Canadian Labour Market and Skills Researcher Network

Working Paper No. 20

How Would One Extra Year of High School Affect Wages? Evidence from a Unique Policy Change

Harry Krashinsky
University of Toronto

April 2009

CLSRN is supported by Human Resources and Social Development Canada (HRSDC) and the Social Sciences and Humanities Research Council of Canada (SSHRC).
All opinions are those of the authors and do not reflect the views of HRSDC or the SSHRC.

How Would One Extra Year of High School Affect Wages? Evidence from a Unique Policy Change

Harry Krashinsky*

University of Toronto

Abstract

This paper uses a unique policy change in Canada's most populous province, Ontario, to provide direct evidence on the effect of reducing the length of high school on labour market outcomes for high school graduates. In 1999, the Ontario government eliminated the fifth year of education from its high schools, and mandated a new four-year program. This policy change created two cohorts of students who graduated from high school together with different amounts of education, thus making it possible to identify the effect of one extra year of high school education on earnings. Using restricted survey data, the results demonstrate that students who receive one less year of high school education receive wages that are approximately ten percent lower than their counterparts one year after graduation, and these effects persist two years after graduation. Using birth year to instrument for educational attainment produces estimates that are even higher than the cross-sectional findings, but quite consistent with the existing literature on the return to education. These results are a significant contribution to the literature on the return to education because unlike prior changes to the educational system, this change in schooling laws results in two cohorts entering the labour market simultaneously. As such, business cycle effects do not confound the results.

JEL Code: I20, I28, C10

Keywords: Human Capital, Returns to Education, Years of Schooling, Ontario Double Cohort

*Corresponding author: Harry Krashinsky, 121 St. George Street, Centre for Industrial Relations, University of Toronto, Toronto, Ontario, Canada, M5S 2E8. Telephone: (416) 978-1744. Fax: (416) 978-5696. Email: harry.krashinsky@utoronto.ca I would like to thank a number of people who assisted me in this project. First, I am indebted to Joshua Angrist, Dan Lang, Robert McMillan, Phil Oreopoulos, Diane Whitmore Schanzenbach, Elizabeth Stuart, and the University of Toronto for very helpful comments on this project. This project was funded by the Canadian Labour Market and Skills Researcher Network (CLSRN).

Executive Summary

This paper uses a unique policy change in Canada's most populous province, Ontario, to provide direct evidence on the effect of reducing the length of high school on labour market outcomes for high school graduates. The key change to the Ontario high school system was enacted in 1999 by the provincial government. Prior to this time, students could graduate high school by completing 30 credits: taking eight credits per year for the first three years and then six Ontario Academic Credits (OACs) in grade twelve, a student could, if he or she chose, complete their high school degree in four years. However, the majority of high school students chose to complete their degrees in five years and take more than 30 credits, since a four-year program typically did not allow the student enough time to take all of the courses he or she would like.

But the change enacted by the government instituted a standard four-year, 30-credit program for all students, which essentially forced all students to graduate high school within this time. This policy change created two cohorts of students who graduated from high school together with different amounts of education, thus making it possible to identify the effect of one extra year of high school education on earnings. Using restricted survey data, the results demonstrate that students who receive one less year of high school education receive wages that are approximately ten percent lower than their counterparts one year after graduation, and these effects persist two years after graduation for less-able workers from these cohorts. The findings are generally consistent with the existing literature on the effect of education on earnings, and the results are also reasonably robust in their magnitude. In addition, the results from this particular policy change are important because this particular change in schooling laws results in two cohorts entering the labour market simultaneously. As such, they were competing for equivalent jobs in similar labour markets and under the same macroeconomic conditions. As such, business cycle effects do not confound the results.

1 Introduction

The causal return to additional year of education represents the extra earnings derived from receiving one more year of education. To properly estimate this effect, many recent papers have sought the use of instrumental variables, which alter an individual's educational attainment, but are otherwise unrelated to his or her earnings. Ideally, an experiment to gauge the causal impact of education on earnings would randomly assign additional education to a treatment group, while denying it to an otherwise identical control group. Some of the instrumental variable approaches that have been used have exploited existing administrative rules about educational attainment such as the minimum age of schooling and a student's quarter of birth (Angrist and Krueger (1991)), or policies which raise the minimum age of schooling (Harmon and Walker (1995), Oreopolous (2006)). More generally, instruments used to estimate the returns to education have considered individuals who have been induced to obtain more education for various reasons that are unrelated to their potential earnings, such as proximity to colleges or universities (Kane and Rouse (1993), Card (1995), Conneely and Uusitalo (1997)), or within-twin differences in educational attainment (Ashenfelter and Krueger (1991), Ashenfelter and Rouse (1995)). All of these approaches are predicated on the notion that they compare otherwise identical individuals who differ only through their educational attainment.

But a common criticism to the instrumental variable approach is whether or not the treatment and control groups are, in fact, identical. For instance, Bound, Jaeger and Baker (1995) have criticized Angrist and Krueger's work by arguing that individual ability and background may differ by quarter of birth, thus invalidating this instrument. Similar criticisms have been leveled at instruments which use college proximity, or within-twin variation in educational attainment, to generate randomized variation in education. In addition, instruments which rely upon changes in the minimum age of schooling are prone to potentially contaminating business cycle effects. Specifically, these studies consider wage differences between two cohorts of students who attempt to leave school as soon as possible before and after the minimum schooling laws take effect; the resulting two cohorts of students have different levels of education that are randomly assigned due to the alteration of the law. However, the cohorts in these types of experiments enter the labour market in different years: as an example, a one-year increase in the minimum schooling laws would result in two cohorts entering the labour market two years apart from one another. Any business cycle fluctuations over these two years which had a differential effect on the wages received by either cohort would bias the return to education estimated from these cohorts.

