0.076** (0.008)	0.076** (0.008)	0.076** (0.008)	0.076** (0.008)
Exp^2	-0.002** (0.000)	-0.002 ** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)	-0.002** (0.000)
Ln(weeks)	0.700** (0.041)	0.699 ** (0.041)	0.700** (0.041)	0.699** (0.041)	0.698** (0.041)	0.700** (0.041)	0.699** (0.041)
Constant	5.048** (0.192)	4.993 ** (0.210)	5.077** (0.226)	4.073** (0.781)	5.049** (0.192)	5.046** (0.192)	5.024** (0.193)
Litot1 ($\times 10^3$)		0.123 (0.193)					
Litot1 ($\times 10^3$) - Truncated			-0.055 (0.226)				
Lnlitot1				0.159 (0.123)			
Litrnk1 ($\times 10^4$)					0.095 (0.220)		
$(\text{Litot1})^3$ ($\times 10^9$)						-0.023 (0.088)	
$(\text{Litot1})^9$ ($\times 10^{28}$)							-0.303 (0.270)
Adjusted R^2	0.353	0.352	0.352	0.353	0.352	0.352	0.353

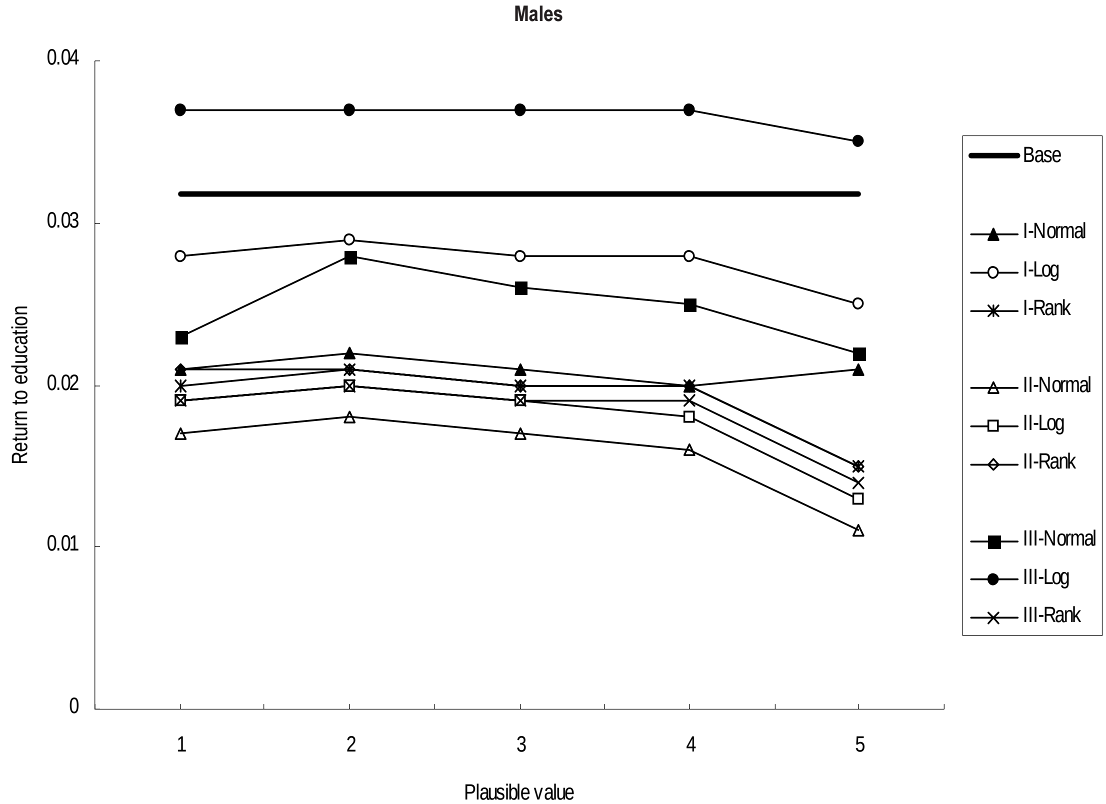
Notes: Values in parentheses are standard errors.
** = significant at the 5% level
Edyrs = Years of formal education (beginning with grade 1)
Exp = Age - Years of Education - 5
Ln(weeks) = The natural log of weeks worked in the past year
Litot1 = Total literacy using plausible value 1 = Prose1 + Document1 + Quantitative1
Litot1Truncated = Total literacy, measured with truncated literacy distributions: all scores above maximum valued question were replaced with the maximum value and all scores below mimimum valued question were given the minimum value.
Lnlitot1 = Natural log of Litot1
Litrnk1 = Rank in total sample ordered by Litot1
($\times 10^?$) = Value of coefficient and standard error have been multiplied for reporting by 10 raised to the given power

