The Effect of High School Duration on Educational Attainment and Labor Market Outcomes: Evidence from Ontario

Xian Zhang†
Andrei Munteanu§

October 30, 2020

Abstract

This paper studies a policy change in Ontario, Canada, which reduced the high-school duration from five to four years. By comparing the difference between pre- and post-reform cohorts in Ontario with those in the rest of Canada, we show that the policy led to a slight reduction in the high school graduation rate. Conditional on graduating from high school, shortened school length increased college and university enrollment. However, the increased enrollment in postsecondary education is offset by increases in dropout rates, leaving the educational attainment distribution unchanged. At the same time, our results suggest that the policy caused a significant wage penalty, approximately 5 to 10 p.p., conditional on educational level, for those who have graduated from high school under the new system. We propose two possible explanations for these findings: a composition effect due to the misallocation of individuals across educational levels, resulting from students being less informed about their ability ("orientation effect") and a decrease in human capital accumulation ("performance effect"). We construct a dynamic discrete choice model to assess these two channels' relative importance, finding that the orientation effect accounts for $11 - 20\%$ of the decrease in the wage premia across educational categories.

Keywords: Education, Labour, Policy Analysis
JEL Classification Codes: I21, I28, J24

†Department of Economics, McGill University.
§Department of Economics, McGill University.

1 Introduction

The choice of optimal duration of schooling is a fundamental issue for education policy (Meyer et al., 2015). On the one hand, as one type of input for academic outcomes, time spent in school contributes to human capital accumulation (Krashinsky, 2014) and ultimately, higher productivity and wages. On the other hand, staying in school increases costs along multiple dimensions. For individuals, longer program duration means delaying labor market entrance and thus entails opportunity costs. Moreover, from the public perspective, longer school curricula are associated with higher direct government expenditure, higher indirect cost from straining tax and social security systems, increased shortages of skilled workers and delayed intergenerational transfers of know-how (Meyer and Thomsen, 2018).

In this paper, we first quantify the effects of schooling duration on educational attainment, future educational decisions and labor market outcomes. Starting in 1999, the Ontario government implemented a reform that abolished the final year of high school while leaving the graduation requirement unchanged. This policy provides us exogenous variation in school duration, enabling us to analyze the causal impact of school length on students' educational choices and their labor market outcomes.

We then study two channels through which schooling duration impacts these outcomes. The first channel is "performance effect". If we believe that years of schooling affect human capital accumulation, then reducing high school length would adversely affect post-secondary-bound students' overall human capital level. This policy would lead to a reduction in productivity and, thus, wages. It may also explain the increase in the post-secondary dropout rate.

The second channel is the "orientation effect". First proposed by Meyer and Thomsen (2016), the orientation effect refers to the fact that time spent in school helps students better understand their abilities and preferences. A reduction in high school length means that students, on average, make worse decisions as to the optimal education level they should pursue. This can lead to sub-optimal high school dropout decisions, higher dropout rates in university and generally, a misallocation of students across different education levels. This misallocation could also explain lower wage premia for higher education.

Here is a preview of our results. We first examine how policy reform affected high school graduation rates, post-secondary educational decisions and, ultimately, wages. By employing data from various sources, we exploit the variation in high school length over time and across provinces. By comparing the difference between pre- and post-reform cohorts in Ontario with those in the rest of Canada, we show that the policy led to a 1.4% reduction in the high school graduation rate. Conditional on graduating from high school, the fraction of individuals enrolled in post-secondary education increased. It promotes immediate enrollment into university by 3% and college enrollment with a one year delay by 5%. The increased enrollment in post-secondary education is offset by increases in dropout rates, leaving the educational attainment distribution unchanged when measured at age $28-36$. There is 5% to 10% wage penalty conditional on education for those who graduated from high school under the new system.

Determining which of these two explanations is valid or, if they are both valid, the relative importance of either channel, is not possible within a reduced-form framework. Indeed, both changes in educational attainment decisions and lower human capital accumulation for a given education level can explain decreases in the average wages. Likewise, it is not clear if the change in educational distributions is due to wage rate changes or to students being less informed upon graduating from high school. Therefore, we construct a dynamic discrete choice model in which students learn their ability by going to school and receiving grades. We find that the relative importance of the two effects on wages depends on the educational level. The orientation effect accounts for 11% to 20% of the decrease in wage premia[1]. The performance effect can explain the remainder of the difference.

Regarding the policy effect on the education distribution, the orientation effect predicts an increase in college enrollment so that students insure themselves against uncertainty. Wage rate changes further distort their post-secondary education decision. Specifically, the decreases in college graduates' wage premia relative to that of university graduates contribute to an overall increase in university enrollment in the new cohorts.