But a new educational policy for high schools in Canada's most populous province, Ontario, has created a new instrumental variable for educational attainment which may circumvent previous problems encountered in the returns to education literature. Unlike prior studies which have relied upon policy changes that have induced students to obtain more education, this work will focus on a policy which caused students to obtain less education after it was enacted (that is, the policy decreased the amount of education given to treatment group). The policy analyzed in this paper relies upon a restructuring of the Ontario high schools system, which graduated most of its students after five years of study. But in 1999, the government of Ontario changed the structure of the high school program to graduate students in four years. This caused two cohorts of students to graduate simultaneously from high school in June, 2003 – the last class of five-year graduates, and the first class of four-year graduates. When these students left high school and entered the labour market together in 2003, it was possible to compare the labour market outcomes of four- and five-year graduates to consider the impact of having an extra year of education on earnings. More importantly, since assignment to the four- or five-year cohort is determined by birth year, this policy change provides a natural experiment for considering the impact of variation in education on wages. Restricted survey data demonstrates that four-year graduates earn approximately 10% less than five-year graduates in the labour market one year after graduation, and for less-able workers, there is still a significant difference in earnings two years after graduation.

One complication for the analysis is that although most of the students in the five-year system graduated high school after five years, this system also provided the option to graduate after four years. To accommodate this issue of noncompliance, the elimination of a fifth year of high school can be viewed as a policy with a random encouragement design[1], and its causal impact can be analyzed with a two-stage least squares approach, which focuses on students who complied with the program. The evidence shows that most students did, in fact, comply with the policy change imposed on the high school system, and its impact on labour market performance is similar to this paper's initial findings: for compliers, five-year graduates earn wages that are approximately 10 to 15% higher than four-year graduates one year after graduation, and less-able compliers still exhibit a significant wage difference two years after graduation.

2 Literature Review

The impact of variations in educational attainment on earnings has been considered by many analysts, and all of these studies seek to overcome the underlying problem of endogeneity that plagues any analysis of the return to education. The standard wage equation is specified as:

$$\log(wage) = \beta(education) + \gamma X + \varepsilon$$

where X represents observable characteristics, and ε is an error term. The endogeneity of education within this equation relates to its correlation with the error term, since the unobservable portion of ability may be correlated with educational attainment. The instrumental variable approach has relied upon using a variable, Z, which is correlated with *education* but uncorrelated with the error term; intuitively, this implies that Z has an effect on education that is independent of the wage. In the context of the wage equation, the instrumental variable increases education but does not, by itself, affect the wage rate. Card's

[1] Please refer to Angrist (2004), Angrist, Imbens and Rubin (1996), Hirano (2000), and Imbens and Rubin (1997) for a discussion of this approach.

(1999) review of this literature discusses how the optimal choice of educational attainment is derived by equating the marginal cost of education to its marginal benefit, and as such, many of the quasi- and natural experiments used to generate instrumental variables for the returns to education use randomized changes in marginal cost to observe how the resulting change in the optimal choice of education affects earnings. There are three main classes of such studies: the first group (Kane and Rouse (1993), Card (1995), Conneely and Uusitalo (1997)) uses variation in the marginal cost of education, such as proximity to a college, to obtain a causal effect of education on earnings. Assuming that proximity to a college has no independent effect on wages aside from its incentive to obtain more education, this is a legitimate instrument. However, a critique of this approach is that although it does provide variation in the marginal cost of college, it does not do so in a random manner. Specifically, the types of individuals who live closer to a college may be more able than those who do not. Expanding the above equation, it could be written that:

$$\log(wage) = \beta(education) + \gamma X + A + \delta$$

where A represents the unobservable component of individual ability, so that $\varepsilon = A + \delta$, where δ is an error term. In this case, cov(college proximity, A) $\neq 0$, which would invalidate the instrument.

The second class of papers (Ashenfelter and Krueger (1991), Ashenfelter and Rouse (1995), Krashinsky (2004)) have used a within-family approach, the most recent of which uses twins, who are assumed to be equally able. In that case, within-twin differences in earnings and education yield unbiased estimates of the causal effect of education on earnings because it completely eliminates the unobservable component of ability from the model:

$$\Delta\log(wage) = \beta(\Delta education) + \gamma\Delta X + \Delta\delta$$

where $\Delta\log(wage)$, $\Delta education$, ΔX, and $\Delta\delta$ are the within-twin differences in wages, education observable characteristics and error terms. However, some have viewed this assumption as tenuous, and have questioned why equally-able twins would obtain different levels of education (since their marginal benefits and costs should also be the same). In particular, if the within-twin difference does not completely remove the unobservable component of ability from the model, then this approach does not yield an estimate of the causal return to education.

The third class of papers (Harmon and Walker (1995), Oreopolous (2006)) relies upon changes in minimum schooling laws. Although these do not alter an individual's marginal cost function, they can cause an individual to obtain more education than he otherwise would if he had unconstrained choice. Specifically, if one cohort is subject to a lower minimum schooling law than another, then birth year (which determines assignment to a particular cohort) will serve as a legitimate instrumental variable. However, one problem with this approach is that the treatment and control groups enter the labour market in different years, and if business cycle effects are present, then it can confound the instrumental variable strategy. In particular, suppose the wage equation re-written as:

$$\log(wage) = \beta(education) + \gamma X + BC + \delta$$

where BC represents the impact of the business cycle. In this case, since $cov(change\ in\ minimum\ schooling,\ BC) \neq 0$, then this instrument can be invalid, because it would improperly attribute business cycle effects to the change in minimum schooling laws.[2]

[2] Some other recent studies have considered the impact of taking specific fields or additional high school courses and their individual impact on wages. In this literature, the results vary quite widely. Altonji (1995) used the NLS and found that an additional high school course has a very small impact on wages – an additional year of math, science, English, social studies and a foreign language only have a 0.3 percent increase on wages. This effect rises to about three percent if the negative impacts of English and social studies are excluded, but he concludes that the impact of an individual course is much less than a full year of study on wages, which is evidence in favour of education being more consistent with signalling than effective training. Levine and Zimmerman (1995) used the High School and Beyond data, and like Altonji, found that an additional semester of math courses has a very small effect on wages.

The elimination of the fifth year of high school in Ontario can be an important contribution to this literature because it does not fall prey to the problems outlined for the other instrumental variable strategies in this literature. First, the assignment to a five- or four-year high school cohort is random and not correlated with wages; it is based upon birth year. Second, unlike other studies which use changes in schooling laws, the treatment and control groups can not be differentially affected by business cycle effects, since they enter the labour market simultaneously. The specifics of the changes made to the Ontario high school system will be detailed in the following section.

