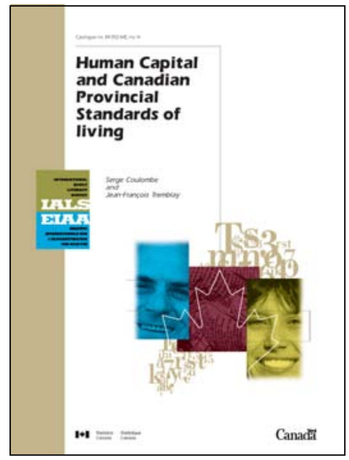
Catalogue no. 89-552-MIE, no. 14

International Adult Literacy Survey

Human Capital and Canadian Provincial Standards of Living

Serge Coulombe and Jean-François Tremblay



Statistics Canada Statistique Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Client Services, Culture, Tourism and the Centre for Education Statistics, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: (613) 951-7608; toll free at 1 800 307-3382; by fax at (613) 951-9040; or e-mail: educationstats@statcan.ca).

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 website.

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
Website	www.statcan.ca

Information to access the product

This product, catalogue no. 89-552-MIE, is available for free. To obtain a single issue, visit our website at www.statcan.ca and select Our Products and Services.

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 that 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. The service standards are also published on www.statcan.ca under About Statistics Canada > Providing services to Canadians.

International Adult Literacy Survey

Human Capital and Canadian Provincial Standards of Living

Serge Coulombe and Jean-François Tremblay
Department of Economics, University of Ottawa

Published by authority of the Minister responsible for Statistics Canada

April 2006

Catalogue no. 89-552-MIE, no. 14

ISSN 1480-9516

ISBN 0-662-43122-7

Frequency: Irregular

Ottawa

Cette publication est disponible en français (N° 89-552-MIF au catalogue, n° 14)

Statistics Canada

Library and Archives Canada Cataloguing in Publication
Data

Coulombe, Serge.
Human capital and Canadian provincial standards of living.

(International Adult Literacy Survey)
Issued also in French under title: Le capital humain et les niveaux de vie dans les provinces
Canadiennes.
ISBN 0-662-43122-7 (Internet)
CS89-552-MIE no. 14

1. Human capital – Canada – Provinces – Econometric models.
2. Income distribution – Canada – Provinces – Econometric models.
3. Literacy – Canada – Provinces – Statistics. 4. College graduates – Canada –
Provinces – Statistics. 5. Wages – Effect of education on – Canada – Provinces – Statistics.
I. Tremblay, Jean-François.II. Statistics Canada. III. Canada. Human Resources Development
Canada. IV. Title: Human capital and Canadian provincial standards of living. V. Series.

HD4907.4 C68 2006 339.20971
C2006-988026-3

Acknowledgements

We are grateful to Scott Murray from Statistics Canada for his encouragement to work on this research project and for his numerous insights throughout. Earlier versions of the paper have benefited from comments by Angel de la Fuente, Bob Fay, Seamus McGuinness, Benoit Robidoux, and the economists of the Economic and Fiscal Policy Branch from Finance Canada. We thank Statistics Canada for its financial support and for its help in retrieving and aggregating data. We are grateful to Han You for her very valuable research assistance and to Patricia Buchanan for the English editing. The views expressed in this paper are those of the authors.

Note of Appreciation

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

Table of contents

Table of contents

Summary

This paper examines the role of human capital accumulation in explaining the relative levels of income per capita across Canadian provinces. We use principally two different types of human capital indicators based respectively on university attainment and literacy test scores. A synthetic time series of the average literacy level of labour market entrants for each period between 1951 and 2001 is constructed from the demographic profile of literacy test scores taken from the 2003 Adult Literacy and Lifeskills Survey. The percentage of the working-age population holding a university degree is available since 1951 from the census figures. Our main results are the following. First, both human capital indicators are strong predictors of the relative levels of per capita income (minus government transfers) across provinces, along with the relative rates of urbanization and specific shocks in Alberta and Quebec. Second, the skills acquired by one extra year of schooling result in an increase in per capita income of around 7.3 percent. Third, we find that our literacy indicator does not outperform the university attainment indicator. This contrasts sharply with our recent result found at the cross-country level (Coulombe, Tremblay, and Marchand [2004]) and suggests substantial measurement error in cross-country schooling data. Fourth, by focusing on regional economies that have similar levels of social infrastructure and social development, our analysis provides potentially more reliable estimates of the contribution of human capital accumulation to relative living standards.

Highlights

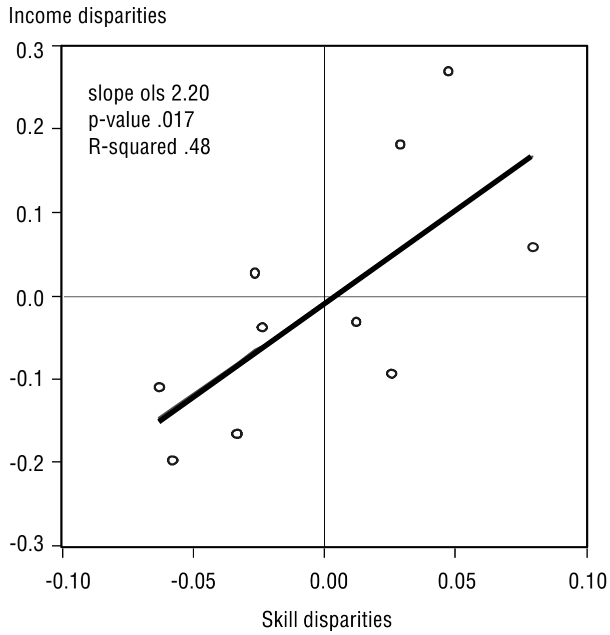
This paper examines the role of human capital accumulation in explaining relative levels of income per capita, excluding government transfers to individuals, across Canadian provinces between 1951 and 2001. The empirical analysis improves our understanding of Canadian provincial disparities and makes a contribution to the measurement of human capital by using different indicators, based on university attainment, literacy test scores, and years of schooling. Among other things, we find that both literacy and university attainment indicators are strong predictors of the relative levels of per capita income and that the macroeconomic returns from human capital are substantial. By focusing on the Canadian provinces whose social infrastructure is relatively similar, our analysis also allows us to disentangle the contribution of human capital from that of institutional improvements or investment in social infrastructure, both of which have received considerable attention in recent cross-country growth literature.

The underlying framework

The theoretical framework underlying the empirical analysis is that of an open economy with perfect physical capital mobility, which captures fairly well the reality of Canadian provinces. In small open economies with perfect capital mobility for the financing of investment in physical capital but a binding constraint for the financing of human capital investment, economic theory predicts that human capital accumulation is the main determinant of per capita income (Barro, Mankiw, and Sala-i-Martin 1995). Because of the complementarity between physical and human capital in the production process, the rate of return on physical capital will be determined by the ratio of human capital to labour. Therefore, even though capital can flow freely across economies in response to small differences in the rate return, the physical capital to labour ratio may not be equalized across economies. Thus the accumulation of human capital will be the main determinant of physical capital accumulation and of per capita income.

The scatter diagram on the next page provides a first rough look at the relationship between income per capita and human capital across provinces, measured by our indicator based on literacy test scores. This figure shows a clear positive relationship between income per capita and skills in each province, both measured relative to the cross-sectional mean. Moreover, the result of a simple regression shows that skills disparities can potentially explain up to 48 percent of income disparities. Of course, the slope coefficient will be biased as long as human capital is correlated with important variables omitted in this simple regression. In particular, our empirical analysis shows that the relative rate of provinces' urbanization, which is positively correlated with literacy scores, has a positive and significant effect of income per capita. As a result, the omission of urbanization in the simple regression presented in the next figure leads to an upward bias in the estimated effect of literacy skills. Although admittedly too simple, this regression nonetheless provides support to the theoretical framework underlying the more sophisticated empirical analysis presented in the paper.

Skill and income disparities



Notes: Skills of non-migrants and personal income (minus transfers to individuals). Ten provinces, 2003. Logarithm deviations from the cross-sectional sample mean.

The data

Three different types of human capital indicators are used. The first is the percentage of the working-age population who hold a university degree, information that is available from the census on a consistent basis since 1951. The second indicator, following a methodology proposed by Coulombe, Tremblay, and Marchand (2004), is a synthetic time series of human capital accumulation constructed from the demographic profile of literacy tests scores provided by the 2003 Adult Literacy and Lifeskills Survey (ALL). Specifically, as a measure of a province's relative investment in human capital in each 5-year period (1951–1956, …, 1996–2001), we use the average literacy score of the cohort of individuals who were entering the labour market during that period relative to the cross-sectional mean. One obvious limitation of this indicator is ignoring changes in the human capital stock of individuals that occur over their lifetimes due to labour market training, learning-by-doing, and skills depreciation, among other things. However, to the extent that the pattern of human capital accumulation over the life cycle of individuals is relatively similar across provinces, this limitation is largely mitigated by the fact that the average literacy score of each province is taken as a ratio of the cross-sectional mean. The third indicator, analogously to the synthetic literacy variable, is a synthetic time series of the average years of schooling of labour market entrants in each period, constructed from the years of schooling reported in ALL.

The main results

The estimated effect of human capital on per capita income, whether measured by the literacy or the university attainment indicator, is positive and significant. In contrast, the effect of the synthetic schooling variable is positive but non-significant in several regressions. Moreover, the effects of both literacy and university attainment are found to be positive and significant when the two indicators are used simultaneously as measures of human capital. This result may reflect the fact that some of the skills typically learned at university are not captured by ALL.

From a quantitative point of view, the average estimated elasticity of per capita income to literacy across the 20 regressions performed (using different specifications and regression methods) is 1.43, which is close to the results of Coulombe, Tremblay, and Marchand (2004) for 14 OECD countries. Using this elasticity and existing evidence on the marginal return to years of education in terms of literacy scores (OECD 2000) based on the International Adult Literacy Survey, we can construct an estimate of the macroeconomic return from the skills acquired through an extra year of schooling. Specifically, we find that the skills acquired by an extra year of schooling result in an increase in per capita income of 7.3 percent. Interestingly, this falls in the 5 to 15 percent range reported by Psacharopoulos (1994) for the individual return in terms of wages of an extra year of education. Thus our macroeconomic findings are consistent with the microeconomic Mincerian literature, especially as the individual return to education in a highly developed economy such as Canada is likely to fall in the lower part of the 5 to 15 percent interval. This is due to the fact that the marginal return to education will tend to be lower in economies where the average level of education is high and therefore human capital is relatively abundant.

Regression results also show that human capital indicators based on literacy test scores do not outperform those based on university attainment. This finding contrasts sharply with our recent result found at the cross-country level (Coulombe, Tremblay, and Marchand 2004). This suggests that literacy indicators outperform schooling indicators at the cross-country level because literacy test scores are more comparable than years of schooling.

Much of the recent literature on the determinants of relative living standards across economies has focused on human capital and on institutional improvements or investment in social infrastructure. Empirically, however, it is difficult to disentangle the relative contributions of each of these factors, given the problems associated with measuring the quality of institutions and social infrastructure. Of course, not accounting for differences in these factors will lead to biased estimates of the effect of human capital on per capita income in the likely event that human capital is correlated with the quality of institutions and social infrastructure. But to the extent that institutions and social infrastructure are similar across provinces, our analysis provides a more reliable estimate of the contribution of human capital to the relative standards of living across economies.

The analysis controls for the urbanization structure and for specific shocks in Alberta and Quebec. As mentioned above, the relative rate of urbanization across provinces is found to have a positive and highly significant effect on relative income per capita. Moreover, the point estimates are remarkably stable under different specifications and regression methods. Finally, shocks in Alberta and Quebec, respectively the 1973 oil shock and the start of the Anglophone exodus from Montreal around 1970, also explain part of the cross-sectional and time-series heterogeneity in the evolution of relative per capita income.