Our paper contributes to the literature in three ways. First and foremost, this is the first paper that studies the Ontario reform comprehensively. Some studies focus on double cohorts[2] given their educational choices: Krashinsky (2014) and Morin (2013) restrict their sample to be double cohorts who were admitted to universities, Krashinsky (2009) investigates wage differences among double cohorts who entered the labor market immediately after high school graduation. One problem with these studies is that educational attainment, namely the selection criteria for the study sample, could also be affected by the policy reform. Our paper extends the analysis along the following dimensions. First, we analyzed the policy's effect on students' labor market performance conditional on educational attainment and its impact on individuals' educational choices and the underlying channels. In addition, our treatment group includes all the cohorts who were exposed to the policy change (Cohorts born between 1985 and 1988[3]) to rule out the possibility of a temporary effect. Finally, the outcomes covered in this project are a lot more extensive compared with the previous literature, ranging from their high school graduation rate to their labor market performance

[1]The wage premia are relative to the average high school dropouts' wage.

[2]Double cohorts in the Ontario reform case refer to the last old cohort, who experienced the five-year program curriculum and the first new cohort, who were the first ones to take the new curriculum after the reform. Notice that the reform was implemented so that in 2003 the double cohorts graduate from high school in the same year, though there was a one-year difference when they entered school.

[3]Cohorts born after 1988 were affected by another Ontarian policy change, which increased minimum school leaving age to 18 starting in 2007.

measured in their later life cycle. All these aspects together help support a comprehensive evaluation of the policy change.

Secondly, this is the first paper to quantify the orientation effect and disentangle it from the performance effect in education. The estimated orientation effect is essential in understanding the policy reform and provides insights into future improvements in high school "curriculum" design. After all, compressing high school duration without changing the curriculum will not merely shift individuals' schooling decisions by one year as the policymaker would expect. We cannot ignore its impact on students' beliefs regarding their ability as it induces a long-run effect on students' post-secondary education decisions.

Finally, this paper sheds light on the role of school duration as an input in the educational production function. The number of years of schooling provided to a student is often endogenous and studies rely on exogenous shocks to school duration. One of the most studied policies of this type took place in Germany in 2001. Most federal states in Germany shortened high school length from 13 to 12 years to promote earlier high school graduation and earlier labor market entry. Büttner and Thomsen (2015) and Huebener and Marcus (2017) document decreases in math scores, increased grade repetition rates and lower final grade point average. Hofmann et al. (2017) and Dahmann and Anger (2013) find negative effects on psychological dimensions. Meyer and Thomsen (2016) and Meyer and Thomsen (2018) find evidence of delayed, rather than lower university enrolment rates and no changes in university dropout probabilities.

Although this German policy is in many ways similar to the 1999 Ontarian policy, no article analyzes how labor market outcomes are affected by it. In this study, we conduct a long-term labor market analysis. Moreover, the German policy is slightly different in that the curriculum was unchanged, which significantly increased students' learning intensity. In the Ontarian policy, as we will see in more detail, the curriculum was modified such that the learning intensity remained mostly unchanged. This allows for an analysis of schooling duration that is not contaminated by the effects of learning intensity.

The rest of the paper is structured as follows. Section 2 provides additional information regarding the Ontarian educational system and the policy change in 1999, Section 3 and 4 summarize the data and empirical results, Section 5 presents the the stylized structural model, Section 6 includes the calibrated results and some counterfactual analysis. Section 7 shows the results of a policy exercise we conduct using the structural model. Finally, Section 8 concludes.

2 Policy Description in Ontario

Before 1999, Ontarian high school students typically completed their high school studies in five years. All students were required to complete 30 credits. While there was no restriction on the types of courses required to obtain a high-school diploma, university-bound students had to complete at least six of these 30 credits from university-preparatory courses (called OACs - Ontario Academic Credits).

Although it was technically possible for all students to graduate in four years, for example, by completing eight courses per year during the first three years of high school and six OACs in the fourth year, the students enrolling in post-secondary education rarely did so. It was in part because course requirements made it difficult to complete all desired OACs within four years. Course requirements made it impossible to graduate in four years and complete six OACs without taking at least one course during summer semesters.

Moreover, university admissions were based on the best six OAC scores, regardless of the total number of OAC courses a student completed. Thus, students had an incentive to complete more than the required six OACs to improve their university admission chances. Indeed, between 20% to 25% (Casas and Meaghan, 1996) of students chose to repeat one or more OACs to boost their grades. This practice was called "OAC grazing" (Brady and Allingham, 2010). Typically, university-bound students completed eight credits per year in their first three years and six OACs per year for the subsequent two years, for a total of 36 credits. Thus, in practice, most students spent five years in high school. This pattern was especially true for university-bound students (Krashinsky, 2014).