3 Program

In 1995, Ontario's provincial government announced that it was making a fundamental change to its high school system for students entering high school in 1999. Prior to this, students could graduate high school by completing 30 credits: taking eight credits per year for the first three years and then six Ontario Academic Credits (OACs) in grade twelve, a student could, if he or she chose, complete their high school degree in four years. However, the majority of high school students chose to complete their degrees in five years and take more than 30 credits, since a four-year program typically did not allow the student enough time to take all of the courses he or she would like.[3] Students who graduated in five years typically took eight courses per year for the first three years, and then at least six OACs in each of the final two years of high school. But the change enacted by the government instituted a standard four-year, 30 credit program for all students, which essentially forced all students to graduate high school within this time. Most importantly, the new system

[3] In most cases, a four-year graduate could not take very many courses in "elective" subjects, such as drama, music or art. And, by necessity, a four-year program was more limiting in the number of upper-year classes the student could take in other fields outside of math, science and english. Five-year graduates, on the other hand, had the flexibility to take such courses, and most completed high school with 36 to 38 credits.

basically left the educational curriculum unchanged for the first three years of high school for both the four- and five-year groups; the main effect of the program was to decrease the number of OAC courses available to the four-year group. As such, the elimination of the fifth year of high school significantly decreased the human capital of students who intended to cease their education after high school.

The overall effect of this policy change was to create two cohorts of students who would graduate from high school in June of 2003: the last of the five-year high school graduates, and the first group that was mandated to graduate from high school in four years. Those who did not obtain any further education entered the labour market immediately, and as such, the two cohorts entered the labour market together, creating a unique opportunity to assess the impact of variation in educational attainment on labour market performance: two groups were competing for jobs in the same market. The only difference between these two groups was that one was assigned to a four-year high school curriculum, while the other graduated from a system which overwhelmingly produced five-year graduates. This difference in educational attainment, which was randomly assigned by the year of birth, provides a natural experiment to estimate the causal return to a year of education on earnings.

4 Data and Results

The data set used in this study is the Survey of Labour and Income Dynamics (SLID), and is a restricted-use data set that is administered by Statistics Canada. The survey is designed to be a representative sample of Canadians, and is conducted longitudinally for a subset of the respondents. To maximize the sample size for this study, the cross-sectional group of the SLID was used in this study, and two subsamples were considered for the paper: the five-year graduates and four-year graduates from the double cohort who did not obtain any further education after leaving high school. Since the SLID interviews respondents throughout the year, the 2004 wave of the survey was used in order to ensure that neither cohort of students was being interviewed while they were in school. Sample means of some key variables for the two groups of high school graduates surveyed in 2004 are displayed in Table 1. The first row displays average age as of December 31, 2004 for both cohorts, and the results show that the five-year cohort is approximately one year older than the four-year cohort. This is due to the fact that assignment to the two cohorts was based upon birth-year: students in Ontario must be enrolled in first grade as of the September of the year in which they turn six years old, so the last class of five-year high school graduates began first grade in the fall of 1990, and the first four-year class began first grade in the fall of 1991. Thus, at the end of 2004, graduates of the five-year program would be 20 years old, while the four-year cohort would be 19 years old.

The second row of Table 1 displays the difference in hourly wages earned by respondents from the two cohorts, and the results demonstrate that the five-year cohort earn wages that are approximately 15% higher than their four-year counterparts. This is basic evidence to suggest that receiving an additional year of education in high school causes the five-year cohort to earn higher wages. More importantly, this is roughly consistent with many existing studies which consider the return to education: a ten to fifteen percent return to education is the most common cross-sectional estimate within this literature. This consistency with the existing literature is important because it suggests that the analysis of the double cohort is not an unrepresentative sample within the labour market. The remaining rows in the table describe other observable characteristics of the two cohorts, which are remarkably similar. Both groups have similar: unionization rates, percentage of female workers, percentage of individuals in urban centres, and generally similar industry concentrations. Aside from wages, there appear to be little observable differences between the groups, which is also important because it validates the empirical exercise. If the two groups were different in some observable way, it might raise the possibility that wages differed for reasons other than the treatment of an additional year of education.

These wage differences are considered in a more rigorous way in Table 2, which displays wage regression results for the two samples. The regression for this table takes the following specification:

$$\log(wage) = \beta(five\ year) + \gamma X + \varepsilon$$

where $five\ year$ is an indicator variable equal to one if the respondent is a member of the five-year cohort, and zero if they are a member of the four-year cohort, X represents observable characteristics and ε is an error term. The first row of the table confirms the general findings of table one, which showed a significant wage difference between the two cohorts. In the first column of table two, a simple bivariate regression is conducted, in which $five\ year$ is the only independent variable in the model, and the regression confirms that there is a significant difference in earnings between the two cohorts; the five-year cohort earns approximately ten percent more than the four-year cohort, and this difference is statistically significant. As mentioned earlier, this simple finding is important because it affirms the correspondence between the results from this policy change with the rest of the literature on the returns to education. The next two columns display regressions which include other typical covariates included in wage regressions, and two facts are evident from columns two and three: first, the return to a fifth year of high school education has not been significantly altered by the inclusion of other covariates, suggesting that the significant wage differential between the cohorts is not caused by differences in observable characteristics. The second point that can be observed from these columns is that the signs and magnitudes on the other covariates included in the regressions appear to be consistent with estimates from standard wage regressions. In particular, the regressions report that women earn significantly less than men; there is a positive coefficient on the indicator variable which captures union membership, married men earn more than unmarried men, but married women do not exhibit a significant difference in earnings compared to single women.

An additional concern with the analysis is that the sample is extremely limited in definition, which limits the analysis to 19- and 20-year-old individuals who were members of the double cohort. To determine if some of the individuals sampled by the SLID are representative of the double cohort, the results from the first three columns of the table are replicated with regressions which incorporate sampling weights. As is clear from the results in columns four through six, the sample weights basically leave the results unchanged, suggesting that the significant wage difference between the four- and five-year cohorts is not due to the sampling of the SLID.