1. Introduction

What determines the differences in living standards across economies in the long run? The study of this central question in economics regained the front stage of mainstream economics in the last two decades with the “growth revival” pioneered by the works of Baumol (1986), Romer (1986), Lucas (1988), and Barro and Sala-i-Martin (1992) on endogenous growth theory, growth empirics, and convergence. According to Glaeser et al. (2004), after years of empirical and theoretical studies and debate, two answers to this central economic question stand out as candidates: (A) human capital accumulation (Mankiw, Romer, and Weil 1992; Barro, Mankiw, and Sala-i-Martin 1995; among others); and (B) institutional improvements or investment in social infrastructure as emphasized by, for example, Hall and Jones (1999).[1]

From an empirical point of view, testing candidates A and B using cross-country data is subject to some problems. First, it is very difficult to construct indicators of institutional quality that are comparable across countries and across time. According to Glaeser et al. (2004), the most often used indicators of institutional quality in cross-country growth studies are conceptually deficient. However, *not* using indicators of social infrastructure raises a second important problem in cross-country studies. Supporters of candidate answer B might interpret a positive estimated effect for the human capital indicator as a missing variable bias if the human capital indicator is positively correlated with the quality of institutions. An alternative empirical strategy that avoids both potential problems would be to focus on the role of human capital, using within-country regional data sets. Since the social infrastructure is relatively similar across regions of homogeneous countries, the process of human capital accumulation should account for most differences in standards of living if candidate answer A is at least a part of the story.[2] Empirical practitioners, however, have not used this approach as much because of the lack of reliable data on human capital and standards of living at the subnational level in most countries for a sufficiently long period of time.

In this paper, we follow the regional strategy using Canadian provincial data and test whether human capital accumulation can account for income differences across economies. To this end, we focus on two very different measures of human capital. The first is university achievement, measured as the percentage of the working-age population holding a university degree. This measure, used in recent studies on Canadian provincial convergence (Coulombe and Tremblay 2001; Coulombe 2003), is available at 10-year intervals from census data since 1951.[3] The contribution of this paper is that we also use a new direct indicator of human capital based on literacy test scores. In collaboration with Statistics Canada, we computed and aggregated the new literacy data for the 10 Canadian provinces for the 1951–2001 period from the 2003 Adult Literacy and Lifeskills Survey (ALL) and the 1994 International Adult Literacy Survey (IALS). Following the methodology proposed recently in an OECD cross-country study by Coulombe, Tremblay, and Marchand (2004), we provide the cross-sectional data with a time-series dimension inferred from the demographic structure of the ALL and IALS data. The data are intended to capture the mean literacy level of labour market entrants aged 17 to 25 for each of the 10 Canadian provinces.

The data are also broken down by gender to capture a possible gender-gap effect. Unlike data on school achievement, literacy data might be viewed as a direct measure of human capital that is more comparable across time and across economies (Coulombe, Tremblay, and Marchand 2004). The new data are then used in pooled time-series and cross-sectional (the 10 provinces) empirical models (TSCS) to estimate the mean effect of human capital on aggregate provincial per capita income, minus government transfers.

Our empirical analysis in general provides support for the human capital explanation. Both human capital indicators — based on university achievement and literacy — are found to exert a positive and significant effect on per capita income level when entered in separate TSCS regressions. Furthermore, the effects of literacy and university achievement indicators are both found to be positive and significant when they are used simultaneously as indicators of human capital in TSCS analyses.

A large body of empirical works has recently focused on the role of human capital accumulation in cross-country growth.[4] Generally speaking, when the cross-country sample includes a large set of developed and underdeveloped countries, schooling achievement is found to generate a positive and significant effect on transitory growth and the long-run level of labour productivity or per capita income in growth regressions (Barro 2001). However, when the data set is restricted to developed countries, the effect of various schooling variables has not usually been significant and has sometimes even been negative (Islam 1995; Barro 2001). These divergent results might be interpreted in at least two different, but not necessarily alternative, ways. First, as we argued before, in a broad set of heterogeneous countries, schooling indicators may be positively correlated with missing variables related to social infrastructure. The effect of education may vanish in a sample of OECD countries since developed countries are more homogeneous in this dimension (Coulombe 2001). Second, measurement error on schooling data can be a big issue, especially in cross-country studies (Krueger and Lindahl 2001). De la Fuente and Doménech (2002) address the measurement error issue extensively in their study of OECD countries; their corrected education variable appears to exert a positive and significant effect on living standards. Coulombe, Tremblay, and Marchand (2004), using the same type of synthetic human capital indicators derived from literacy data that are used in this paper, found that literacy indicators systematically outperformed de la Fuente and Doménech's (2002) corrected data in standard convergence growth regressions in a restricted set of 14 highly homogeneous OECD countries. (This country set excluded Spain, Portugal, and Greece).

The key contribution of this paper is not limited to the regional dimension of the analysis, however. From the regional analysis point of view, we show that the accumulation of human capital, whether measured by an indicator based on literacy, university achievement, or a combination of both, can explain a substantial portion of the evolution of differences in per capita income across Canadian provinces since 1951. In line with Coulombe (2000, 2003), other variables of course also matter, variables such as the relative urbanization that appears to be a time-invariant parameter in the Canadian regional case. Two other region-specific factors, the oil shock for Alberta in 1973 and the start of the so-called Anglophone exodus from Montreal around 1970,[5] are also responsible for cross-sectional and time-series heterogeneity in the evolution of relative per capita income. These results are not new since they basically concur with our earlier findings. What is new is that both literacy and university matter for the aggregate well-being of a regional economy. This result contrasts sharply with Coulombe, Tremblay, and Marchand's (2004) main result and suggests that the reason their literacy indicators outperform the best available schooling indicator in that cross-country study is measurement error bias. Literacy data appear to be more comparable across time and across countries than schooling data.

In Section 2, we present the data and compare the reliability of our different indicators of human capital. Section 3 discusses the theoretical foundation of our analysis and the empirical methodology. Results are presented and discussed in Section 4 and Section 5 concludes.

2. Measures of human capital

2.1 Data

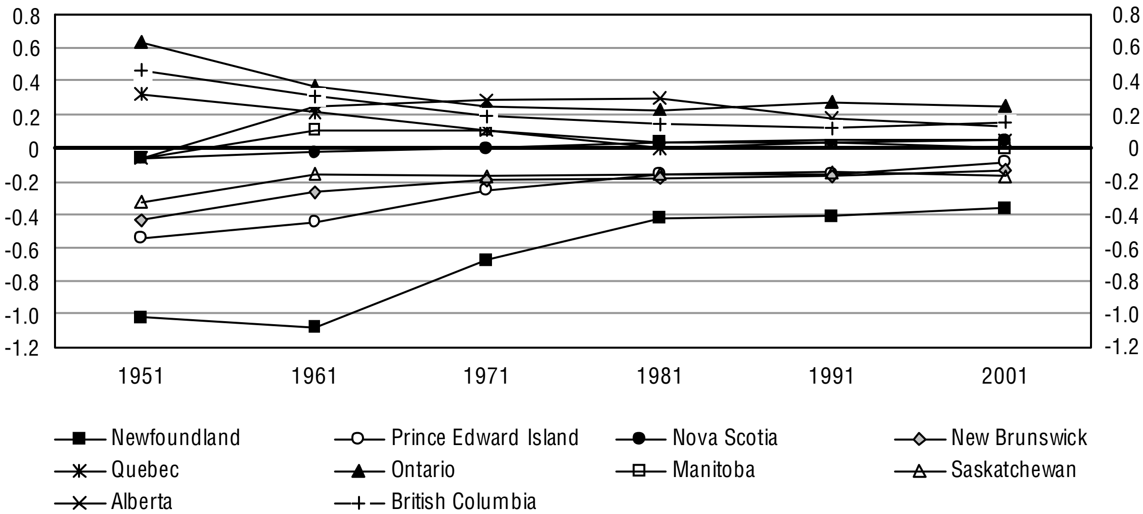
Various measures based on years of schooling have been widely used in cross-country studies with differing levels of success. De la Fuente and Doménech (2002) argue that the negative results might be ascribed to measurement error and their corrected schooling measures do appear to perform better than measures taken from other data banks such as in the Barro and Lee study (2000). Using a restricted set of 14 OECD countries, Coulombe, Tremblay, and Marchand (2004) show that literacy data derived from the demographic profile of the 1994 IALS survey systematically outperformed de la Fuente and Doménech's corrected data in various growth regression frameworks.

One of the main reasons why measurement error might be a serious issue with cross-country schooling data is that raw national data are assembled by various statistical agencies using different methodologies. Schooling data for the Canadian provinces should be much more comparable. However, as elegantly illustrated by de la Fuente and Doménech's case study (2002, Section 3.1) on various Canadian indicators dealing with educational achievement, most Canadian data are not consistent on a time-series basis. The census questionnaires have changed their format over time and the only official *consistent* data that can be used in this paper are those dealing with university attainment.[6] More precisely, our benchmark schooling indicator is based on the percentage of the population in the 15-to-65 age group with at least one university degree. In principle, this measure is less directly related to the mean human capital stock of an economy than the usual measure based on the mean number of years of schooling that is widely used in micro-econometric Mincerian studies. However, our earlier studies dealing with the human capital accumulation process across Canadian provinces have shown well that university attainment data might be viewed as appropriate proxies for the relative level of human capital across the provinces in TSCS analyses (Coulombe and Tremblay 2001; Coulombe 2003).

Our first human capital indicator, the percentage of the population aged between 15 and 65 with at least one university degree, is shown for each province in figure 1. The percentages are transformed in logarithms of deviations from the unweighted provincial mean. Clearly, these indicators have converged across provinces over the period. The percentage of university graduates has grown faster in provinces that initially had a relatively low percentage of individuals with a university degree. Despite this convergence pattern, the relatively rich provinces of Ontario, Alberta, and British Columbia have had the highest percentage of university graduates throughout the period. Newfoundland, Prince Edward Island, and New Brunswick have been trailing behind. In contrast to the other Atlantic Provinces, Nova Scotia has been, and still is, relatively well endowed in university graduates. We will return to the case of Nova Scotia later. Finally, the percentage of university graduates in Quebec initially was substantially above the Canadian average but declined steadily during the 1960s and 1970s due to the exodus of the relatively well-educated Anglophone population. The effect of this exodus on the relative growth of GDP per capita in Quebec is tested in the empirical analysis that follows.

Figure 1

Percentage of the population aged between 15 and 65 with a university degree (log of deviations from the mean)



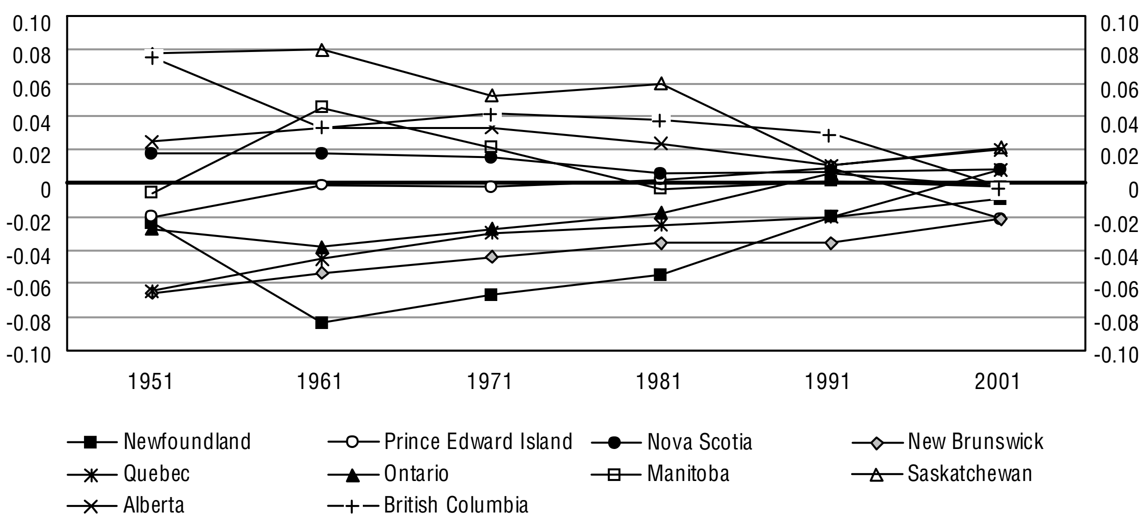
The second indicator of human capital used in this study is based on literacy test scores and is intended to capture the mean literacy level of labour market entrants. Literacy test scores are taken from the 1994 IALS and the 2003 ALL, which tested the literacy skills of individuals between 16 and 65 years of age. Three domains of literacy (prose, document, and quantitative) were tested in IALS and four (prose, document, numeracy, and problem solving) in ALL. The literacy level of an individual is equal to the average score over these literacy domains. Using the demographic profile of test scores, we constructed a synthetic time series of the mean literacy level of individuals aged between 17 and 25 for each period starting in 1951 (1951, 1956, 1961, …, 2001). Using both the 1994 IALS and the 2003 ALL allows us to derive two indicators of the average literacy level for the same 17-to-25-year-old cohort for each period with the exception of 2001. The reliability of these two indicators will be compared in the next section.