Starting in 1999, the Ontario provincial government eliminated the fifth year of secondary education and mandated a new four-year system. The change enacted by the government instituted a standard four-year, 30 credit program for all students. This change affected university-bound students in particular. For these students, this educational reform implied a 1/5 reduction in the time spent in high school and a 1/6 reduction in the number of credits. This reduction in credits is mainly in university-preparatory courses, which are relatively more difficult.

This policy contrasts with the high school reduction policy in Germany in 2001. While that reform significantly increased student workload, the Ontarian policy reduced high school duration without significantly affecting the workload. This reform enables us to study the effects of schooling duration without worrying about picking up effects due to changes in workload intensity.

3 Data Description

This paper mainly uses the Labour Force Survey (LFS) and the National Longitudinal Survey of Children and Youths (NLSCY) to analyze the effect of shortening high school duration along multiple dimensions.

3.1 LFS

The Labor Force Survey (LFS) is an ongoing monthly survey used by the Canadian government to produce unemployment and other standard labor market indicators such as the participation rate and employment rate. We use LFS data covering the years 2000-2017 (the latest wave available).

The variables essential to model identification are the province where the respondent attended secondary education and their birth year. Since it is impossible to accurately observe which curriculum an individual experienced in Ontario, we assume all the individuals born after 1985 from Ontario were exposed to the new regime. The threshold in the birth year comes from the fact that students were required to be in Grade 1 the year when they turned 6 in Ontario and assuming the absence of grade repetition.[4] We further exclude individuals born after 1988 since the new compulsory school attendance law enforced since 2006 would contaminate the treatment effect otherwise. Cohorts born before 1976 would not enter the sample either because they were exposed to different secondary school structures before "OAC" started. On the other hand, there is no information about where the individual received their secondary education in LFS. Thus we use the province of residence instead.[5] To sum up, the treatment group defined in this project include individuals from Ontario and born between 1985 and 1988. Cohorts born between 1976 and 1984 in Ontario and all the cohorts born between 1976 and 1988 in other provinces in Canada[6] serve as the control group.

The LFS has large sample sizes and provides information both on educational attainment and on labor market outcomes. Moreover, it collects data on the treated cohorts until 2017. The data allow us to measure the policy effect in the long run. The education variables include a high school graduation indicator, the highest education level, and the highest grade completed for high school drop-outs. Moreover, the LFS also asks respondents' current educational activities, which proves to be essential to understanding the policy effect on individuals' postsecondary educational choices.

The labor market outcome information provided in the LFS is very rich. It helps bridge the gap between human capital accumulation and labor market outcomes. We mainly use the hourly wage, labor market status and hours worked in our analysis.

Because the attainment information in the LFS contains broad educational categories, we use an alternative data source to impute educational attainment more precisely and avoid measurement error biases.

[4] This definition is consistent with Krashinsky (2014), where the author documented the age distribution of the double cohorts was generally separated at 18.75 years of age measured at the beginning of October in 2003.

[5] In Survey of Labor Income Dynamic (SLID), the correlation between living in Ontario at age 23 and receiving secondary education in Ontario is around 0.96, and it decreases to 0.93 when measured at age 28.

[6] We exclude the province of Quebec because of its unique education structure.

We use the Survey of Labor Income Dynamic (SLID), which contains similar yet more detailed information regarding individuals' educational attainment. Covering the years 1993 to 2011, SLID provides annual panel survey data on labor market activity and income information of individuals living in Canada.[7] Each respondent is tracked for six consecutive years.

The survey further asks the exact number of years of postsecondary education the respondent received and certificates individuals received in addition to the highest education level. This detailed information makes SLID a reliable resource to measure the population's educational distribution, which helps correct the measurement error in LFS data. We provide details of the imputation of educational attainment using SLID in the Appendix.

3.2 NLSCY

The National Longitudinal Survey of Children and Youths (NLSCY) is a longitudinal study of Canadian children and follows their life from birth to early adulthood. It is designed to provide information regarding the factors influencing a child's social and behavioral development. NLSCY started in 1994-1995 and continued biennially until 2008-2009 (Cycle 8).

In NLSCY, cohorts born between 1982 and 1988 were included in the analysis,[8] among which those born between 1985 and 1988 and who went to secondary school in Ontario are the treated group. Thanks to the panel feature, NLSCY documented rich historical information regarding students' school performance. The relevant variable to this project is the mathematics test score of are 14 to 15 years old students, right before entering secondary school. In the structural model analysis, we use students' mathematics test scores to approximate their initial beliefs regarding ability. We also investigate student high school graduation rates using NLSCY data to complete the LFS data analysis.