5 Noncompliance

The difficulty in assessing the impact of the one-year reduction in the length of high school on labour market performance is that compliance with this program was not perfect. As was previously discussed, the former five-year high school system actually allowed students to choose whether they completed their high school studies in four or five years. Although an extremely large percentage of these students opted to graduate in five years, others chose to do so in four years. This lack of perfect compliance introduces a complication into the analysis, but it can be dealt with using established statistical procedures. In particular, if all students who entered high school before 1999 were regarded as though they were assigned to a five-year program (as a control group), while all students who entered high school in 1999 and after are regarded as being assigned to a four-year program (a treatment group), then a standard noncompliance framework can be applied, where the policy change is viewed as a random encouragement design. That is, students in the treatment and control groups were encouraged to complete their studies in four and five years, respectively, but any individual student could opt for a four- or five-year program.

To formalize this notion, suppose that Z represents a dummy variable equal to one if the student has been assigned to the four-year program, and zero if he or she has been assigned to the five-year program. Also, suppose that the dummy variable $D(Z)$ is a function of the treatment that the student was assigned, Z, and is equal to one if the student actually chose to complete a four-year program, and zero otherwise. For instance, if $D(Z=1)=0$, then this represents a student who was assigned to the four-year program, but actually chose to take five years of high school education. This notation is useful, because it allows for a classification of student i, C_i, to one of four possible types of students – a complier (c), a "never-taker" (n), an "always-taker" (a), and a defier (d). Specifically, the student is a complier if he or she actually takes their assigned four- or five-year program. Always-takers are classified as students who will always opt for treatment (the four-year program), regardless of whether or not they were officially assigned to it, and never-takers will never take treatment and opt for a five-year program regardless of their assignment. Defiers are students who simply do the opposite of what they are assigned. In terms of the variables Z and $D(Z)$, these four types can be classified as follows:

$$
\begin{aligned}
C_i &= c \text{ if } D(Z) = Z \text{ for } Z = 0, 1 \\
C_i &= a \text{ if } D(Z) = 1 \text{ for } Z = 0, 1 \\
C_i &= n \text{ if } D(Z) = 0 \text{ for } Z = 0, 1 \\
C_i &= d \text{ if } D(Z) = 1 - Z \text{ for } Z = 0, 1.
\end{aligned} \tag{1}
$$

The result of interest in this case is the impact of the removal of the fifth year of high school education on labour market wages. In randomized trials which ignore compliance information (or assume perfect compliance), the researcher seeks to estimate the Intention-to-Treat (ITT) effect, which is $Y(Z=1) - Y(Z=0)$, where Y is some outcome measure. In the case of the removal of a fifth year of high school education, the difference in grades attained by the students in both groups has been the effect of interest – if there were perfect compliance in this case, the average difference in grades is known as the Average Causal

Effect (ACE), which is equal to:[4]

$$ACE = \frac{1}{N_{5yr}} \sum Y_i(Z_i = 1) - \frac{1}{N_{4yr}} \sum Y_i(Z_i = 0) \tag{2}$$

where N_{5yr} and N_{4yr} represent the sample sizes of the five-year and four-year samples, respectively. The results from the previous section calculated this estimand for the sample, which contains a mixture of compliers and always-takers.[5] Although this is an important effect to estimate, since it captures all types of students affected in 2004 by the policy change, a slightly different effect must also be estimated. Because the system allowed for some students to opt out of their encouraged programs, it is important to determine the relative performance of those who chose to graduate high school in the program they were encouraged to use. In particular, denoting the outcome measure for student i as $Y_i(Z_i, D_i(Z_i))$, the effect of interest is the Complier Average Causal Effect (CACE) – the ITT for compliers:

$$CACE = \frac{1}{N^C_{5yr}} \sum Y_i(Z_i = 1, D_i(Z_i = 1) = 1) - \frac{1}{N^C_{4yr}} \sum Y_i(Z_i = 0, D_i(Z_i = 0) = 0) \tag{3}$$

where N^C_{5yr} and N^C_{4yr} represent the sample sizes for compliers in the five-year and four-year samples, respectively. This effect represents the impact of the policy change for all four- and five-year graduates who complied with the policy change, and is relevant because it restricts the sample to a group whose behavior can actually be altered in expected ways by the policy change.

To compute the CACE, Angrist (2004) demonstrates that a two-stage least squares approach is sufficient. In this case, the variable that can be used to instrument for whether

[4] The indicator variable Z, which classifies assignment, has been subscripted with an i for each student. This is due to the fact that it is being assumed that an individual's assignment status is independent of any other student's assignment status. This is a relatively innocuous assumption if there is random assignment – in this case, since assignment is based upon birth year, this is a fairly safe assumption. Formally, this is known as the Stable Unit Treatment Value Assumption (SUTVA), and is discussed in more detail in Rubin (1978, 1980), Imbens and Rubin (1997) and Angrist, Imbens and Rubin (1996).

[5] For the purpose of this study, it will be assumed that defiers do not exist in the sample. This is a commonly invoked assumption in noncompliance studies and aids in the identification strategy of the model.

or not a student completed high school in four or five years is the student's birth year. As previously discussed, students born in 1985 were assigned to the four-year cohort, and those born in 1984 were assigned to the five-year cohort. The first stage of the instrumental variable procedure involves the following regression:

$$Four\ Year\ Graduate_i = \alpha(Born\ in\ 1985)_i + \beta X_i + \varepsilon_i$$

where $Four\ Year\ Graduate_i$ is an indicator variable equal to one if student i graduated high school in four years and zero otherwise, $(Born\ in\ 1985)_i$ is another indicator equal to one if student i is born in 1985, X_i captures other observable characteristics of the student, and ε_i is an error term. The results from the first stage are presented in Table 3 for the unweighted and weighted samples considered in the analysis. In both samples, regardless of whether or not other covariates are included in the first stage, the indicator for being born in 1985 is highly significant, and in all cases exhibits an F-statistic which implies that this variable is not a weak instrument.[6]