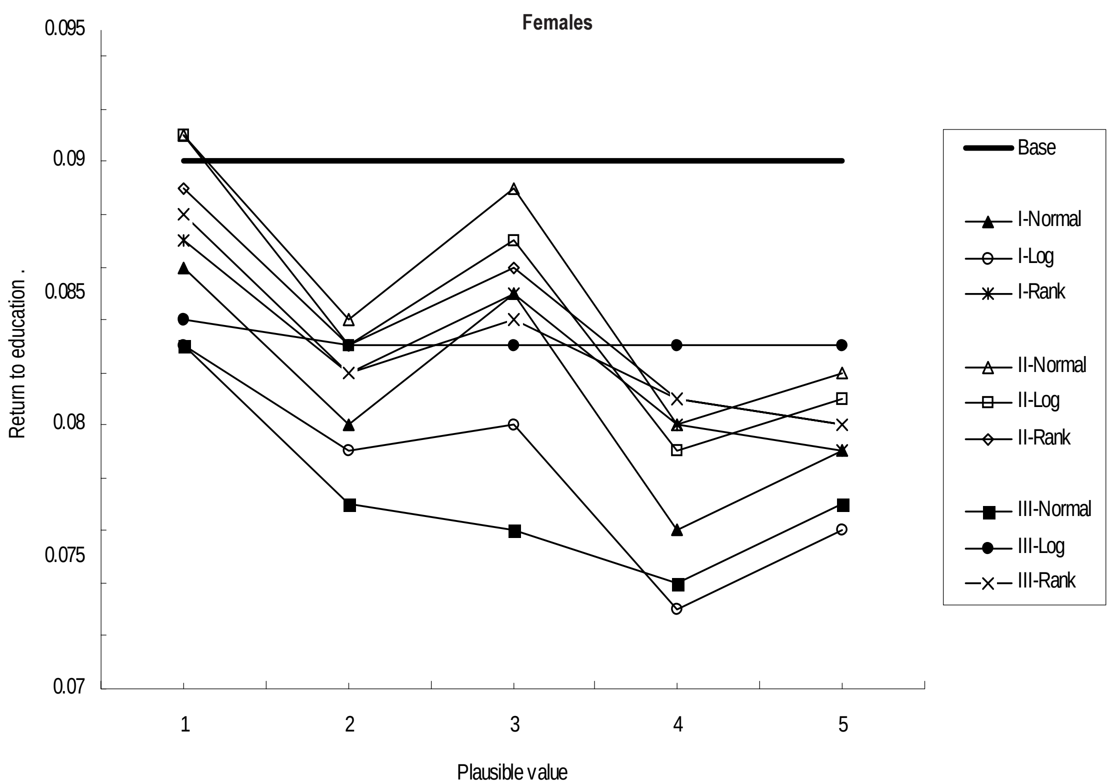
Figure 4 Rate of return to education, males and females 20–65, full-time, full-year



I = Normal
II = Truncated: top pushed down, bottom pushed up
III = Truncated, top pushed down, bottom add to zero

Figure 5 **Rate of return to education, males and females 20–65, full-time and part-time**





I = Normal
II = Truncated: top pushed down, bottom pushed up
III = Truncated, top pushed down, bottom add to zero

Section 3.3

Meta-analysis

When a large number of possible combinations of measurement choices exist, meta-analysis can help detect patterns in the implications of measurement choices. Table 4 presents the implications of measurement choices for the estimated rate of return of years of education. The results are based on OLS regressions. The dependent variable is the rate of return of years of education and the independent variables summarize the characteristics of the regression that produced that estimated rate of return. In column 3, for example, one can interpret the coefficient .729 on the dummy variable for Plausible Value 1 as indicating that relative to the base case (Plausible Value 5), using Plausible Value 1 increases the estimated rate of return per year of education by 0.729%.