Implicit in the construction of these indicators is the assumption that the skill level of individuals remains constant during their lives in the labour market. Hence, the indicator does not take into account the changes in skills that result from labour market training, learning-by-doing through labour market experience, and skills depreciation. Moreover, it does not take into consideration the migration flows that occurred over the period. These limitations should be kept in mind when interpreting the results. On the other hand, these indicators provide a direct measure of the quality of human capital and might be more comparable across time and space than schooling indicators.

Figure 2 shows these indicators for each province, again expressed as logarithms of deviations from the provincial mean. Similar to the university graduate indicator, our literacy indicator also exhibits a clear convergence pattern. Interestingly, however, there are notable differences in the university and literacy indicators of some provinces. In particular, the literacy indicator for Quebec has increased steadily over the period in sharp contrast to the evolution of its university graduates indicator. While Quebec's labour market entrants in 1951 had the lowest level of literacy in the country, along with those in New Brunswick, the Quebec labour market entrants were above the provincial average in 2001. Ontario's literacy indicator has surprisingly been below average for all periods except for 1991 despite the province's very high level of university graduates. Western provinces have generally had the best performance throughout the period with the labour market entrants of Saskatchewan and Alberta having the highest average score in 2001. Finally, Newfoundland and New Brunswick have been trailing behind for most of the period; Prince Edward Island has been oscillating around the provincial average; and Nova Scotia has remained steadily above it.

Figure 2

Average literacy scores of population aged between 17 and 25 (log of deviations from the mean)



Figures 3 and 4 depict the evolution of the standard deviation of the university graduates indicator and the literacy indicator, respectively. The dispersion in both indicators of human capital across provinces has clearly been decreasing over the period. Interestingly, the relative decrease in the dispersion of the university indicator occurred somewhat earlier. Most importantly, however, the dispersion of the university graduates indicator is approximately 10 times that of the literacy indicator. The low dispersion of the literacy indicator results from the arbitrary 0–500 scale that was chosen to report the literacy performance of individuals. We return to this issue below.

Figure 3

Standard deviation of the university graduates indicator

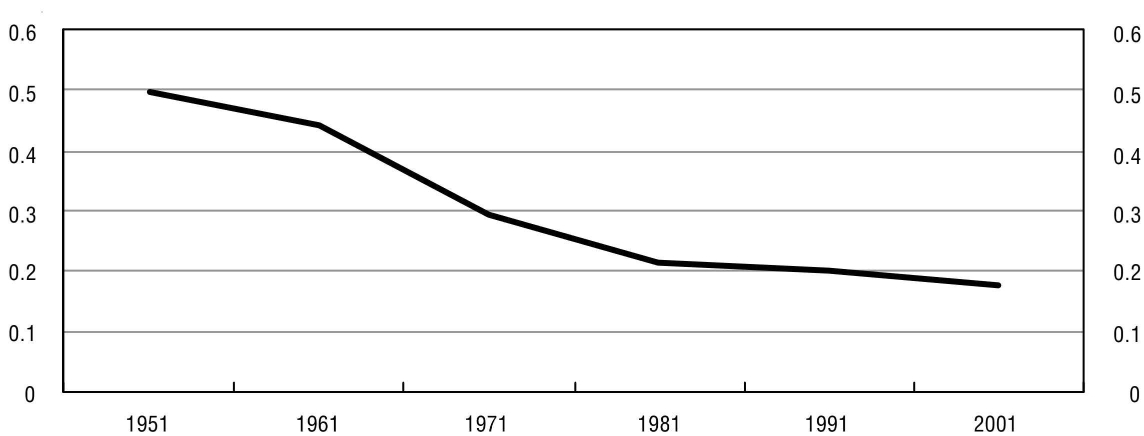
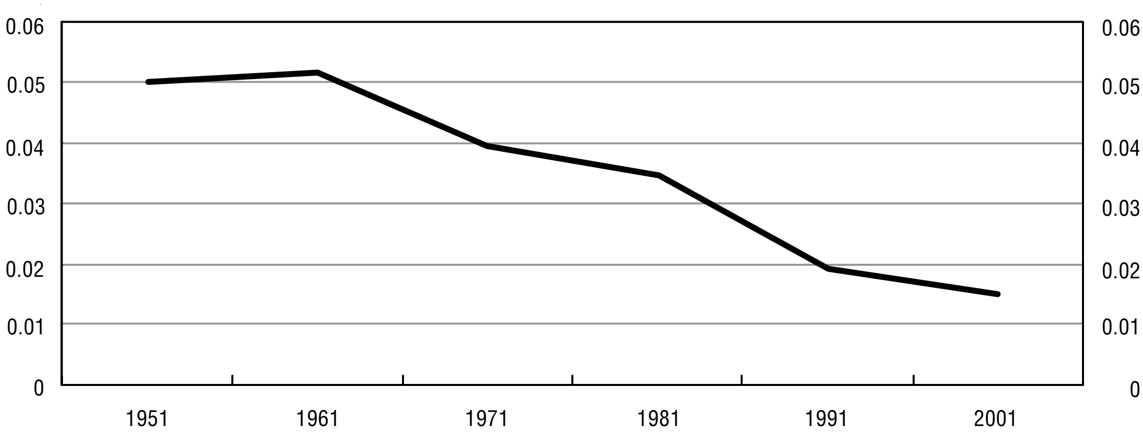


Figure 4
Standard deviation of the literacy indicator



Both types of indicators, based on university attainment and literacy, will be used below to estimate the effect of human capital on the relative level of provincial per capita income minus government transfers to individuals. These income measures, expressed as logarithms of deviations from the mean, are shown for each province in figure 5. Although there has been substantial convergence, Ontario and the western provinces (except for Saskatchewan) have had the highest levels of per capita income minus government transfers throughout the period; the Atlantic Provinces have stood at the lower end. Figure 6 shows that dispersion across provinces has decreased in all subperiods except during the 1970s when relative per capita income minus transfers increased substantially in Alberta and decreased in Newfoundland, New Brunswick, and Nova Scotia.

Figure 5
Income per capita minus government transfers (log of deviations from the mean)

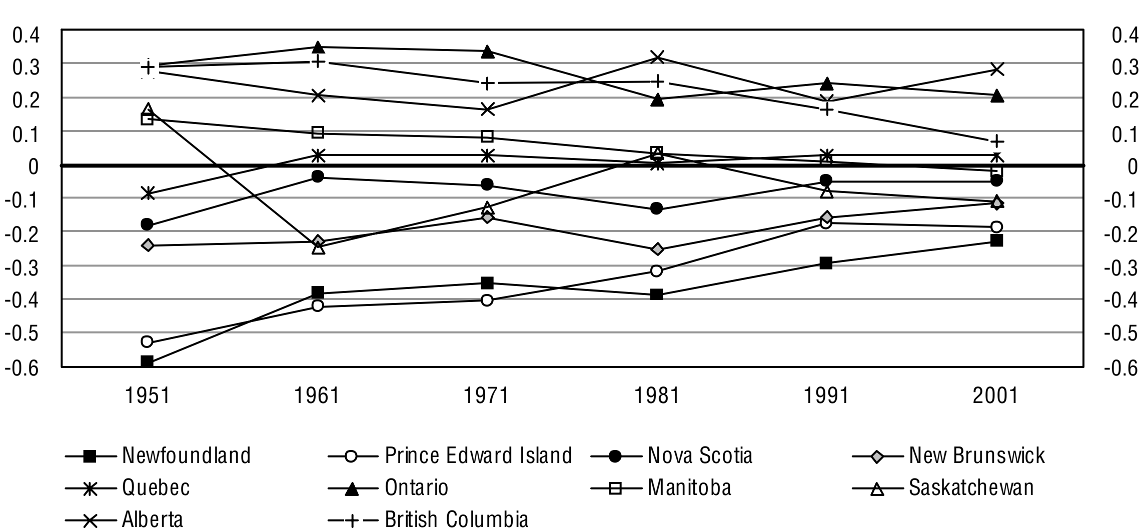


Figure 6
Standard deviation of the per capita income indicator

Finally, we also use in this paper a third type of human capital indicator, constructed from the reported years of schooling in the ALL 2003 survey. We generated synthetic time series to capture years of schooling of labour market entrants, using the same procedure that was based on the demographic profile to generate the time-series dimension in the literacy variable. The effect of this synthetic schooling variable, although positive, was found not significant even at the 10 percent level in many regressions. We report some results using this schooling variable but we keep the emphasis on the university achievement data coming from census data in the following discussion.

2.2 Reliability ratios

We have at our disposal four indicators of human capital: the literacy data from the ALL 2003 survey ($L03_{i,t}$), the literacy data from the IALS 1994 survey ($L94_{i,t}$), the university achievement census data ($U_{i,t}$), and the synthetic schooling data derived from the ALL 2003 survey ($S03_{i,t}$), where i = 1,…, 10 stands for the 10 Canadian provinces and t denotes the year.

It is possible to compare the amount of information on human capital (the signal) with the measurement error (the noise) contained in these alternative human capital measures with the *reliability ratio* concept. Krueger and Lindahl (2001) have estimated the reliability ratios of alternative schooling data used in cross-country studies to illustrate the extent of measurement error in human capital data sets. De la Fuente and Doménech (2002) have also used this approach in a TSCS framework for a variety of human capital indicators.

Let us suppose that H_{it} is the true measure of human capital per person across province i at time t, and that we have two alternative proxy variables of this true measure of human capital, $H1_{it}$ and $H2_{it}$:

$$H1_{i,t} = b_1 H_{i,t} + e^1_{i,t}$$
$$H2_{i,t} = b_2 H_{i,t} + e^2_{i,t}$$

where $e^1_{i,t}$ and $e^2_{i,t}$ are the measurement errors. The reliability ratio ρ_X of indicator HX (where X stands for either indicator 1 or 2) is the ratio of the signal to the signal plus noise and corresponds to

$$\rho_X = \frac{\mathrm{var}(H_{i,t})}{\mathrm{var}(HX_{i,t})} = \frac{\mathrm{var}(H_{i,t})}{b_X^2\left[\mathrm{var}(H_{i,t}) + \mathrm{var}(e_{i,t}^X)\right]}.$$

Note that a reliability ratio varies between 0 (complete noise) and $1/b_X^2$ (no measurement error). Reliability ratios for the two indicators can be estimated under the rather restrictive assumption that both are independent noisy estimates of the true concept to measure. To this end, we assume that $b_1 = b_2 = 1$ and that the measurement errors are “classical”: they are uncorrelated (across time and across provinces) and are treated as white noise disturbances.

Reliability ratios are estimated by the slope coefficient in bivariate regressions of one human capital measure on the other. In our case, given the time-series and cross-sectional nature of our data set, reliability ratios are estimated with time dummies in all specifications:

$$H2_{i,t} = \rho_1 H1_{i,t} + \lambda_t^1 + \varepsilon_{i,t}^1$$
$$H1_{i,t} = \rho_2 H2_{i,t} + \lambda_t^2 + \varepsilon_{i,t}^2$$

Three pairs of human capital indicators will be compared using this framework. In the first pair, the data coming from the IALS 2003 survey ($L03_{i,t}$) are compared with the same type of data coming from the 1994 survey ($L94_{i,t}$). In the second pair, the $L03_{i,t}$ data are then compared with university achievement census data ($U_{i,t}$). Finally, in the third pair, the data coming from the IALS 2003 survey ($L03_{i,t}$) are compared with the synthetic schooling data derived from the same survey ($S03_{i,t}$).

There are two potential complications associated with the estimation of reliability ratios for two different human capital measures. First, as pointed out in Krueger and Lindahl (2001), the measurement errors might be correlated. When comparing data sets such as the $L03_{i,t}$, the $S03_{i,t}$, and the $L94_{i,t}$ data, the errors might indeed be positively correlated since they are derived from the demographic profile of IALS surveys using the same assumptions. If measurement errors are positively correlated, the reliability ratios of both data sets are biased upward. However, this problem should not apply when the measurement error analysis deals with the comparison of $L03_{i,t}$ and $U_{i,t}$ since they are derived from completely different raw data and methodologies.

The second potential complication is that b_1 might not be equal to b_2. This complication applies specifically to the analysis dealing with the comparisons between $L03_{i,t}$ on the one hand and $U_{i,t}$ and $S03_{i,t}$ on the other. Three factors might generate this problem.

