

Comments on
“The Compelling Effects of Compulsory Schooling: Evidence from Canada”
by Philip Oreopoulos

This paper follows a series of recent (and less recent papers) that have looked at the effects of compulsory schoolings laws on education, earnings, and other outcomes. One important contribution of this line of research has been to produce relatively convincing estimates of the causal effect of education on earnings and other outcomes. Most of the existing work is based on US data, though there is also some interesting work for the UK. Viewed from this angle, the main contribution of the paper is to conduct a relatively similar study for Canada. The main findings are relatively similar to what has been found for the US and UK.

While the paper generally uses a sound empirical approach, I am not convinced that the main findings are as compelling as the author claims. I was also annoyed by a large number of typos, mislabeling of tables, etc. It makes it look as if the paper had been written a little too quickly.

The author wins points for honesty by showing the detailed trends in educational achievement by province (Figure 1) that he attempts to explain using the compulsory education variables in Figure 2. It is difficult to see, however, any visual impact of compulsory schooling in these figures. The author acknowledges this difficulty and mentions at the bottom of page 10 that “...We should not expect large discontinuities because the education attainment variable used here (grade attainment) does not correspond precisely with the law changes...” I don’t fully understand this point. It seems that despite these issues we should see a relatively clear inter-cohort break. More importantly, however, I don’t find they evidence that compulsory attendance laws have a significant effect on grade attainment that convincing. Here’s a couple of suggestion on how to improve this part of the paper:

1. It is clear from figure 1 that the provinces that used to have low levels of schooling (Newfoundland, Quebec, New Brunswick and PEI) have pretty much converged to the national average for younger cohorts. These provinces have also converged to the national average in terms of the compulsory schooling laws. So the big question is whether the convergence is due to the laws or to other factors. One obvious ways of checking for this is to add province-specific trends and see whether results are robust. Law changes should generate discrete increases that should be identifiable from general trends. Finding that the results are robust to the inclusion of province-specific trends would make the paper much more convincing.
2. The author should discuss more specifically the sources of identification. Looking at Figure 1 suggests that only 8 changes in minimum school leaving age are used to identify the model for cohorts age 14 between 1920 and 1970. For instance, there are no changes between 1920 (or perhaps 1921) and 1970 in Ontario, British Columbia and Nova Scotia which means these provinces are not used at all in the identification of the school leaving age effects.

Perhaps it would be helpful to show more explicitly how the 8 changes in laws contribute to the identification of the 3 parameters (effect of the 14 years, 15, years, and 16 years dummy). For example, only PEI, NB and Newfoundland had compulsory schooling below age 14 in 1920. This means that these provinces early in the 20th century play a critical role in the effect of compulsory schooling (relative to the base case of less than 14 years). It would be interesting to estimate separately (using dummies) the effect of each of the 8 law changes to show which ones go in the right direction (increase education significantly) and which cause problems. It would really help to clarify this point since the finding that increasing compulsory schooling from age 15 to 16 decreases educational achievement raises serious questions about the validity of the whole approach.

My other main point has to do with specification issues related to the cells approach used by the author. From what I understand, the author constructs cell means by pooling information from the same cohort (in a given province) for several censuses. This leads to an unusual specification that controls for birth cohort effects, census year effects, but not age or experience. There are two problems with this specification choice. First, it does not allow for an explicit effect of age in the educational achievement models. This is a problem since average education keeps going up steadily in the early twenties and more slowly later one. I would suggest using a quadratic or cubic specification in age in the educational achievement models to take account of this factor. Though the linear in age is not identified separately from cohort and census year effects, the higher order terms will be identified. The same point applies to the earnings (or income) models. Here the standard approach is to use potential experience as regressor but one could also try to run the model with controls for age (up to a cubic or quartic).