The second stage of the instrumental variable process is reported in Table 4. The results in this Table display the instrumented effect of obtaining a fifth year of high school education on earnings relative to the four-year class, and as previously discussed, this can be interpreted as the CACE. The first three columns of the Table demonstrate that the instrumented relative wage for five-year high school graduates in 2004 is significantly higher than the four-year graduates; in fact, the results suggest that compliers from the five-year class exhibit earnings that are between 18 and 19 percent higher than their four-year counterparts. To account for potential impacts on these results due to sampling strategies, the last three columns of the table display the second-stage results from the instrumental variable process which use the same specifications as in the first three columns. The findings are similar to results in the first three columns, which suggest that sampling strategies for the SLID are not responsible for the significant difference in earnings estimated by the model. Overall, the findings in the two-stage procedure are not significantly different than findings

[6] See Staiger and Stock (1997) and Bound, Jaeger and Baker (1995) for a discussion of weak instruments.

in other studies which use an instrumental variable procedure to calculate the causal return to a year of education. The estimates of the causal return to education reported by Harmon and Walker (1995) and Oreopoulos (2006), which use changes in minimum schooling laws as an instrument for educational attainment, are as high as 14 to 15 percent, which is well within the confidence intervals of any of the two-stage estimates reported in Table 4. Similarly, many other studies report a two-stage estimate of the return to education in excess of ten percent, including Ashenfelter and Krueger (1994), Ashenfelter and Rouse (1998), Card (1995), Maluccio (1997), amongst others. The consistency of the findings presented here with the existing literature implies that the findings from the double cohort are not uncommon results that are caused by the evaluation of an unusual sample. Also, like other instrumental variable studies of the causal return to education, this paper has found that education has a significant and large effect on wages.

6 Longer-run results

One concern for the interpretation of the results in Tables 4 and 5 is whether or not they are due to the five-year cohort being more productive than the four-year cohort, or if other factors are contributing to the wage differential between the cohorts. For instance, if labour market demand does not adjust sufficiently quickly to accommodate the additional supply of workers from the double cohort, then it is possible that the wage differential between the cohorts is due to differences in productivity as well as short-run adjustments by firms. To address this possibility, the analysis was replicated using data from 2005 wave of the SLID, and an interesting trend is revealed. The results in column one demonstrate that there is no longer a significant difference in earnings between the two cohorts; the first column of Table 6 shows that the average wage difference between the cohorts is a statistically insignificant 3 percent. The source of this difference, however, is revealed in a nonparametric analysis of the two cohorts in Figure 2. Unlike Figure 1, the kernel densities in Figure 2 from the four-and five-year cohorts hourly wages are virtually the same in the upper halves, suggesting that more able workers from the four-year cohort have caught up to more able workers from the five-year cohort. However, a gap in the densities remains in their lower halves. A regression analysis confirms these observations; the second and third columns of Table 6 display results from wage regressions for workers with wages above and below the median of their distributions, and although the wage difference is not significant in column two for the workers with wages above the median, the wage difference of 5 percent is significant for workers with wages below the median. The same wage regressions in the first three columns of the table are estimated using a two-stage least squares approach in the final three columns of Table 6 in order to estimate the CACE. The results in columns 4 through 6 confirm that although there is not a significant instrumented wage difference between the cohorts in the pooled sample and the subsample of workers with relatively high wages, there is a significant 5 percent difference in earnings between the lower-paid workers in both cohorts.

Overall, the results suggest that there has been a convergence of wages for more-able workers, but a difference in earnings remains for less-able workers. This is consistent with the hypothesis that more able workers from the four-year cohort were negatively affected by a temporary lack of demand in the labour market, but this negative effect dissipated after two years. However, less-skilled workers from the five-year cohort still exhibit higher wages than less-skilled workers from the four-year cohort even two years after entering the labour market. This result is consistent with heterogeneous returns to education by skill – the effect of education is initially significant for all workers, but more-able workers are able to overcome this effect after two years. More, for this sub-group, it appears that the initial wage difference between the two cohorts is mostly due to labour market demand adjustments, not inherent productivity differences. Were there significant differences in productivity conveyed by the additional human capital acquired by the more-able members of the five-year cohort, it would be expected that two years after graduating from high school, they would still exhibit significantly higher wages than more able members of the four-year cohort. However, there do appear to be significant differences in earnings that persist between the cohorts for less-able workers. Less-able workers from the four-year cohort still exhibit significantly lower wages than less-able workers from the five-year cohort, even two years after graduation, which is consistent with a productivity difference between the two groups that is due to extra education received by the five-year cohort.

7 Conclusion

Many studies have considered the impact of an additional year of education on earnings, and to circumvent the inherent endogeneity problem of this analysis, most papers have relied upon an instrumental variable approach to obtain an estimate of the causal effect of education on earnings. But many of these instrumental variables have been criticized for a variety of reasons, which focus on how the instrument may be endogenous itself. For instance, changes in mandatory schooling laws have been criticized because they analyze cohorts that enter the labour markets at different times (or that the cohorts themselves are not fundamentally similar in dimensions other than educational attainment). However, the elimination of a fifth grade of high school in Ontario presented a unique opportunity to use a random encouragement policy experiment to consider how variation in the years of high school education affects labour market performance. A simple analysis of four- and five-year high school graduates who entered the labour market simultaneously after graduation revealed that four-year graduates performed significantly worse than five-year graduates. Four-year graduates earned wages that were approximately ten percent lower than wages earned by graduates from the five-year high school cohort one year after graduating from high school, and less-able members of the four-year cohort still exhibited lower wages than less-able members of the five-year cohort two years after graduation. This differential was robust to the inclusion of other covariates and sample weights, and represented important new evidence on the return to education. Unlike prior studies which used changes in mandatory schooling laws to instrument for educational attainment, the two cohorts in the "double cohort" entered the labour market simultaneously and differed only by their educational attainment, which was randomly encouraged on the basis of birth year.

But one complication for the analysis is that high school students in the five-year program could graduate in four years, if they so desired. However, this issue can be addressed with a two-stage least squares estimation strategy. Classifying the five-year group as the control group, and the four-year group as the treatment group, then a two-stage least squares approach can determine the effect of eliminating a fifth year of high school on wages for those who "comply" with the change in high school graduation times. The results demonstrated that the complier average causal effect of removing one year of high school education caused significantly lower wages for four-year graduates, compared to the five-year class, and this effect persisted for less-able graduates two years after graduation. This significant effect on wages is consistent with other papers which use an instrumental variable approach to estimate the return education, suggesting that the findings in this paper are not the result of analyzing an uncommon sample that is unrepresentative of the overall population.