The first factor is that university achievement, years of schooling, and literacy data are not measured in the same scale. The first two are measured as a percentage (of the population) and in number of years respectively; the literacy data are measured in a somehow arbitrary 0–500 scale. We tackle this problem in estimating reliability ratios for the second and third pairs of data by reporting *beta coefficients*. We do this also for all regressions pertaining to the 10-year data set where the effect of literacy data is compared with the effect of schooling. For this purpose, the literacy and the university achievement data are *standardized*. To this end, the variables are transformed by subtracting the cross-sectional mean (in each time period) and by dividing the result by its sample standard deviation.

The second factor is that the inequality between b_1 and b_2 might result from the fact that the literacy data are measures of the human capital intensity of a fraction of the overall labour force whereas the university achievement data are related to the whole labour force. This problem is potentially serious if the skill level of the young cohort is not a good proxy of the mean skills of the labour force. In such a case, both the estimated reliability ratios of $L03_{i,t}$ and $U_{i,t}$ will be biased. Of course, the primary reason for this is that, over time, the skills of the young cohort tend to increase with the general improvement in education observed over the period under study. But this common trend in all provinces is eliminated from the panel data analysis with the introduction of time dummies, which implies that the variables are transformed by the cross-sectional demeaning procedure.

From an econometric point of view, the skill level of the young cohort is a good proxy for the skill of the overall labour force if the slope coefficient s in a bivariate regression that links the skills of the young cohort L_Y to the mean skills of the other cohort L_{ALL-Y} is 1:

$$L_{ALL-Y} = sL_Y + \varepsilon_Y$$

If s is smaller (greater) than one, the reliability ratio of $L03_{i,t}$ will be underestimated (overestimated) and the reverse is true for $U_{i,t}$. We could not test the hypothesis that $\hat{s} = 1$ using time-series and cross-sectional information since the skills of the overall labour force are measured at only one point in time in the 2003 IALS survey. However, the slope parameter can be estimated only with the cross-sectional data. We have regressed the mean skill level of the other cohort on the skill level of the young cohort. With just 10 observations at hand, one has to bear in mind that this estimation is not very precise. However, the estimated $\hat{s}$ are just slightly larger than one (estimated with pooled least squares [PLS] and generalized least squares [GLS]) and a Wald test clearly shows that the null hypothesis ($\hat{s} = 1$) is not rejected at the 5 percent level.

The third factor that could explain a potential inequality between b_1 and b_2 is that literacy and university achievement or schooling data might, generally speaking, reflect different aspects of the true human capital concept. For example, literacy is a direct measure of skills and might be viewed as a function of both schooling achievement and the quality of schooling. In such a case, the $\hat{\rho}$ estimated from regressions (2) can be interpreted as reliability ratios under a set of very restrictive assumptions that cannot be tested. However, in this case, the estimated $\hat{\rho}$ still provide useful information on the statistical relationship between the indicators.

The reliability ratios estimated from the bivariate regressions between $L03_{i,t}$ and $L94_{i,t}$ are presented in the first two columns of table 1. The results are particularly revealing since they clearly indicate that the measurement error is much larger in the 1994 survey than in the 2003 survey. The estimated reliability ratios are very close to 1 for the $L03_{i,t}$ variable (first column) and are close to 0 for the $L94_{i,t}$ variable (second column). Under the assumption that the measurement errors for the two variables are not correlated, the analysis suggests that the noise-to-signal ratio is very small in the 2003 data set whereas measurement errors dominate the data in the 1994 survey.

Table 1

Reliability ratios of literacy and schooling data

		(1) $L94_{i,t}$ regressed on $L03_{i,t}$	(2) $L03_{i,t}$ regressed on $L94_{i,t}$	(3) $U_{i,t}$ regressed on $L03_{i,t}$	(4) $L03_{i,t}$ regressed on $U_{i,t}$	(5) $S03_{i,t}$ regressed on $L03_{i,t}$	(6) $L03_{i,t}$ regressed on $S03_{i,t}$
PLS	Coefficient	0.97 [a]	0.28 [a]	0.33 [b]	0.35 [b]	0.69 [a]	0.72 [a]
	Standard error	0.15	0.10	0.17	0.14	0.10	0.09
	R^2	.27	.27	.10	.10	.50	.50
GLS	Coefficient	0.94 [a]	0.25 [a]	0.36 [a]	0.40 [a]	0.83 [a]	0.78 [a]
	Standard error	0.12	0.06	0.07	0.09	0.05	0.05
	R^2	.55	.55	.43	.43	.78	.78

Notes: Ten-year periods, resulting in 60 panel observations. Beta coefficients are reported in columns 3 to 6. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients. $L03_{i,t}$ are the mean literacy data from the ALL 2003 survey, $L94_{i,t}$ are the same type of data coming from the 1994 IALS survey, $U_{i,t}$ are university achievement census data, and $S03_{i,t}$ are the schooling data from the ALL 2003 survey.

a significant at 1% level;

b significant at 5% level;

c significant at 10% level.

This striking result might be explained by two facts. First, the sample of individuals used in the 2003 survey is much larger than in the 1994 survey. Second and more importantly, contrary to the 1994 survey, the smaller provinces (such as Prince Edward Island) have been greatly oversampled in the 2003 survey.[7] Consequently, our empirical analysis dealing with the relationship between literacy and provincial disparity will focus primarily on the data coming from the IALS 2003 survey. In some regressions with instrumental variables (IV), we will use the data from the 1994 survey as instrument for the $L03_{i,t}$ variable.

The reliability ratios estimated from the pair of variables composed by $L03_{i,t}$ and $U_{i,t}$ are displayed in columns (3) and (4) of table 1. Interpreted within the framework of the classical measurement error, the reported point estimates, which are beta coefficients since the two variables are not measured on the same scale, indicate that the signal/(signal plus noise) ratio is very comparable in the two databases. Both estimated reliability ratios are slightly higher with GLS than with PLS and range between 0.33 and 0.40. If both variables are considered as alternative noisy estimates of the true human capital variable, the results indicate that measurement error is an important component of both. When a regressor is measured with an additive “classical” noise error, its point estimates in least-squares regressions are biased toward zero, a problem known as the attenuation bias. Consequently, we can expect that the point estimate of the human capital variable in least-squares regressions will be biased toward zero to some extent.

Interpreted outside the classical measurement error framework, the point estimates reported in the bivariate regressions in columns (3) and (4) of table 1 indicate that the $L03_{i,t}$ and $U_{i,t}$ variables are strongly positively linked in a TSCS model. However, the relatively low reliability ratios reported in columns (3) and (4) can also be interpreted by the fact that both variables might measure different aspects of the human capital concept that are not perfect proxies of the overall concept. This interpretation indicates that it might be interesting to verify in the empirical analysis of the effect of human capital accumulation whether both the $L03_{i,t}$ and $U_{i,t}$ variables can be combined in some ways.

Finally, the reliability ratios estimated from the pair of variables composed by $L03_{i,t}$ and $S03_{i,t}$ are displayed in the last two columns of table 1. The two variables are strongly positively correlated and the estimated reliability ratio is very comparable for both variables. They are around 0.70 with PLS and 0.80 with FGLS. As pointed out earlier, the fact that the estimated reliability ratio is higher for the $S03_{i,t}$ than the $U_{i,t}$ variable might result from a positive correlation between the measurement errors for $L03_{i,t}$ and $S03_{i,t}$.

3. Theoretical foundations and empirical methodology

3.1 From the production function

Our benchmark results of the effect of human capital accumulation on income differences come from an empirical framework that has many similarities with the traditional growth accounting framework.[8] Both frameworks are based on the production function. Let us suppose that output Y of economy i at time t is described by the following Cobb-Douglas production function:

$$Y_{i,t} = A_{i,t} K_{i,t}^{\alpha} H_{i,t}^{\eta} L_{i,t}^{1-\alpha-\eta}, \quad \text{with: } 0<\alpha<1, 0<\eta<1, \text{ and } \alpha+\eta<1. \qquad (1)$$

In this set-up, inputs K and H are respectively the stock of physical and human capital, input L is "raw" labour, and A is the state of the technology. With constant returns to scale, the production function might be written in units of labour:

$$y_{i,t} = A_{i,t} k_{i,t}^{\alpha} h_{i,t}^{\eta}, \quad \text{where: } y = Y/L, k = K/L, \text{ and } h = H/L. \qquad (2)$$

Taking the logarithms on both sides of (2) yields:

$$\ln y_{i,t} = \ln A_{i,t} + \alpha \ln k_{i,t} + \eta \ln h_{i,t}. \qquad (3)$$

In growth accounting, assumptions are made regarding parameters α and η. The growth rate of technological progress $\Delta \ln A_{i,t}$ is measured as the Solow residual when the effect of human and physical capital accumulation ($\alpha \Delta \ln k_{i,t} + \eta \Delta \ln h_{i,t}$) is withdrawn from the growth of output ($\Delta \ln y_{i,t}$). As pointed out in Topel (1999), the private return of human capital in the growth accounting approach is implicitly assumed to be equal to the social return. The macro-econometric literature on the effect of human capital, pioneered by Mankiw, Romer, and Weil (1992), is less restrictive; the effect of human capital is estimated freely in regressions based on (3) using (cross-country) macro-data in which the technology parameter $A_{i,t}$ is included in the error term:

$$\ln y_{i,t} = \beta_0 \ln k_{i,t} + \beta_1 \ln h_{i,t} + \ln A_{i,t}. \qquad (3')$$

With appropriate indicators of the stock of human and physical capital, equation (3´) can be estimated from a pure cross-section of countries when the data are available in just one point in time, as was done in Mankiw, Romer, and Weil (1992). In our executive summary, the cross-section approach is used as an illustrative example of the correlation between human capital and per capita income disparities across the ten provinces in 2003. However, equation (3´) could be estimated more efficiently with TSCS data when the data are also available through time, as it is the case for the Canadian provincial data. With TSCS regressions, the error term that embodies the technology parameter can be modelled in a much more general way than in a pure cross-sectional study.

The main complication in estimating parameter β_1 (the social return to human capital) in (3´) is finding reliable data on the stock of human capital and physical capital that have time-series and cross-sectional dimensions. If the capital stock is measured with error, the estimator of $\hat{\beta}_1$ will generally be biased. We do not need data on the capital stock in a TSCS framework, however, if the capital/output ratio is assumed to be constant through time. This hypothesis is consistent with Kaldor's (1963) stylized facts on growth and has been used in TSCS analysis of this type in Coulombe and Tremblay (2001) and Lange and Topel (2004). This stylized fact might be explained in a Cobb-Douglas production function framework by the assumption that the marginal product of capital is constant through time. Barro, Mankiw, and Sala-i-Martin's (1995) model of a small open economy with perfect (physical) capital mobility and a binding constraint for the financing of human capital generates this result. Coulombe and Tremblay's (2001) analysis illustrates that this assumption describes well the evolution of Canadian regional economies in the 1951–1996 sample. In this case, $\ln y_{i,t} - \ln k_{i,t} = c_i$ and equation (3´) simplifies to:

$$\ln y_{i,t} = \beta_h \ln h_{i,t} + \bar{c}_i + \ln \bar{A}_{i,t}, \qquad (4)$$

where $\bar{c}_i$ and $\ln \bar{A}_{i,t}$ are renormalizations of c_i and $A_{i,t}$ and $\beta_h = \left(\beta_1 \big/ 1 - \beta_0 \right)$.