The first column of Table 4 shows the results when all estimated regressions are in the sample. The base case is the earnings equation for all men when no literacy variable is included. Columns 2 to 4 restrict attention to regressions including literacy variables. Relative to Plausible Value 5, the use of other plausible values produces higher estimates of the rate of return to education, except for women who work full time throughout the year.

Raising the literacy score to successively higher powers is a monotonic transformation which increasingly emphasizes the importance of differences at the top end of the literacy distribution. For full-time, full-year males, magnifying the importance of literacy differentials at the top end of the literacy distribution has a statistically significant effect on the measured return to education, but the effect is non-linear. In the meta-analysis, entering the power to which literacy scores are raised, and the square of that power, tests for non-linearities, and column 3 of Table 4 can be interpreted to mean that the estimated rate of return to education is minimized when the square of the literacy score is entered as an explanatory variable,[18] alongside years of education.

Conversely, a logarithmic transform of literacy scores compresses the influence of differentials at the top end of the literacy distribution, and that appears to inflate the measured influence of years of schooling on earnings. In addition to the fact that the specific plausible value used typically matters for the estimated rate of return to education (conditional on literacy), it also matters whether the scaling of literacy emphasizes differences at the top end of the distribution, relative to the bottom.

It also seems to matter how one treats very low literacy scores. The dummy truncation 1 indicates the measurement choice that all literacy scores above or below the maximum and minimum difficulty actually asked are set to those maximum and minimum levels. Truncation 2, on the other hand, sets to 0 the literacy score of any person assigned a score less than the least difficult question asked. This implicitly accentuates measured literacy differences in the bottom tail, and such a choice measurably improves the estimated return to education.

Table 4 Meta-analysis results

	OLS regressions – estimated return to education is dependent variable for each regression			
	All regressions	Regressions including literacy only		
Sex:	Males and females	Males and females	Males	Females
Labour Status:	Full-time and part-time	Full-time and part-time	Full-time, full-year only	Full-time, full-year only
Observations:	284	280	70	70
Constant	4.656**	2.419**	2.368**	6.839**
	-0.289	-0.169	-0.059	-0.110
D-plaus.val=1	-1.804**	0.366**	0.729**	-0.236**
	-0.327	-0.104	-0.038	-0.071
D-plaus.val=2	-1.990**	0.180 *	0.714**	-0.443**
	-0.327	-0.104	-0.038	-0.071
D-plaus.val=3	-2.022**	0.148	0.557**	-0.793**
	-0.327	-0.104	-0.038	-0.071
D-plaus.val=4	-2.115**	0.055	0.386**	-0.214**
	-0.327	-0.104	-0.038	-0.071
D-plaus.val=5	-2.170**			
	-0.327			
D-female	4.447**	4.414**		
	-0.095	-0.093		
D-fem.&pooled	0.966**	1.010**		
	-0.095	-0.093		
D-male&pooled	-0.866**	-0.861**		
	-0.095	-0.093		
power	-0.059	-0.059	-0.123**	0.115
	-0.143	-0.139	-0.051	-0.095
power^2	0.039	0.039	0.031**	0.062**
	-0.026	-0.025	-0.009	-0.017
D-log	0.481**	0.481**	0.111**	0.096
	-0.095	-0.092	-0.039	-0.074
D-rank	0.053	0.053	0.153**	0.197**
	-0.095	-0.092	-0.034	-0.064
D-truncation1	0.013	0.013	0.053	0.350**
	-0.095	-0.092	-0.034	-0.064
D-truncation2	0.416**	0.416**	0.071*	0.303**
	-0.095	-0.092	-0.039	-0.074
D-trunc.2 & log			0.925	2.302**
			-0.067	-0.125

Notes: ** = significant at the 5% level
* = significant at the 10% level

Section 4

Conclusion

The development of direct measures of skill attainment, such as the IALS data, offers labour economists a powerful new tool to help explain labour market outcomes. There is a great deal of useful information in such test scores—as this paper has demonstrated, the literacy test scores of men have a statistically and empirically significant relationship with individual earnings, and that effect is robust to a large variety of measurement choices.