References

[1] Abadie, A. and G. Imbens, "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects." NBER technical working paper #283, 2002.

[2] Altonji, Joseph G. "The Effects of High School Curriculum on Education and Labor Market Outcomes." *Journal of Human Resources*, 30(3), Summer 1995, pp. 409-438.

[3] Angrist, Joshua D. "Treatment Effect Heterogeneity in Theory and Practice." *Economic Journal*, 114(494), March, 2004, pp. C52 - C83.

[4] Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434), June 1996, pp. 444-455.

[5] Angrist, J. and A. Krueger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings," Quarterly Journal of Economics, 106, 979-1014.

[6] Ashenfelter, O. and A. Krueger (1994), "Estimates of the Return to Schooling from a New Sample of Twins," American Economic Review 84, 1157-1173.

[7] Bound, John and David A. Jaeger. "Do Compulsory School Attendance Laws Alone Explain the Association between Quarter of Birth and Earnings?" Research in Labor Economics,

[8] Bound, John; David A. Jaeger and Regina M. Baker. "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Variable is Weak." Journal of the American Statistical Association, June 1995, 90(430), pp. 443-450.

[9] Bound, John and Gary Solon. "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling." *Economics of Education Review,* April 1999, 18(2), pp. 169-182.

[10] Card, David. "The Causal Effect of Education on Earnings." Handbook of Labor Economics. Volume 3A, 1999, pp. 1801-1863.

[11] Card, D. (1995), "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp, ed. by L. N. Christofides, E. K. Grant, and R. Swidinsky. Toronto: University of Toronto Press, 201-222.

[12] Card, David and Thomas Lemieux. (2001) "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis" Quarterly Journal of Economics, 116(2), pp. 705-46

[13] Conneely, Karen and Roope Uusitalo. "Estimating Heterogeneous Treatment Effects in the Becker Schooling Model." Princeton University Industrial Relations Section, Working Paper, 1997.

[14] Dearden, Lorraine, Javier Ferri and Costas Meghir. "The Effect of School Quality on Educational Attainment and Wages." *Review of Economics and Statistics*, 84(1), February 2002, pp. 1-20.

[15] Grogger, Jeff and Eric Eide. "Changes in College Skills and the Rise in the College Wage Premium." *Journal of Human Resources*, 30(2), Spring 1995, pp. 280-310.

[16] Harmon, Colm and Ian Walker. "Estimates of the Economic Return to Schooling for the United Kingdom." American Economic Review, 85(5), pp. 1278-1286.

[17] Hirano, Keisuke et. al. "Assessing the Effect of an Influenza Vaccine in an Encouragement Design." *Biostatistics*, January 2000, 1(1), pp. 69-88.

[18] Hsieh, C and D. E. Clark (2000), "Schooling and Labor Market Impact of Taiwan's 1968 Nine-Year Education Program," Unpublished manuscript.

[19] Imbens, Guido W. and Donald B. Rubin. "Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance." *The Annals of Statistics*, 25(1), January 1997, pp. 305-327.

[20] Jones, Ethel B. and John D. Jackson. "College Grades and Labor Market Rewards." *Journal of Human Resources*, 25(2), Spring 1990, pp. 253-266.

[21] Kane, T. J. and C. Rouse (1993), "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?" NBER Working Paper #4268.

[22] Katz, Lawrence F and Kevin M. Murphy (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors" Quarterly Journal of Economics, 107(1), pp. 35-78

[23] Krashinsky, Harry A. (2004), "Do Marital Status and Computer Usage Really Change the Wage Structure? Evidence from a Sample of Twins," Journal of Human Resources, 39(3), pp. 774-791.

[24] Krashinsky, Harry A. (2005), "How Would One Extra Year of High School Affect Academic Performance in University? Evidence from a Unique Policy Change", Mimeo, University of Toronto.

[25] Levine, Phillip B. and David J. Zimmerman. "The Benefit of Additional High-School Math and Science Classes for Young Men and Women." *Journal of Business and Economic Statistics*, 13(2), April 1995, pp. 137-149.

[26] Orepolous, Phil (2006). "Average Treatment Effects of Education when Compulsory School Laws Really Matter", American Economic Review, 96 (1), pp.152–175.

[27] Rubin, Donald B. "Discussion of 'Randomization Analysis of Experimental Data in the Fisher Randomization Test'." *Journal of the American Statistical Association*, 75, pp. 591-593.

[28] Sweetman, Arthur. "What if High School were a Year Longer? Evidence from Newfoundland." Mimeo, 1999, 31 pages.

[29] Staiger, Douglas and Stock, James H. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65(3), May 1997, pp. 557-586.

[30] Welch, Finis (1979) "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust", Journal of Political Economy, 87(5), Part 2, October 1979, pp. S65-97

Table 1: Sample Means for Five- and Four-year High School Graduates from the Double Cohort

	Five-Year Graduates	Four-Year Graduates
Age	20.52 (0.28)	19.50 (0.29)
Hourly Wage	10.81 (7.58)	9.36 (5.13)
Female	0.492 (0.502)	0.482 (0.501)
Union Members	0.131 (0.338)	0.108 (0.312)
Urban	0.454 (0.500)	0.524 (0.501)
Construction Industry	0.300 (0.460)	0.271 (0.445)
Manufacturing Industry	0.092 (0.291)	0.084 (0.279)
Service Industry	0.346 (0.478)	0.361 (0.482)
N	130	166

Standard deviations are listed in brackets beneath the sample means. The samples are composed of high school graduates from Ontario's double cohort who earn at least $5/hour.