The TSCS structure of our empirical analysis allows us to model the technology parameter $\ln \bar{A}_{i,t}$ in a general way. It is decomposed in three components: the initial technology levels $A_{i,0}$ that are allowed to vary across provinces; a technological growth component $g(t)$ that is assumed to be the same across provinces but allowed to vary through time; and idiosyncratic disturbances $\varepsilon_{i,t}$:

$$\ln \bar{A}_{i,t} = A_{i,0} + g(t) + \varepsilon_{i,t}. \qquad (5)$$

The term embedding the growth rate of technological progress in (5) is treated as an unobservable time-specific fixed effect λ_t in all TSCS models in this paper. This widely used procedure in TSCS analysis implies that all variables in the regressions are transformed using a cross-sectional demeaning procedure or that T-1 time dummies are entered in the regressions. Common shocks to all provinces, such as the productivity slowdown, must be eliminated in an analysis of this type in order to get unbiased estimates of β_h. With this modelling, equation (4) can be written as:

$$\ln y_{i,t} = \beta_0 \ln h_{i,t} + \bar{c}_i + A_{i,0} + \lambda_t + \varepsilon_{i,t}. \qquad (6)$$

3.2 Benchmark empirical set-ups

All empirical set-ups based on equation (6) take the following structure:

$$\ln y_{i,t} = F(h^*_{p,i,t}, Z_{i,}, Z_t, Z_{i,t}, \varepsilon_{i,t}),$$

for $p = u$ (human capital indicator based on university achievement data) and/or l (indicator based on literacy); i = 1,…,10 for the ten Canadian provinces; t = 1,…, 6 for TSCS of 10-year periods; and t = 1,…,11 for TSCS of 5-year periods. The Z_i are cross-sectional specific controls (time-invariant) and the Z_t are time-period specific controls (cross-sectional invariant). The $Z_{i,t}$, which are cross-sectional and time-varying dummies, are used in certain structural specifications to capture the effect of specific shocks that affected the provinces of Alberta and Quebec in the period under study (1951–2001).

For our first empirical TSCS model, $\bar{c}_i + A_{i,0}$ of equation (6) are amalgamated into an unobservable cross-sectional specific effect γ_i:

$$\ln y_{i,t} = \beta_p \ln h^*_{p,i,t} + \gamma_i + \lambda_t + \varepsilon_{i,t}. \qquad \text{(R1)}$$

With this fixed-effects approach, the y and h variables are transformed by taking the differences, in all periods, from the time-period mean. The fixed-effects transform procedure is the straightforward one to use if one is interested in getting estimates of only the $\hat{\beta}_p$. However, that procedure cannot be used for estimating the effect of time-invariant determinants of the technology level and the capital/output ratio in (6).

In line with previous studies (Coulombe 2000, 2003) on the conditional convergence of Canadian provinces, we also use a "structural" version of (6). In this version, a rural/urban structure variable, UR_i, is used as a proxy variable for the time-invariant determinants (capital/labour ratios and initial technology levels) in the production function:

$$\ln y_{i,t} = \beta_p \ln h^*_{p,i,t} + \varphi_1 UR_i + \varphi_2 AB_{i,t} + \varphi_3 QU_{i,t} + \varphi_4 NS_i + \lambda_t + \varepsilon_{i,t}. \qquad \text{(R2)}$$

The UR_i variable captures the relative (measured as the logarithm deviation from the cross-sectional sample mean) degree of urbanization of the 10 provinces. As shown in Coulombe (2000), even if the post-1950 period is characterized by a steady urbanization in all provinces, the relative degree of urbanization is quite stable for all provinces during the period for which these data are available. Also in line with our previous studies, we account for province-specific shocks that disturbed the growth patterns of Alberta in 1973 (the first oil shock) and Quebec around 1970 (the Anglophone exodus and the relative decline of Montreal). The *AB* and *QU* variables take the value 0 for the other nine provinces and the value 1 for Alberta (*AB*) after the oil shock and Quebec (*QU*) after the Anglophone exodus. For these two provinces, their respective shock variables take also the value 0 prior to the shock. As shown in Coulombe (2003), a Nova Scotia (*NS*) dummy variable, which takes the value 1 for only this province and the value 0 for the other nine provinces, will also be found significant in some structural regressions. As discussed in Coulombe (2003), the Nova Scotia dummy captures the various specific characteristics of that province's university system and the tendency of educated Nova Scotians to remain in their relatively poor and less urbanized province even if the return on education is less than in richer provinces.

The third empirical set-up used in this paper might be viewed as an intermediate specification between R1 and R2. In the fixed-effects transform, the effect of time-invariant variables (or variables that change very slowly through time) such as the rural/urban structure cannot be estimated since the effect is eliminated by the time-demeaning transformation. Random-effects estimation, however, allows the estimation of the effect of time-invariant controls since it involves only a quasi-time-demeaning transformation of the data. In the random-effects TSCS model R3, the unobserved effect re_i is assumed to be uncorrelated with the human capital indicator and the rural/urban structure variables:

$$\ln y_{i,t} = \beta_p \ln h^*_{p,i,t} + \varphi_1 UR_i + re_i + \lambda_t + \varepsilon_{i,t}. \qquad \text{(R3)}$$

The random-effects approach is not often used in TSCS analysis where the number of cross-sectional units is fixed and determined by the number of countries or regions under study. The interest in the random-effects approach in this investigation is that the effects of both the human capital and the rural/urban structure can be estimated under the assumption of unobserved heterogeneity across the 10 provinces. The analysis will illustrate the robustness of the estimations of $\hat{\beta}_p$ and $\hat{\varphi}_1$ in specifications R1 and R2.

3.3 The convergence-growth framework

In recent empirical literature in macroeconomics on the effect of human capital accumulation, a significant portion has taken a different approach than the one suggested by equation (6). The alternative approach, proposed by Barro and Sala-i-Martin (1995) in their work on convergence, has been implemented in numerous recent studies dealing with the effect of human capital. These studies include Barro (2001), de la Fuente and Doménech (2002), and Coulombe, Tremblay, and Marchand (2004). The approach is based on the convergence equation, the fundamental property of transitional dynamics in neoclassical growth models (Barro and Sala-i-Martin 1995). Convergence stipulates that the transitory growth rate (in excess of the growth rate of technological progress, which is the growth rate of per capita GDP on steady state) is proportional to the gap between the initial level of per capita GDP, $y_{i,t-1}$, and its steady-state level, y_i*, both being measured in efficiency units of labour:

$$\Delta \ln y_{i,t} = \psi(\ln y_i * - \ln y_{i,t-1}),$$

where parameter ψ is the speed of convergence toward steady state. The convergence equation has been widely estimated in cross-sections of countries or regional data such as the U.S. states (Barro and Sala-i-Martin 1995) or in TSCS frameworks in both types of data sets (Islam 1995; Coulombe and Lee 1995). In a pure cross-section of data, convergence regressions adopt the following structure:

$$\ln y_{i,T} - \ln y_{i,0} = -\psi y_{i,0} + \theta_h h_{i,0} + \theta' z_{i,0} + \varepsilon_i. \quad \text{(R4)}$$

The growth rate during the period between time 0 and time T is regressed on the initial level of per capita income, the initial level of human capital, and a set of other controls $z_{i,0}$. In this set-up, the initial human capital $h_{i,0}$ and the other controls $z_{i,0}$ are viewed as proxy variables for the steady-state level y_i*. Since ψ is a positive fraction, the long-run solution of equation (R4) is:

$$y_i * = \left(\frac{\theta_h h_i * + \theta' z_i *}{\psi}\right). \quad (7)$$

A change in any controls, including human capital, exerts a transitory effect on the growth rate of the economy and a long-run effect on the level of per capita GDP. From a growth regression that takes the form of equation (7), the estimated long-run elasticity of per capita output to human capital is $-\hat{\theta}_h / \hat{\psi}$.

Topel (1999) and Krueger and Lindahl (2001) have argued that the point estimate $\hat{\theta}_h$ in the convergence growth regression cannot be interpreted in a straightforward way. In this vein, Coulombe (2000) argues that in the open-economy framework of Barro, Mankiw, and Sala-i-Martin (1995) — certainly a model better able to account for the growth of Canadian provinces than the closed-economy version of the neoclassical growth model from which the convergence-growth equation is derived — the regression model is mis-specified when both the initial human capital and the initial per capita output are included in the list of controls. If Barro, Mankiw, and Sala-i-Martin's (1995) framework is the true model, $y_{i,t} = f(h_{i,t})$ and model (R4) suffers from perfect collinearity. If regression (R4) is run on a set of data that is best described by Barro, Mankiw, and Sala-i-Martin's (1995) model, least-squares estimation of parameter $\hat{\theta}_h$, and potentially of the other slope parameters is driven by measurement error and missing variable biases. It was for this reason that our earlier empirical studies on the role of human capital in Canadian provincial growth (Coulombe and Tremblay 2001; Coulombe 2003) focused either on the convergence of human capital where $\Delta h_{i,t} = f(h_{i,t-1},\ldots)$ or on the return of human capital derived from an econometric growth accounting framework described by equation (6).

In this paper, we report some results from convergence growth regressions based on equation (R4). We show that the point estimate of $\hat{\theta}_h$ is generally not significant. We also show that when the initial human capital variable is taken out of R4, it can be used as an efficient instrument for the initial per capita income in IV estimations.

3.4 Other details on estimation techniques

We use appropriate econometric techniques to tackle the time-series and cross-sectional heteroscedasticity problems underlying this type of TSCS analysis. A set of results comes from pooled least squares (PLS) for which we report White heteroscedasticity consistent standard errors (HCCME). A second set of results comes from iterated feasible generalized least squares (FGLS) estimations using cross-sectional weighted regressions to account for cross-sectional heteroscedasticity. For this set of results, we also report HCCME standard errors to allow for asymptotically valid inferences in the presence of the remaining time-series heteroscedasticity.

A third set of results was produced using system estimations with instrumental variables (IV) performed with two-stage least squares (TSLS) and weighted two-stage least squares (WTSLS), again to account for cross-sectional heteroscedasticity. For WTSLS, we used iterative techniques for updating coefficients and the weighting matrix. In all IV estimations, independent variables other than literacy are used as their own instrument. We report results for two alternative instruments used for the literacy variable. In the first set, to deal with a possible measurement error, we used the 1994 IALS data as the instrument for the ALL 2003 literacy data. In a second set of results, the lagged literacy data from ALL 2003 are used as the instrument for the contemporaneous variable. This procedure might help mitigate a possible endogeneity problem.[9] The skill of the young cohort might result from the current income level. Instrumenting with the lagged variable is a common way to cope with the possible reverse causation.

We use the variance components method (GLS) in the four sets of results for random-effects estimations. Furthermore, in the 5-year model for which serial correlation appears to be a serious problem, the disturbances were modelled in two regressions (referred as AR(1) in table 4) as serially correlated and cross-sectional heteroscedastic:

$$\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + u_{i,t}.$$

Finally, in some tables, we also report results on beta coefficients using standardized variables (for literacy, university achievement, per capital income, and the urbanization variable). These beta coefficients are reported within brackets in the same column as the associated usual estimation. As mentioned before in our discussion of reliability ratios, the beta coefficients are reported to allow a straightforward comparison between results dealing with literacy and university achievement since the literacy variable is measured on an arbitrary scale. To compute each standardized variable, we first calculated the standard deviation of the cross-sectional demeaned transformed variable in each time period. Second, we computed the time-series mean of the standard deviation series. Third, the cross-sectional demeaned transformed variable has been divided by this mean. With this procedure, the information contained in the time-series evolution of the cross-sectional variance (the sigma-convergence process observed for most variables) is not eliminated by the standardization process.

Given the limited number of time-series observations at hand, it was not possible to perform more general methods of estimation such as seemingly unrelated regression estimation (SUR) and three-stage least squares. Note that these estimation methods are not preferable to the one used in this paper; such estimation methods are known in TSCS analysis to minimize standard errors and lead to extreme confidence when the number of time series is limited (Beck and Katz 1995).

4. Results

4.1 Literacy or university achievement as indicator of human capital

Results of TSCS estimations of regression models R1, R2, and R3 when literacy, university achievement, or synthetic schooling data are used as indicator of human capital are presented in tables 2 to 5.

4.1.1 Estimates of human capital indicators

The effect of literacy is positive and significant at least at the 5 percent level in 18 of 20 TSCS regressions. In the other 2 regressions using PLS, the effect of literacy is positive and significant with a p-value around 10 percent.[10] Note that the effect of all human capital indicators is estimated less precisely with PLS than with GLS. Furthermore, the effect of university achievement is also positive and significant at least at the 5 percent level in the 5 TSCS regressions. Hence, human capital, whether measured from the literacy or university achievement aspect, appears to have a clear, positive effect on the relative level of per capita income across provinces. In our view, the fact that both indicators have positive and significant effects on per capita income, in contrast to our earlier cross-country study, is the main result of this paper.