The other specification issue is that the OLS estimates from the cell models are not generally comparable to the usual cross-sectional OLS estimates. The cell means approach completely discards the covariation between education and earnings within a given cohort-province-year cell. As a result, the OLS estimates reported here are really “between-group” estimates. They can be larger or smaller than the usual OLS depending on circumstances. For example, when education is measured with error, the OLS model on cells can be viewed as a Wald estimator that yields consistent estimates of the return to education that are larger than the standard OLS. Indeed, I find the OLS estimates reported in Table 5-7 quite large relative to the existing literature. This may explain why, contrary to what the author suggests on the bottom of page 19, the IV estimates (for earnings in the full sample) are generally lower than the OLS.

This whole part of the paper has to be reworked substantially. The list of “to-do” things includes: A) the author needs to explain clearly what his OLS estimator really estimates. B) the standard cross-sectional OLS estimates need to be reported (public use files can be used for this). C) standard controls for age and experience need to be included.

Other points:

1. On page 15 it is said that “As shown in the next tables, this finding (that the age 15 effect is larger than the age 16 effect) may, in part, be due to the enrolment patterns in Quebec after 1950.” I do not quite see this point since the compulsory age does never get to 16 (in Quebec) during the sample period. Also, I kept looking for a discussion of this point around the “next tables” but never found it. This has to be clarified.
2. On a related point, I am surprised to see that there was a compulsory schooling age in Quebec as early as Figure 1 seems to indicate. The common wisdom is that compulsory schooling was only introduced in the “*Loi de fréquentation scolaire obligatoire de 1943*”. Thanks to Google and its automatic english translation, number of easy to check references clearly suggest that a compulsory schooling age of 14 was only introduced in 1943 (so that the first cohorts affected were those born in 1929 or so). For example, the web site of the Quebec National Assembly tells the story of a debate on this in 1912 and explains why it took until 1943 before the law was passed, while France and Ontario passed similar laws in 1882 and 1891, respectively .
3. As mentioned above, the finding that increasing compulsory schooling from age 15 to 16 decreases educational achievement raises serious questions about the validity of the whole approach. The author sort of brushes this aside by saying that the estimates generally go in the right direction. I don’t find this argument convincing, however. Hopefully the above suggestions (including province specific trends, etc.) will help straighten up the results.
4. There is a serious disconnect between the numbering of tables and figures in the text and in the tables and figures per se. In particular: A) “Figure 3” on the bottom of page 6 should be “Figure 1”. B) Footnote 8 says that results for men and women are not shown but are available on request while they are in fact shown in Table 4. C) All the tables beyond table 4 are mislabeled in the text (e.g. the text says “table 4” for what is really “table 5”, etc.)
5. The heading of Table 6 is incorrect. We are shown the effects for income only, not education and income.

Final comment:

This paper is currently written as a not so innovative contribution to the literature on the education earnings. The innovation is to use Canadian data, but otherwise the paper follows a very conventional approach.

The author should consider a completely different way of “pitching” his paper. Several papers (Barro-Lee, Krueger-Lindhal, etc.) have looked at the effect of education in cross-country growth. Substantive question here is whether cross-sectional estimates of the return to education can really be used to compute how education growth contributes to the economic growth of nations. This has long been an issue in economics. In fact, Griliches’ original work on ability biases in the return to education was motivated by this growth accounting literature.

In one sense, this paper looks at this very issue by analyzing whether aggregate cross-province growth in income and earnings can be explained by education. As I mentioned earlier, the OLS estimates based on cell means are not comparable to the usual cross-sectional estimates. How are these OLS estimates identified? Since the author includes a full set of province, cohort and year dummies, the identification comes from differential inter-cohort variation between education and earnings within provinces over time. A prime suspect here is that provinces with the strongest growth in education achievement (Quebec and the Atlantic provinces) have also experienced the strongest growth in earnings. If this conjecture is correct, it means that this paper presents compelling evidence that convergence in education across Canadian provinces has contributed to the convergence in earnings and income across provinces.

So in one sense the OLS estimates presented here are really “macro-estimates” that are relevant for looking at whether education is relevant for provincial growth. The fact that they appear to be similar (or larger perhaps) than standard cross-sectional OLS estimates (not reported in the paper) is an important finding. I am not an expert on education and growth but I think that this paper would have more impact by addressing these issues and clarifying how the different estimates should be interpreted.