Table 2: Wage Regressions for Five- and Four-year High School Graduates from the Double Cohort

	Unweighted			Weighted		
Member of the Five-Year Cohort	0.101 (0.040)	0.102 (0.036)	0.107 (0.035)	0.103 (0.045)	0.105 (0.042)	0.099 (0.041)
Female		-0.168 (0.034)	-0.106 (0.038)		-0.184 (0.042)	-0.140 (0.043)
Union		0.082 (0.062)	0.101 (0.057)		0.025 (0.063)	0.045 (0.063)
Married		0.312 (0.032)	0.415 (0.036)		0.269 (0.623)	0.377 (0.607)
Married*Female		-0.302 (0.115)	-0.409 (0.120)		-0.166 (0.684)	-0.307 (0.667)
Other Observable Characteristics	No	No	Yes	No	No	Yes

Standard errors are listed in parentheses beneath the coefficient estimates, and all regressions use a dependent variable equal to the log of the respondent's hourly wage. The estimates from the first three columns are derived from unweighted regressions and the last three columns display results from regressions which use sampling weights; all regressions include the covariates listed in the first column of the table. "Other Observable Characeristics" include six indicator variables to capture wage effects at the one-digit industry level, as well as an indicator variable equal to one if the respondent resides in an urban centre. The samples from both data sets are composed of individuals from Ontario's double cohort who earn at least $5/hour.

Table 3: First-Stage Regressions for I.V. Model

	Unweighted			Weighted		
Born in 1984	0.545 (0.049)	0.553 (0.049)	0.559 (0.050)	0.545 (0.050)	0.547 (0.050)	0.559 (0.050)
Female		0.009 (0.049)	-0.030 (0.053)		0.038 (0.050)	0.004 (0.052)
Union		-0.037 (0.076)	-0.042 (0.077)		0.022 (0.075)	0.001 (0.077)
Married		-0.803 (0.423)	-0.861 (0.426)		-0.814 (0.741)	-0.888 (0.738)
Married*Female		0.518 (0.473)	0.590 (0.479)		0.660 (0.813)	0.738 (0.811)
Other Observable Characteristics	No	No	Yes	No	No	Yes
F-Statistic	121.9	25.3	11.22	119.5	24.2	10.99

Standard errors are listed in parentheses beneath the coefficient estimates, and all regressions use a dependent variable equal to one if the respondent is a member of the five-year graduating class from the double cohort, and zero otherwise. The estimates from the first three columns are derived from unweighted regressions and the last three columns display results from regressions which use sampling weights; all regressions include the covariates listed in the first column of the table. “Other Observable Characeristics” include six indicator variables to capture wage effects at the one-digit industry level, as well as an indicator variable equal to one if the respondent resides in an urban centre. The samples from both data sets are composed of individuals from Ontario’s double cohort who earn at least $5/hour.

Table 4: Second-Stage Regressions for I.V. Model

	Unweighted			Weighted		
Member of the Five-Year Cohort	0.192 (0.071)	0.185 (0.067)	0.192 (0.064)	0.196 (0.080)	0.177 (0.074)	0.177 (0.074)
Female		-0.170 (0.036)	-0.100 (0.040)		-0.145 (0.043)	-0.141 (0.043)
Union		0.089 (0.067)	0.109 (0.063)		0.053 (0.063)	0.045 (0.063)
Married		0.460 (0.042)	0.581 (0.054)		0.559 (0.609)	0.534 (0.606)
Married*Female		-0.397 (0.110)	-0.522 (0.121)		-0.458 (0.669)	-0.438 (0.666)
Other Observable Characteristics	No	No	Yes	No	No	Yes

Standard errors are listed in parentheses beneath the coefficient estimates, and all regressions use a dependent variable equal to the log of a respondent's hourly wage. The estimates from the first three columns are derived from unweighted two-stage regressions and the last three columns display results from two-stage regressions which use sampling weights; all regressions include the covariates listed in the first column of the table. "Other Observable Characeristics" include six indicator variables to capture wage effects at the one-digit industry level, as well as an indicator variable equal to one if the respondent resides in an urban centre. The samples from both data sets are composed of individuals from Ontario's double cohort who earn at least $5/hour.

Table 5: CACE Regressions for I.V. Model

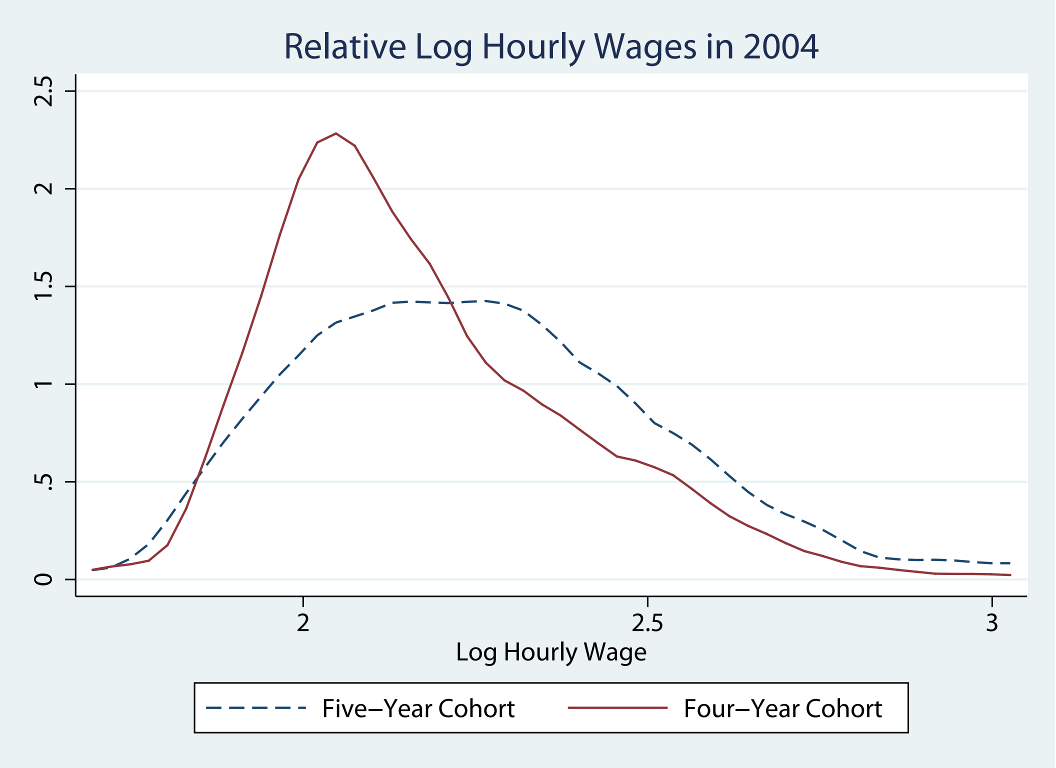
	First Stage			Second Stage		
Born in 1984	0.492 (0.044)	0.492 (0.040)	0.516 (0.044)			
Five-Year Cohort				0.123 (0.065)	0.123 (0.063)	0.135 (0.058)
Female		0.017 (0.043)	0.046 (0.044)		-0.156 (0.030)	-0.119 (0.031)
Union		0.087 (0.069)	0.066 (0.068)		0.103 (0.048)	0.104 (0.047)
Married		-0.475 (0.488)	-0.478 (0.474)		0.246 (0.340)	0.326 (0.327)
Married*Female		0.557 (0.473)	0.618 (0.549)		-0.137 (0.394)	-0.219 (0.379)
Other Observable Characteristics	No	No	Yes	No	No	Yes
F-Statistic	122.9	25.0	13.34			