Nevertheless, this paper has also emphasized that some caution is in order in the use of direct measures of skill attainment in statistical analysis. Labour economists have developed over the years many complex statistical techniques for working with data, but the underlying concepts have typically been clearly observable magnitudes (such as number of children, hourly wage or marital status), which can be measured either as cardinal numbers, or as discrete states. Literacy, and skill attainment more generally, is not like that—literacy is a complex concept, for which there is no natural unit of measurement. Although direct measures of literacy proficiency, such as the IALS, can rank individuals in literacy attainment, literacy scores are the product of complex statistical procedures, which involve many of the same variables (such as education or age) that labour economists would usually expect also to play an independent role in determining labour market outcomes. Literacy scores are also inherently ordinal numbers, and a variety of monotonic transformations of those scores may be equally plausible.

The method of calculating literacy scores may therefore matter considerably for the perceived impact of literacy on labour market outcomes. More generally, many public services have a "quality" dimension that is similarly complex, and similarily inherently ordinal. Hence, the issue of how best to measure the impact of literacy on individual earnings may be an example of a more general problem—that the method of calculation of "quality" measures of public sector outcomes may be central to the perceived "success" of public policies.

This paper has attempted to make, by example, the methodological point that research using direct measures of skill attainment should test for the robustness of statistical results by examining a variety of monotonic transformations of skill attainment. When this is done with literacy scores, one observes that rankings of Canadian provinces in average literacy attainment may change, sometimes quite dramatically. Assessment of the relative success, or failure, of public sector policies (such as education) by the criterion of average literacy levels in different jurisdictions may therefore be excessively dependent on measurement choices and scaling assumptions.

Particularly for males, it is clear that, whatever the transformation of literacy scores, much of the return to education is due to a return to literacy skills—perhaps as much as 40% to 45%, although the exact proportion of the return to education that can be accounted for by literacy skills depends somewhat on measurement choices and scaling assumptions. The rate of return to education for women is both higher and less influenced by literacy proficiency.

There is also some suggestion in the data that the relative influence of literacy may vary at different points in the distribution of literacy attainment—specifically, that literacy may account for a higher proportion of the impact of education on earnings among those with high literacy skills. The differing roles played by literacy proficiency among men and women, and at different points in the distribution of literacy, remain important issues for future research.

Endnotes

1. A "monotonic transformation" is a transformation that preserves the rank order of the initial variables.

2. Some reviewers reject this latter finding, arguing that some of the tested monotonic transformations are implausible.

3. If average student attainment is used to determine incentive pay, or otherwise allocate resources, grade weights acquire the role of shadow prices. If so, they can be expected to influence the effort that teachers place on improving the achievement of marginal students, compared with the effort that they devote to assisting top students. If the scaling is A=1000, B=100, C=10, and if the average score influences teachers' pay, one can expect most teachers to pay a lot more attention to potential A students than to C students. Whether or not it is socially desirable to concentrate on moving B students into the A category, compared with preventing C students from slipping to D or F, the weights assigned to achievement represent an incentive system.

4. As Cliff (1996:92) notes, in the standard OLS regression model Y = XB + E, "Monotonically transforming X changes not only its covariance with Y but its covariance with the other predictors as well, and such transformation is likely to do so in unpredictable or irregular ways. This will alter its coefficient b, and affect ϕ (the sum of squared residuals) and X's contribution to its reduction. Thus, results are not invariant if the variables are transformed."

5. I.Q. scores have, for example, long been used as an explanatory variable in earnings regressions (e.g., Taubman 1973, 1976). Even if it is not known what the shape of the distribution of intelligence is, it is known that I.Q. scores are scaled so as to follow a normal distribution.

6. Adding an arbitrary constant—for example, 1 billion—to the I^{th} observation's score, and to all higher scores, would preserve the ordering of observations, but would clearly (depending on the value of I) dominate any multiple regression results.