Table 2

Fixed-effects estimations of regression model R1

		Dependent variables: per capita income minus government transfers							
		(1) PLS	(2) FGLS	(3) PLS	(4) FGLS	(5) PLS	(6) FGLS	(7) TSLS	(8) WTSLS
Literacy	Coefficient	1.05	1.35	d	d	d	d	3.03	1.81
	Standard error	0.70	0.46 a	d	d	d	d	1.27 b	0.81 b
	Beta	0.15	0.20	d	d	d	d	...	...
University achievement	Coefficient	d	d	0.24	0.23	d	d	d	d
	Standard error	d	d	0.09 b	0.08 a	d	d	d	d
	Beta	d	d	0.31	0.30	d	d	d	d
Schooling	Coefficient	d	d	d	d	0.36	0.67	d	d
	Standard error	d	d	d	d	0.09	0.22 a	d	d
	Beta	d	d	d	d	0.08	0.14	d	d
Adjusted R^2		.86	.88	.87	.90	.85	.88	...	...
Durbin-Watson		1.53	1.80	1.90	1.80	1.50	1.57	...	...

Notes: Ten-year periods resulting in 60 panel observations. In columns (7) and (8), the 1994 IALS data are used as IV for the ALL 2003 literacy data. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients. Beta coefficients are reported in the row below the standard errors of the coefficients.

a significant at 1% level;
b significant at 5% level;
c significant at 10% level;
d variable not in model.
... not applicable

From a quantitative point of view, the estimated effects of literacy and university achievement on per capita income are quite different. The point estimate of the non-standardized literacy variable is much larger than the point estimate of university achievement. However, in the five regressions using standardized variables (beta coefficients), the effect of university achievement on personal income is systematically larger (56 percent on average) than the effect of literacy. On average, a one standard deviation increase in literacy in one province translates into an increase of 0.19 standard deviations in personal income (minus government transfers). For university achievement, the increase is 33 percent.

Table 3
Estimations of structural regression model R2

		(1) PLS	(2) FGLS	(3) PLS	(4) FGLS	(5) PLS	(6) FGLS	(7) TSLS	(8) WTSLS
Literacy	Coefficient	1.04	1.80	d	d	d	d	1.28	1.92
	Standard error	0.52 c	0.35 a	d	d	d	d	0.49 b	0.28 a
	Beta	0.15	0.26	d	d	d	d	...	...
University achievement	Coefficient	d	d	0.27	0.26	d	d	d	d
	Standard error	d	d	0.12 b	0.08 a	d	d	d	d
	Beta	d	d	0.36	0.34	d	d	d	d
Schooling	Coefficient	d	d	d	d	0.54	1.39	d	d
	Standard error	d	d	d	d	0.53	0.15 a	d	d
	Beta	d	d	d	d	0.11	0.29	d	d
Urban	Coefficient	0.91	0.60	0.59	0.59	0.89	0.57	0.89	0.59
	Standard error	0.09 a	0.07 a	0.15 a	0.11 a	0.12 a	0.04 a	0.08 a	0.05 a
	Beta	0.83	0.55	0.54	0.55	0.82	0.53	...	...
Quebec	Coefficient	-0.11	-0.05	-0.09	-0.09	-0.13	-0.08	-0.11	-0.05
	Standard error	0.02 a	0.02 b	0.02 a	0.02 a	0.02 a	0.02 a	0.06 c	0.02 b
	Beta	...	...	...	...	...	...	...	...
Alberta	Coefficient	0.10	0.12	0.10	0.10	0.11	0.13	0.10	0.12
	Standard error	0.04 a	0.04 a	0.04 b	0.04 a	0.04 a	0.04 a	0.06	0.04 a
	Beta	...	...	...	...	...	...	...	...
Nova Scotia	Coefficient	-0.01	-0.05	-0.03	-0.03	-0.01	-0.05	-0.02	-0.05
	Standard error	0.03	0.02 b	0.02	0.02	0.03	0.04	0.05	0.02 b
	Beta	...	...	...	...	...	...	...	...
Adjusted R^2		.76	.85	.77	.88	.75	.87	...	...
Durbin-Watson		0.89	1.40	1.06	1.46	0.88	1.60	...	...

Notes: Ten-year periods resulting in 60 panel observations. In columns (7) and (8), the 1994 IALS data are used as IV for the ALL 2003 literacy data. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients. Beta coefficients are reported in the row below the standard errors of the coefficients.

a significant at 1% level;
b significant at 5% level;
c significant at 10% level;
d variable not in the model.
... not applicable

Table 4
Estimation results for the 5-year model, R1 and R2

		Dependent variables: per capita income minus government transfers									
		(1) PLS	(2) FGLS	(3) PLS	(4) FGLS	(5) TSLS	(6) WTSLS	(7) TSLS	(8) WTSLS	(9) PLS	(10) FGLS
Literacy	Coefficient	0.90	1.21	1.16	1.89	1.37	2.01	0.86	1.84	0.98	1.03
	Standard error	0.42 [b]	0.29 [a]	0.31 [a]	0.19 [a]	0.31 [a]	0.17 [a]	0.29 [a]	0.16 [a]	0.40 [b]	0.27 [a]
Urban	Coefficient	d	d	0.88	0.59	0.87	0.59	0.87	0.65	0.62	0.58
	Standard error	d	d	0.06 [a]	0.04 [a]	0.05 [a]	0.03 [a]	0.05 [a]	0.03 [a]	0.07 [a]	0.08 [a]
Quebec	Coefficient	d	d	-0.11	-0.04	-0.10	-0.04	-0.11	-0.06	-0.05	-0.05
	Standard error	d	d	0.02 [a]	0.01 [a]	0.04 [b]	0.01 [a]	0.04 [a]	0.01 [a]	0.02 [b]	0.02 [b]
Alberta	Coefficient	d	d	0.09	0.11	0.09	0.11	0.10	0.10	0.12	0.12
	Standard error	d	d	0.02 [a]	0.02 [a]	0.04 [b]	0.03 [a]	0.04 [a]	0.02 [a]	0.05 [b]	0.05 [b]
Nova Scotia	Coefficient	d	d	-0.02	-0.05	-0.02	-0.05	-0.01	-0.04	-0.03	-0.04
	Standard error	d	d	0.01	0.01 [a]	0.03	0.01 [a]	0.03	0.01 [a]	0.01 [b]	0.01 [a]
Fixed-effects estimations		Yes	Yes	No	No	No	No	No	No	No	No
AR(1) estimations		No	No	No	No	No	No	No	No	Yes	Yes
Adjusted R^2		.89	.90	.79	.86	...	...	...	...	.94	.95
Durbin-Watson		0.90	1.04	0.48	0.88	...	...	...	...	2.44	2.50

Notes: Five-year periods resulting in 110 panel observations in columns (1) to (6), and 100 panel observations in columns (7) to (10). In columns (5) and (6), the 1994 IALS data are used as IV for the ALL 2003 literacy data. In columns (7) and (8), the lagged literacy data from ALL 2003 is used as instrument for the contemporaneous variable. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients.

a significant at 1% level;
b significant at 5% level;
c significant at 10% level;
d variable not in model.
... not applicable

However, the effect of the synthetic schooling variable constructed from the ALL 2003 survey is always positive but is not significant at the 10 percent level in three out of five regressions. This lack of significance, compared with the data based on university attainment and literacy, might be interpreted in at least two ways. First, literacy and university attainment matter more than years of schooling. This is the straightforward interpretation if one abstracts from measurement errors in human capital. Second, the signal contained in the synthetic literacy data is correlated more with per capita income than are the synthetic schooling data derived using the same methodology and survey. This might be because, among other things, a substantial proportion of the population in many provinces are foreign-born and have completed most of their schooling in their country of origin. If, as the conclusion of this study suggests, years of schooling are less comparable across countries than literacy, the literacy data within Canadian provinces might also be more comparable than years of schooling data. Furthermore, the methodology for deriving synthetic cohorts might introduce measurement errors in both the synthetic literacy and schooling data. The attenuation bias pushes down the point estimates of both slope coefficients but literacy remains significant because of its higher human capital signal. This also explains why the university attainment data are highly significant since their time-series dimension is not derived from the demographic profile.

From a methodological point of view, a few additional results are worth noting. First, in all six regressions where literacy data from the 1994 survey are used as instruments for the 2003 data (tables 2, 3, and 4, comparing columns (1) with (5) and columns (2) with (6)), the point estimate of the literacy variable is larger with IV. This can be interpreted as meaning that the use of another indicator of human capital as instrument decreases the attenuation bias resulting from the presence of measurement errors. It is important to note that (non-reported) results with IV estimations where the university attainment variable is used as instrument for the literacy variable yield very similar results.

Second, the effect of literacy appears to be consistently estimated across the 5-year and the 10-year data sets. The estimations seem more precise in the 5-year set-up (all literacy coefficients are significant at least at the 5 percent level even using PLS and TSLS).

Third, the positive effect of literacy remains highly significant when its lagged value is used as instrument in columns (7) and (8) of table 4. Comparing columns (5) with (7) and columns (6) with (8) in table 4, we see that the lagging procedure decreases the point estimate of the literacy variable slightly when comparable IV estimation techniques are used. These results indicate that the potential reverse causation problem (endogeneity) is not driving the positive results of the synthetic literacy variables.

Fourth, random-effects estimations (table 5) produce very similar results for the literacy and university achievement variables than in regression models R1 and R2. However, the effect of the synthetic schooling variable derived from ALL 2003 is not close to being significant with the random-effects model.

Table 5

Random effect estimation of model R3

		Dependent variables: per capita income minus government transfers			
		(1)	(2)	(3)	(4)
Literacy	Coefficient	1.13	d	d	0.99
	Standard error	0.53 b	d	d	0.36 a
	Beta	0.17	d	d	...
University achievement	Coefficient	d	0.25	d	d
	Standard error	d	0.08 a	d	d
	Beta	d	0.33	d	d
Schooling	Coefficient	d	d	0.42	d
	Standard error	d	d	0.33	d
	Beta	d	d	0.09	d
Urban	Coefficient	0.89	0.62	0.89	0.88
	Standard error	0.15 a	0.17 a	0.16 a	0.14 a
	Beta	0.83	0.58	0.82	...
Adjusted R^2		.86	.87	.85	.89
Durbin-Watson		1.35	1.66	1.31	0.85
Observations		60	60	60	110

Notes: Estimation method is GLS (variance component). Ten-year periods in columns (1) to (3), five-year periods in column (4). White heteroscedasticity standard errors are shown in the second row below the estimated coefficients. Beta coefficients are reported in the row below the standard errors of the coefficients.

a significant at 1% level;
b significant at 5% level;
c significant at 10% level;
d variable not in model.
... not applicable

Finally, the mean point estimate of the literacy variable is 1.43 across the 20 TSCS estimations in tables 2 to 5. This number is the elasticity of per capita income (minus government transfers to individuals) to the literacy variable. It can be interpreted in terms of the microeconomic returns to education estimated in Mincerian labour studies. According to Psacharopoulos (1994), the individual return on wages of one extra year of education ranges from 5 to 15 percent. According to OECD (2000, p. xiv), one additional year of education increases the literacy score by 10 points. This is 3.8 percent of the mean literacy score (265) of the young cohorts across the provinces in our sample. Consequently, an increase in skills corresponding to one additional year of schooling increases per capita income at the provincial level in Canada by 5.4 percent. One can argue that this number is pushed down by the attenuation bias resulting from measurement error as pointed out before. If this is the case, we could rely on the six point estimates for which the IALS (1994) data are used as instruments for the 2003 literacy data (columns (7) and (8) from tables 2 and 3, and columns (5) and (6) from table 4). The mean of the six point estimates for the literacy variable is 1.90 for these IV estimations. This corresponds to a macroeconomic return to education of 7.3 percent. Interestingly, the number falls in the middle of the individual returns summarized by Psacharopoulos (1994).

4.1.2 The urbanization structure

In all cases (20 regressions in tables 3 to 5), the coefficient on the urbanization variable is positive and significant at the 1 percent level. The point estimates are also remarkably stable. In all estimations where university achievement is used as the indicator of human capital, and in all estimations using the literacy or the synthetic schooling data with FGLS and WTSLS, the point estimates of the urbanization variable are around 0.60. In all estimations using the literacy or the synthetic schooling variables with PLS, TSLS, and random effects, the point estimates are a little higher, around 0.90.

From a quantitative point of view, these results concur with Coulombe's (2000) estimate of 0.78 for the long-run elasticity of per capita income (minus government transfers) to urbanization. Our previous results were obtained in a dynamic framework such as R4 in which the initial human capital variable was not entered in the list of controls for the reasons argued above. Our present results imply that, after having accounted for human capital accumulation and other factors, a province that is 10 percent more urbanized than the Canadian mean will be between 6% and 9% richer.