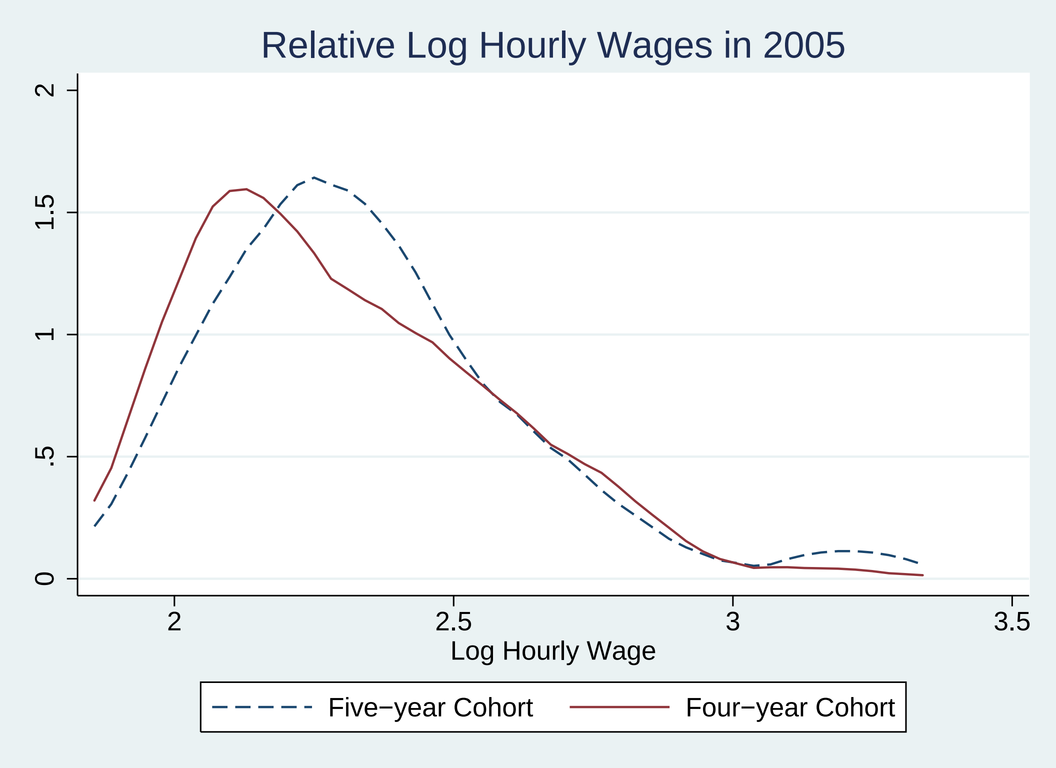
Standard errors are listed in parentheses beneath the coefficient estimates, and all regressions use a dependent variable equal to one if the respondent is a member of the five-year graduating class from the double cohort, and zero otherwise. The estimates from the first three columns are derived from unweighted regressions and the last three columns display results from regressions which use sampling weights; all regressions include the covariates listed in the first column of the table. "Other Observable Characeristics" include six indicator variables to capture wage effects at the one-digit industry level, as well as an indicator variable equal to one if the respondent resides in an urban centre. The samples from both data sets are composed of individuals from Ontario's double cohort who earn at least $5/hour.

Table 6: Wage Regressions from 2005 Data

	OLS Wage Regressions			Two-Stage Least Squares		
	Pooled Sample	Log Wage > Median	Log Wage < Median	Pooled Sample	Log Wage > Median	Log Wage < Median
Five-Year Cohort	0.026	0.029	0.050**	0.022	0.043	0.048**
	(0.031)	(0.051)	(0.021)	(0.038)	(0.065)	(0.023)
Female	-0.067**	0.104	-0.036	-0.096***	0.052	-0.038*
	(0.032)	(0.064)	(0.022)	(0.031)	(0.074)	(0.020)
Union	0.162**	0.120	0.018	0.187***	0.159	0.031
	(0.064)	(0.082)	(0.031)	(0.065)	(0.101)	(0.025)
Married	0.145	0.077	0.126***	0.235**	0.126*	0.139***
	(0.125)	(0.111)	(0.027)	(0.101)	(0.075)	(0.023)
Married*Female	-0.243	-0.013	-0.143***	-0.305***	-0.044	-0.135***
	(0.137)	(0.068)	(0.047)	(0.109)	(0.078)	(0.040)
F-Statistic				50.20	21.51	37.34

Standard errors are listed in parentheses beneath the coefficient estimates, and all regressions use a dependent variable equal to one if the respondent is a member of the five-year graduating class from the double cohort, and zero otherwise. The estimates from the first three columns are derived from unweighted regressions and the last three columns display results from regressions which use sampling weights; all regressions include the covariates listed in the first column of the table. "Other Observable Characeristics" include six indicator variables to capture wage effects at the one-digit industry level, as well as an indicator variable equal to one if the respondent resides in an urban centre. The samples from both data sets are composed of individuals from Ontario's double cohort who earn at least $5/hour.


Relative Log Hourly Wages in 2004
0
.5
1
1.5
2
2.5
2
2.5
3
Log Hourly Wage
Five−Year Cohort
Four−Year Cohort



Relative Log Hourly Wages in 2005
0
.5
1
1.5
2
2
2.5
3
3.5
Log Hourly Wage
Five−year Cohort
Four−year Cohort