7. For example, we may be unsure as to how a score at the top end of the literacy distribution compares with a score in the middle of the distribution—if the median individual scores 250 and the 95^{th} percentile individual scores 400, it would be placing too much credence in a particular scaling of scores to say that the latter individual has 60% more literacy than the former. One could scale the same test to produce scores of 100 and 900, which might also be thought "reasonable." However, a rescaling which produced the values of 100 and 100,000,000,000 might be thought by most observers to be "unreasonable."

8. Level 1 includes scores between 0 and 225; level 2 includes scores between 226 and 275; level 3 includes scores between 276 and 325; level 4 includes scores between 326 and 375; and level 5 includes scores between 376 and 500. In some cases a broad distinction is drawn between adults with literacy levels above the basic level (that is, levels 1 and 2) and those with only level 1or level 2 literacy (Murray 1995).

9. Since the upper bound for level 1 numeracy was also 225, this implies that only one questionnaire item could actually distinguish between level 1 numeracy and level 2 or higher.

10. Questionnaire blocks were assigned randomly into respondents in a "spiral sampling" design. One implication is that only some respondents actually answered the specific questions testing the lowest, and highest, difficulty ratings.

11. Gordon, Lin, Osberg and Phipps (1994) have noted that estimation using logit or probit analysis can imply quite different estimated marginal effects on probabilities of outcomes for individuals with the mean characteristics of the sample, and that sampling variability can be an important issue in samples of 10,000 or less. (Note that all individual test items are asked of (3/7 x 5,660) 2,425 or fewer respondents.)

12. In addition, the alternative scalings of literacy scores discussed below include some which accentuate significantly the relative weight of differences at the extremes of the literacy distribution.

13. One way of testing for the importance of very high level skills is to magnify the impact of differentials in the top end of the literacy distribution, by raising individual literacy scores to successively higher powers.

14. The change in Nova Scotia's average literacy ranking is particularly notable, and has a straightforward explanation—a dramatic increase in school retention (in 1965, the Nova Scotia grade 12 retention rate of the students in Grade 7, five years earlier, was only 33%. By 1992, it had increased to 94%). It is less clear why New Brunswick and Newfoundland, which also had dramatic increases in school retention, continue to trail.

15. For additional discussion of the IALS methodology, see OECD and Statistics Canada 1995; Statistics Canada and Human Resources Development Canada 1996.

16. Clearly, in focusing on a very simple human capital model, this paper is abstracting from the impacts of unionization, industry and occupation of employment, firm size, province of residence or labour market segment—not to mention any compensating differentials due to fringe benefits, workplace hazards, etc. Many studies have discussed these (and other) potential explanatory variables, most of which are unavailable in IALS. Because there is no clear consensus on the complete earnings function, this paper adopts the strategy of simplicity.

17. In column 7 of Table 1, the return to a year of schooling is 3.6%, compared to 4.3% when literacy is not considered—a drop of one-sixth.

18. Interestingly, explained variance (R^2) in the earnings regressions is highest when literacy scores are raised to the third power, and very nearly as high when they are squared.

References

Arrow, Kenneth J. (1973). "Higher Education as a Filter." *Journal of Public Economics*, Vol. 2(3), July: 193–216.

Bradley, R.A., S.K. Katti and Irma J. Coons (1962). "Optimal Scaling for Ordered Categories." *Psychometrika*, Vol. 27(4), December: 355–374.

Borgatta, Edgar F. (1968) "My Student, The Purist: A Lament." *The Sociological Quarterly*, Vol. 9(1), Winter: 29–34.

Cliff, Norman (1989). "Ordinal Consistency and Ordinal True Scores." *Psychometrika*, Vol. 54(1), March: 75–91.

Cliff, Norman (1993). "Dominance Statistics: Ordinal Analyses to Answer Ordinal Questions." *Psychological Bulletin*, Vol. 114(3), November: 494–509.

Cliff, Norman (1996). *Ordinal Methods for Behavioral Data Analysis*. Mahwah, New Jersey: Lawrence Erlbaum Assoc.