Also on quantitative grounds, the estimated beta coefficients for the urbanization variable (in tables 3 and 5) are systematically larger than the estimated beta coefficients for both university achievement and literacy. The mean estimate (across the six regressions) of the beta coefficient of urbanization is 0.65 (against 0.33 and 0.19 for university achievement and literacy respectively).

Interestingly, the results regarding the urbanization variable with random-effects estimations in table 5 also concur with the ones obtained from the structural model R2 in tables 3 and 5. This indicates that the effect of urbanization remained roughly unchanged even when the provincial heterogeneity is modelled in a less constrained (and rather ad hoc) way than with the structural break and dummy variable approach taken in the structural model R2.

The results regarding urbanization also shed light on the estimated effect of human capital in this analysis. In estimating the partial effect of human capital in this study, we control for the positive correlations between urbanization, human capital, and per capita income. Without controlling for urbanization (which is also done in the fixed-effects modelling, given that relative urbanization is stable through time), the effect of human capital would be positively biased as in the pure cross-section that is used as an illustration in the highlights.

4.1.3 Provincial dummies

Results for the provincial dummies (Quebec, Alberta, and Nova Scotia) are presented for 16 regressions in tables 3 and 4. The first 8 regressions (table 3) pertain to the 10-year model and the last 8 (table 4) to the 5-year model.

In the 10-year model (table 3), Quebec and Alberta's structural breaks are both modelled to arrive in period 3 (1970). In the 5-year model, the Quebec shock occurs in period 5 (1970) and the Alberta shock in period 6 (1975). In Coulombe (2000), which used a TSCS model with yearly data and corrected specifically for serial correlation, the dates of those shocks were chosen in order to maximize the *t-statistics* of the dummy. The 5-year set-up is more flexible when it comes to modelling a break that occurs close to the middle of a decade. This emerges, to a certain extent, in the TSLS estimates (in both column (7) of table 3 and column (5) of table 4). For this estimation method of the structural model, Quebec's shock is significant only at the 10 percent level in the 10-year model whereas Alberta's shock is not even significant at 10 percent. In the 5-year set-up, however, both shocks are significant at the 5 percent level. In all other regressions, both shocks are significant at least at the 5 percent level.

On quantitative grounds, point estimates of the Alberta's oil shock are very comparable across the 5-year and the 10-year set-up and vary between 0.09 and 0.12. They are a little lower than the 15.4 percent long-run elasticity estimated in Coulombe (2000) for the same per capita income concept in a dynamic convergence model analog with R4 and using annual data.

The size of the point estimates of the negative 1970 Quebec shock varies across estimation methods. With weighted least squares (FGLS and TSLS), when human capital is captured only by literacy in both the 5-year and the 10-year set-up, the point estimates vary between –0.04 and –0.05 percent. In the other regressions (all regressions with university achievements, and regressions using PLS and TSLS for literacy), the effect of Quebec's shocks is doubled and varies between –0.09 and –0.12. The latter estimates are closer to the long-run elasticity of 10.3 percent estimated in Coulombe (2000) for this variable in his conditional-convergence framework.

The results regarding the Nova Scotia fixed effect are much less robust than the ones for the Quebec and Alberta dummies. The negative coefficient is not significant at the 10 percent level in 7 of 16 regressions. These results concur to some extent with Coulombe (2003) who found that educated Nova Scotians tend to remain in their province even if it is less urbanized and relatively poorer than the Canadian average. In the other three relatively poor provinces of Atlantic Canada (Newfoundland, Prince Edward Island, and New Brunswick), the educated people show a greater tendency to migrate to richer provinces (such as Ontario and Alberta). Interestingly. In the present study, all significant results for Nova Scotia fixed effects are found when the literacy variable is used as the human capital indicator.

4.2 Male versus female literacy

In Coulombe, Tremblay, and Marchand's 2004 cross-country study of 14 OECD countries, human capital indicators based on female literacy systematically outperformed those based on male literacy in a convergence regression set-up analog to R4. This result is rather interesting since it is robust to controlling for cross-country differences in the female labour force participation rate. Point estimates of female indicators are larger than those for the male population; the results are more robust with indicators based on the female population; and perhaps more interestingly, in contrast to indicators based on female literacy, the estimated coefficient for the male population is generally not significant when the literacy of both sexes is entered separately in the list of controls. A number of possible explanations were proposed in Coulombe, Tremblay, and Marchand (2004) for the gender-specific effect. They suggested that an analysis performed at the regional level, such as the one done in the current paper, might clarify which one is the most suitable.

Results for the comparison of the effect of male and female literacy in the structural model R2 are shown in table 6. The general direction of results for other specifications concurs with the ones reported. In the results reported in columns (1) and (2) (for PLS) and (4) and (5) (for FGLS), the indicators based on female and male literacy are entered in the list of controls in separate regressions. In the results reported in columns (3) (for PLS) and (6) (for FGLS), both indicators are entered separately in the list of controls in the same regression. The effects of the gender-specific indicator are all significant at the 1 percent level when the indicators are entered in separate regressions. In this case, the point estimate for the female population is a little higher with FGLS but we find the reverse in PLS estimations. When the indicators based on both sexes are entered separately in the same regression, the effect of male literacy is larger with PLS and significant at the 5 percent level. In this case, the effect of female literacy is not significant even at the 10 percent level with PLS. Again in this case, with FGLS the effect of both indicators is significant at the 1 percent level and the effect of female literacy is a little higher (although not significantly) than male literacy. Overall, we find no systematic evidence of a gender-specific effect in our regional data set.

Table 6

Female versus male literacy

		Dependent variables: per capita income minus government transfers					
		(1) PLS	(2) PLS	(3) PLS	(4) FGLS	(5) FGLS	(6) FGLS
Male literacy	Coefficient	1.09	d	0.90	1.17	d	0.86
	Standard error	0.31 a	d	0.42 b	0.26 a	d	0.24 a
Female literacy	Coefficient	d	0.87	0.27	d	1.47	1.01
	Standard error	d	0.30 a	0.40	d	0.21 a	0.18 a
Urban	Coefficient	0.84	0.93	0.85	0.72	0.71	0.60
	Standard error	0.06 a	0.05 a	0.06 a	0.06 a	0.04 a	0.05 a
Quebec	Coefficient	-0.12	-0.09	-0.11	-0.11	-0.05	-0.04
	Standard error	0.01 a	0.03 a	0.02 a	0.02 a	0.01 a	0.01 a
Alberta	Coefficient	0.10	-0.11	0.09	0.09	0.11	0.11
	Standard error	0.02 a	0.02 a	0.02 a	0.01 a	0.03 a	0.03 a
Nova Scotia	Coefficient	-0.03	-0.01	-0.03	-0.04	-0.03	-0.05
	Standard error	0.02 b	0.01	0.02	0.02 a	0.02 c	0.02 a
Adjusted R^2		.80	.79	.80	.84	.83	.85
Durbin-Watson		0.54	0.46	0.51	0.78	0.84	0.88

Notes: Regression model R2. Five-year periods resulting in 110 panel observations. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients.

a significant at 1% level;

b significant at 5% level;

c significant at 10% level;

d variable not in model.

The differences in the results for the gender-specific effects in our Canadian provincial data set and in the cross-country data set used in Coulombe, Tremblay, and Marchand (2004) help support one of the candidate explanations for this effect proposed by Coulombe, Tremblay and Marchand (2004). In cross-country studies, an indicator based on female literacy can capture the effects of omitted variables such as the one labelled *social infrastructure* by Hall and Jones (1999) or, more generally, the level of social development of a country. Hall and Jones (1999) argued that differences in social infrastructure account for the large differences in per capita GDP levels across countries that cannot be accounted for by differences in human and physical capital (Solow residual) across developed and less-developed countries. As pointed out in our introduction, the concept of social infrastructure is very hard to measure adequately. In cross-country growth regressions, it is usually left in the residual or accounted for by fixed effects. In a cross-country panel data set, the relative evolution of social infrastructure through time across countries can be positively correlated with the evolution of the ratio of female to male literacy. In our Canadian provincial data set, one can argue that social infrastructure differences across provinces are not significant, or that the few differences are captured by the Nova Scotia fixed effect and Quebec dummy variables. The evolution of social infrastructure that is correlated with the evolution of the female/male literacy index might have affected the 10 provinces evenly. If so, this common effect would be eliminated by the cross-sectional demeaning procedure.

The above highlights the general interest in performing empirical analyses using regional data but the interest is not limited to regional studies. The use of relatively homogeneous regional data can also shed light on important aspects of cross-country studies.

4.3 Combining literacy and university achievement indicators

Up to now, the literacy and the university achievement indicators were viewed as alternative proxies of the true human capital concept. Since both indicators are not perfectly correlated, as illustrated by the reliability ratio results presented in table 1, we attempted to verify in numerous empirical investigations whether both indicators could be combined in different ways. In table 7, we report some key results using specifications R1, R2, and R3 where both indicators, or a combination of the two, are entered in the list of controls in the same regression.

Table 7

Combining the literacy and the lagged university achievement indicators

		Dependent variables: (standardized) per capita income minus government transfers				
		(1) R1	(2) R1	(3) R1	(4) R2	(5) R3
Literacy	Coefficient	0.20	0.14	d	d	d
	Standard error	0.08 b	0.04 a	d	d	d
University achievement	Coefficient	0.17	d	d	d	d
	Standard error	0.10	d	d	d	d
Lagged university	Coefficient	d	0.18	d	d	d
	Standard error	d	0.05 a	d	d	d
Combined index	Coefficient	d	d	0.16	0.19	0.18
	Standard error	d	d	0.03 a	0.04 a	0.05 a
Observations		60	50	50	50	50
Adjusted R^2		.89	.93	.93	.92	.91
Durbin-Watson		1.94	2.08	2.03	1.63	1.28

Notes: Ten-year model (R1, R2, and R3) resulting in 60 panel observations. Estimation methods are FGLS (columns 1 to 4) and GLS variance component (column 5). Statistics on structural parameters not shown in columns (4) and (5). The combined index of human capital is the sum of the (standardized) current literacy variable and the lagged university education variable. Beta coefficients for human capital and urbanization in all regressions. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients.

a significant at 1% level;

b significant at 5% level;

c significant at 10% level;

d variable not in model.

Using fixed-effects estimation and FGLS, both the contemporary literacy and the university achievement variables are entered separately in the list of controls in the first regression (column 1). The point estimates (beta coefficients) of both variables are of the same magnitude and the difference between the two is not statistically significant. The coefficient on literacy is significant at the 5 percent level but the coefficient on university achievement is not significant (marginally) at the 10 percent level. Using the same estimation techniques, both estimates are significant at the 1 percent level when the university achievement variable is lagged one period in the second regression (column 2). With the exception of the regressions reported in columns (7) and (8) for which the adjusted R square has been inflated by the AR correction, the adjusted R square (0.93) for this regression is larger than in any previous regressions where the human capital variable is proxied by either the literacy or the university achievement indicators.

Since the point estimates of both human capital indicators in the second regression are again of the same magnitude (and the difference is not significant), we combined the two variables by simply adding the current literacy to the lagged university achievement variable. The results are reported in columns (3) to (5) for this combined index for regression models R1, R2, and R3. Results for the other variables in the structural model and the random-effects model are very comparable with those reported for the previous regressions and are not reported in table 7. In the three cases, the reported coefficients of the combined index are significant at the 1 percent level and the point estimates are very comparable with beta coefficients varying between 0.16 and 0.19. Again, with the exception of the regression using the AR correction, the adjusted R squares are again higher (between 0.91 and 0.93) than those reported in the previous regressions. Finally, with the exception of the regression using random effects, analysis of the Durbin Watson statistic indicates that serial correlation is not a serious problem when literacy is combined with university achievement, whether the latter variable is lagged or not.

These results suggest that the effect of literacy on per capita income is positive and significant, even after controlling for university achievement. The effect of university achievement is also positive and significant when the variable is lagged, even after controlling for literacy. Furthermore, both indicators can be combined in a human capital index that performs extremely well in the three regression models. In a nutshell, both literacy and university achievement appear to matter.

4.4 Results of convergence regressions

Results for the convergence regression model R4 are shown in table 8. For the regression results in column (1), we used the same modelling as in Coulombe (2000) but our data set was extended to 2001. On both qualitative and quantitative grounds, the results of the first regression concur with our earlier studies. The estimated convergence speed is around 5 percent per year. The estimated long-run elasticity (0.69) of the urbanization variable is in the lower middle of the range of those estimated in static models R2 and R3. The long-run elasticity of Quebec's shock (-0.10) is in the upper end of the range of estimates found in our static regressions. The long-run elasticity of Alberta's oil shock variable is higher than in the static regressions but the coefficient is estimated imprecisely.