Grether, David M. (1974). "Correlations with Ordinal Data." *Journal of Econometrics*, Vol. 2(3), September: 241–246.

Gordon, D.V., Z. Lin, L. Osberg and S. Phipps (1994). "Predicting Probabilities: Inherent and Sampling Variability in the Estimation of Discrete-Choice Models." *Oxford Bulletin of Economics and Statistics*, Vol. 56(1), February: 13–32.

Kim, Jae-On (1975). "Multivariate Analysis of Ordinal Variables." *American Journal of Sociology*, Vol. 81(2), September: 261–298.

Kim, Jae-On (1978). "Multivariate Analysis of Ordinal Variables Revisited." *American Journal of Sociology*, Vol. 84(2), September: 448–456.

Kim, Suh-Ryung (1990). "On the Possible Scientific Laws." *Mathematical Social Sciences*, Vol. 20(1): 19–36.

Kirsch, Irwin (1995). "Literacy Performance on Three Scales: Definitions and Results." Chapter 2 in *Literacy, Economy and Society*. Statistics Canada Catalogue no. 89-545-XPE. Paris: OECD, and Ottawa: Statistics Canada.

Kirsch, Irwin (1996). "Literacy Performance on Three Scales: Definitions and Results." Appendix B in *Reading the Future: A Portrait of Literacy in Canada.* Catalogue no. 89-551-XPE. Ottawa: Statistics Canada.

Labovitz, Sanford (1967). "Some Observations on Measurement and Statistics." *Social Forces*, Vol. 46(2), December: 151–160.

Maddala, G.S. (1983). Limited-Dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press.

McCullagh, Peter (1980). "Regression Models for Ordinal Data." *Journal of the Royal Statistical Society*, Series B, Vol. 42(2): 109–142.

Mislevy, Robert J., A.E. Beaton, B. Kaplan and K.M. Sheehan (1992). "Estimating Population Characteristics From Sparse Matrix Samples of Item Responses." *Journal of Educational Measurement*, Vol. 29(2), Summer: 133–161.

Murray, T. Scott (1995). "International Adult Literacy Survey: Insights and Prospects." Chapter 5 in *Literacy, Economy and Society*. Statistics Canada Catalogue no. 89-545-XPE: Paris: OECD, and Ottawa: Statistics Canada.

National Centre for Education Statistics (1997). *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*, NCES 98-053. Washington DC: United States Department of Education.

O'Brien, Robert (1982). "Using Rank-Order Measures to Represent Continuous Variables." *Social Forces*, (61, 1, 144-155).

OECD and Statistics Canada (1995). *Literacy, Economy and Society: Results of the First International Adult Literacy Survey*. Statistics Canada Catalogue no. 89-545-XPE: Paris: OECD, and Ottawa: Statistics Canada.

OECD and Human Resources Development Canada (1997). *Literacy Skills for the Knowledge Society: Further Results from the International Adult Literacy Survey.* Paris: OECD, and Ottawa: Human Resource Development Canada.

Smith, Robert B. (1978). "Nonparametric Path Analysis: Comments on Kim's 'Multivariate Analysis of Ordinal Variables'." *American Journal of Sociology*, Vol. 84(2), September: 437–448.

Spence, A. Michael (1973). "Job Market Signaling," *Quarterly Journal of Economics*, Vol. 87(3), August: 355–374.

Spence, A. Michael (1974). *Market Signaling: informational transfer in hiring and related screening processes*. Cambridge: Harvard University Press.

Taubman, Paul J. and Terence J. Wales (1973). "Higher Education, Mental Ability and Screening," *Journal of Political Economy*, Vol. 81(1), January/February: 28–55.

Taubman, Paul (1976). "The Determinants of Earnings: Genetics, Family and Other Environments. A Study of White Male Twins." *The American Economic Review*, Vol. 66(5), December: 858–870.

Wilson, Thomas P. (1971). "Critique of Ordinal Variables." *Social Forces*, Vol. 49: 432–444.

Yamamoto, Kentaro and Irwin S. Kirsch (1997). "Proficiency Estimation." In National Centre for Education Statistics, *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*, NCES 98-053. Washington DC: United States Department of Education.