Table 8

Estimation results for the convergence regression model R4

		Dependent variables: mean annual growth rate of per capita income minus government transfers		
		(1) FGLS	(2) FGLS	(3) WTSLS
Lagged per capita income	Coefficient	-0.051	-0.054	-0.039
	Standard error	0.009 [a]	0.010 [a]	0.008 [a]
	Long-run elasticity	...	...	...
Lagged literacy	Coefficient	[d]	0.009	[d]
	Standard error	[d]	0.031	[d]
	Long-run elasticity	[d]	0.17	[d]
Urban	Coefficient	0.040	0.040	0.025
	Standard error	0.010 [a]	0.010 [a]	0.010 [a]
	Long-run elasticity	0.69	0.68	0.66
Quebec	Coefficient	-0.005	-0.005	-0.003
	Standard error	0.002 [a]	0.002 [b]	0.002 [c]
	Long-run elasticity	-0.10	-0.09	-0.08
Alberta	Coefficient	0.010	0.010	0.009
	Standard error	0.006 [c]	0.006 [c]	0.005 [b]
	Long-run elasticity	0.19	0.19	0.24
Adjusted R^2		.35	.33	...
Durbin-Watson		2.57	2.56	...

Notes: Five-year regression model resulting in 100 panel observations. In column (3), the lagged literacy variable is used as instrument for the lagged per capita income variable. White heteroscedasticity standard errors are shown in the second row below the estimated coefficients. Long-run elasticity computed from equation (7) is shown in the row below the standard errors of the coefficients.

a significant at 1% level;
b significant at 5% level;
c significant at 10% level;
d variable not in model.
... not applicable

In the second regression, the lagged human capital variable based on literacy is added to the list of controls in the conditional convergence model as was done in many cross-country studies including the one by Coulombe, Tremblay, and Marchand (2004). As discussed in our earlier research on the topic (Coulombe 2000; Coulombe and Tremblay 2001; Coulombe 2003), the results shown in column (2) clearly illustrate that the effect of human capital on per capita income (or per capita GDP) cannot be estimated efficiently when the lagged human capital variable is entered in the list of controls in a conditional convergence regression model of per capita income. As pointed out in our discussion in Section 3.3, in an open-economy growth model with perfect capital mobility and a binding constraint for the financing of human capital, the initial level of per capita income is a log-linear function of the initial level of human capital. If both variables are entered in the list of controls in the same regression, the strong collinearity between the two variables implies that the estimate of the human capital variable will be driven by measurement error and omitted variable biases.

If this interpretation is taken seriously, the lagged human capital variable might be used as an instrument for the lagged per capita income in IV estimations of the convergence growth regression. From an econometric point of view, using human capital as instrument in this set-up was motivated by the possibility that it might lessen the tendency in dynamic regressions to overestimate the effect of the lagged dependant variable because of measurement error in the per capita income variable. The problem will be alleviated if the measurement error in the human capital variable is not correlated with the measurement error in per capita income. This assumption is clearly verified if literacy is used as a proxy for human capital.

Results of IV estimations for which the lagged literacy is used as an instrument for the lagged per capita income are shown in column (3). The regression model performs extremely well. The convergence speed is still significant at the 1 percent level but its points estimate is lower (around 4 percent) than the one reported in column (1). The long-run elasticities of the three structural variables are also the same magnitude as the ones reported in table 1. Literacy is so intrinsically related to per capita income that literacy could be viewed as an efficient instrument in solving problems related to measurement of error bias in per capita income.

5. Conclusions

The empirical analysis presented in this paper consolidates our understanding of the regional growth process. In particular, it shows that the accumulation of human capital has been an important determinant of the relative growth of Canadian provinces over the 1951–2001 period. The relative degree of urbanization of provinces as well as specific shocks in Alberta (the first oil shock in 1973) and Quebec (the Anglophone exodus around 1970) also played a significant role. Our mean estimate of the macroeconomic return on one additional year of education, in terms of skills acquired, is 7.3 percent in IV estimations designed to correct for the attenuation bias caused by measurement error. This number falls to the middle of the 5 to 15 percent range for the estimates of individual return shown in Mincerian studies.

More importantly, our analysis makes an important contribution to the literature on measurement of human capital and on macroeconomic returns from human capital accumulation. It compares in a regional context the effects of two different types of human capital indicators based on university attainment and literacy test scores, respectively. The main insights gained from our analysis are the following.

First, literacy indicators do not appear to outperform schooling indicators based on university attainment at the Canadian *provincial* level. This contrasts sharply with the main result of the Coulombe, Tremblay, and Marchand (2004) study based on cross-country regressions. This is potentially the most novel result of the current analysis. It suggests that literacy indicators outperform schooling indicators at the *cross-country* level because literacy test scores are more comparable than years of schooling. It is important to note that in the present study, the schooling indicator is not based on reported or computed years of schooling but on university achievement. It is also possible that the relative performance of university achievement data in our study comes from the fact that they are reported in census data in a consistent way through time in Canada (see de la Fuente and Domenéch [2002] on this). At the cross-country level, however, data on years of education are often derived from raw data on benchmark educational level using correspondence that might not be consistent through time. The use of years of schooling at the cross-country level as the indicator of human capital — possibly motivated by the desire to link the macro results to the micro Mincerian literature that has focused on the return on an extra year of schooling — might increase measurement errors.

Secondly, to the extent that institutional quality and levels of social infrastructure are similar across Canadian provinces, our regional analysis allows us to derive an unbiased estimate of the contribution of human capital to relative standards of living across economies. In cross-country studies, one can reasonably argue that estimates of the contribution of human capital to income per capita may capture, at least to some extent, differences in the level of social infrastructure. Hence, our results shed new light on this issue and provide support to the view that human capital does matter significantly for the relative long-run well-being of developed economies.

Third, there is no systematic evidence of gender-specific effects of literacy on provincial income per capita. This suggests that the much stronger impact of female literacy on standards of living across countries (found in Coulombe, Tremblay, and Marchand [2004]) may reflect differences in the levels of social infrastructure or social development across countries, variables that may well be correlated with the ratio of female-to-male literacy.

As mentioned earlier, one of the main limitations of the literacy indicator that we use is that it does not take into account the migration flows that occurred over the period. An important avenue for future research would be to construct indicators of human capital transfers across Canadian provinces based on the literacy test scores of migrants. Relative to existing measures of migration flows across provinces, such indicators would take into account the skill level of migrants. These indicators would also allow a better understanding of the determinants of the skilled labour migration and of the effect of such migration on the convergence of standards of living across regions.

References

Barro, R.J. 2001. Education and economic growth. In Helliwell, J.F., ed., *The contribution of human and social capital to sustained economic growth and well-being.* Organization for Economic Cooperation and Development / Human Resources Development Canada, Paris, pp. 13-41.

Barro, R.J., and Lee, J-W. 2000. International data on educational attainment: updates and implications. Working Paper 7911, National bureau of Economic Research, Cambridge, MA.

Barro, R.J., Mankiw, G., and Sala-i-Martin, X. 1995. Capital mobility in neoclassical models of growth. *American Economic Review* 85(1): 103-115.

Barro, R.J., and Sala-i-Martin, X. 1992. Convergence. *Journal of Political Economy* 100(2): 223-251.

Barro, R.J., and Sala-i-Martin, X. 1995. *Economic Growth.* McGraw-Hill, New York.

Beck, N., and Katz, J.N. 1995. What to do (and not to do) with time-series cross-section data. *American Political Science Review* 89(3): 634-647.

Baumol, W.J. 1986. Productivity growth, convergence, and welfare: what do the long-run data show? *American Economic Review* 76(5): 1072-1085.

Coulombe, S. 2000. New evidence of convergence across Canadian provinces: the role of urbanization. *Regional Studies* 34(8): 713-725.

Coulombe, S. 2001. The contribution of human capital and urbanization to Canadian regional growth. In Helliwell, J.F., ed., *The contribution of human and social capital to sustained economic growth and well-being.* Organization for Economic Cooperation and Development / Human Resources Development Canada, Paris, pp. 203-220.

Coulombe, S. 2003. Human capital, urbanization and Canadian provincial growth. *Regional Studies* 37(3): 239-250.

Coulombe, S., and Lee, F.C.. 1995. Convergence across Canadian provinces, 1961 to 1991. *Canadian Journal of Economics* 28(4a): 886-898.

Coulombe, S., and Tremblay, J.-F. 2001. Human capital and regional convergence in Canada. *Journal of Economic Studies* 28(3): 154-180.

Coulombe, S., Tremblay, J.-F., and Marchand, S. 2004. Literacy scores, human capital and growth across fourteen OECD countries. International Adult Literacy Survey. Statistics Canada, Human Resources and Skills Development Canada, Ottawa. Cat. No. 89-552-MIE, no. 11.

De la Fuente, A., and Doménech, R. 2002. Human capital in growth regressions: how much difference does data quality make? An update and further results. Manuscript. Instituto de Análisis Económico (CSIC) and Universidad de Valencia.

Glaeser, E., La Porta, R., Lopes-de-Silanes, F., and Shleifer, A. 2004. Do institutions cause growth? *Journal of Economic Growth* 9(3): 271-303.

Hall, R., and Jones,. C. 1999. Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114(1): 83-116.

Islam, N. 1995. Growth empirics: a panel data approach. *Quarterly Journal of Economics* 110(4): 1127-1170.

Kaldor, N. 1963. Capital accumulation and economic growth. In Lutz, F., and Hague, D. eds., *Proceedings of a conference held by the International Economics Association*. Macmillan, London.

Krueger, A.B., and Lindahl, M. 2001. Education for growth: why and for whom? *Journal of Economic Literature* 39(4): 1101-1136.

Lange, F., and Topel, R. 2004. The social value of education and human capital. Mimeo. Yale University and University of Chicago.

Lucas, R.E., Jr. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22(1): 3-42.

Mankiw, N.G., Romer, D., and Weil, D.N. 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107(2): 407-437.

OECD (Organization for Economic Cooperation and Development). 2000. Literacy in the information age. Final report of the International Adult Literacy Survey. Organization for Economic Cooperation and Development / Statistics Canada, Paris.

Psacharopoulos, G. 1994. Returns to investment in education: a global update. *World Development* 22: 1325-1343.

Romer, P.M. 1986. Increasing returns and long-run growth. *Journal of Political Economy* 94(5): 1002-1037.

Temple, J. 2000. Growth effects of education and social capital in the OECD countries. Economics Department Working Paper No. 263, OECD, October.

Topel, R. 1999. Labor markets and economic growth. In Ashenfelter, O., and Card, D., eds., *Handbook of Labor Economics, volume 3C*, North-Holland, Amsterdam, pp. 2943-2984.

Endnotes

1. In an open-economy growth model with physical capital mobility (Barro, Mankiw, and Sala-i-Martin 1995), investment in physical capital is driven by human capital investment. R&D is also a potential candidate to account for differences in cross-country productivity.
2. It is worth remembering that Hall and Jones (1999, p. 84) define social infrastructure as " the institutions and government policies that determine the economic environment within which individuals accumulate skills, and firms accumulate capital and produce output."
3. They are also available at 5-year intervals since 1976.
4. For a survey of empirical studies dealing with the growth effects of human and social capital across OECD countries, refer to Temple (2000). Also see Krueger and Lindahl (2001) for a broader discussion of empirical studies dealing with human capital and education.
5. Coulombe (2000, 2003) use the term *Anglophone exodus* to describe the shift from Montreal to Toronto of high level tertiary employment.
6. The other measure of education attainment deals with the attainment of grade 9. However, given that virtually 100 percent of the population attained this benchmark level of education in most provinces by the middle of our sample, we cannot use this indicator in the current study as a proxy of relative human capital across provinces.
7. There were 20,451 respondents in the IALS 2003 Canadian survey. In 1994, the number was 5,660. In Prince Edward Island, the number of respondents was 93 in 1994 (1.6 percent of the total Canadian number) and 645 in 2003 (3.2 percent of the total number).
8. See Topel (1999, Section 3.4).
9. We thank Angel de la Fuente for having suggested this point.
10. The literacy variable is not significant at least at the 10 percent level only in the case of column 1 in table 2. In this case, the p-value is 0.13.