4 Empirical Results

We start by estimating Eq. (1):

$$Y_{ijk} = Cons + \text{Policy}_{ijk}\theta + \alpha_j + \beta_k + X'_{ijk}\gamma + \varepsilon_{ijk} \tag{1}$$

where individuals are indexed by i, j indicates province and k stands for cohort. Variable Policy_{ijk} is the treatment dummy indicating if the individuals belong to the Ontarian "new cohort", variable α_j captures

[7]Longitudinal estimates are available starting in 1996 and up to (including) 2010

[8]Cohorts born between 1982 and 1988 happens to be part of the original cohort in NLSCY (children who were $0 - 11$ years old in 1994).

province fixed effect, β_k represents cohort fixed effects defined by birth year and X_{ijk} includes other controls. The coefficient θ identifies the reform effect by comparing the difference between pre- and post-reform cohorts in Ontario to those in the rest of Canada. The choice of controls depends on data availability and varies according to data source.

The dependent variables used in Eq. (1) include the students' high school graduation status, post-secondary education attendance dummy and highest educational attainments measured at age $28-36$. We present these results in detail in the following section.

4.1 High school graduation rate

What does the theory predict regarding the effect of shortening high school length on high school graduation rates? In the human capital framework, two opposing forces play a role. On the one hand, high school graduation's opportunity cost decreases as the forgone income of schooling shrinks under the new structure. On the other hand, the increased learning intensity would require more effort for schooling, discouraging students from finishing high school. These two forces affect high school graduation decisions in opposite ways.

Table 1 presents the policy effect on high school graduation decisions using both the data sources. Column 1 shows that shortening high school duration negatively affects the high school graduation rate by 1.4%. This negative effect implies the increased learning intensity due to the curriculum change outweighs the lowered opportunity cost. The policy effect estimated using NLSCY is also negative yet slightly larger. The quantitative gap could be due to the difference in the population each data set represents.[9] It could also be attributed to the fact that the treatment group in LFS is defined using the province of residence rather than the province of high school attendance. This biases the estimates towards zero, as predicted by the measurement error literature. Statistically, the difference in regression estimates on the high school graduation rate using LFS and NLSCY is not significant.

To explore possible heterogeneity in policy effect among students with different abilities, we categorize them into two groups based on their overall academic performance in Grade 7 or 8. Their ability type is determined before entering high school, making the classification orthogonal to the policy change. In the sample, the high-ability type consists of around 70% of students whose overall performance was described as "very well" or "well", and the low-ability type include the rest 30% who performed as "average or below" in the corresponding school year.

Column 3 in Table 1 suggests that on average, low-type students are 12.8 p.p less likely to graduate from

[9]The population of LFS is the Canadian population in 2011, while that of NLSCY is the children population whose age is between 0-15 in 1994.

Table 1: Policy effect on high school graduation rate

	(1) LFS	(2) NLSCY	(3) NLSCY
Policy	-0.014**	-0.070**	-0.044
	(0.005)	(0.029)	(0.032)
Policy · Low type			-0.038
			(0.050)
Low type			-0.128***
			(0.024)
N	20800	2801	2577
R^2	0.069	0.080	0.092

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is high school graduation indicator. Source I: LFS 2000-2017. Sample includes all the individuals born between 1975 and 1988, with age $22-24$ and who were not attending high school in ref month. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort, province of residence fixed effect. Source II: NLSCY Cycle 1 - Cycle 8. Sample includes all the individuals born between 1982 and 1988 at age $20 - 21$. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, household social economic score at age $14 - 15$, single parent status at age $14 - 15$, year, cohort and province of high school fixed effect. Standard errors are in the parentheses. Variance is clustered within province $\times$ birth year.

high school when measured between the ages of 22 and 24. The estimates do not support the hypothesis that students with different ability types are affected differently when high school duration is reduced.

To sum up, it turns out that individuals' high school graduation decision is negatively affected by the policy. This suggests that the increased learning intensity dominates the decrease in the attached opportunity cost.

4.2 Effect on the post-secondary education choice

Conditional on graduating from high school[10], students have three choices: to enter the labor market directly, to enroll in post-secondary education directly or to enroll later. This subsection will present how this policy affects post-high school educational and labor market decisions.

Table 2: Policy effect on psec activities, LFS

	(1)	(2)	(3)	(4)
	Not in school	School Attendance		
		High school	College	University
HS grad age				
Policy	-0.046***	0.012	0.003	0.032**
	(0.013)	(0.009)	(0.013)	(0.013)
N	73523	73523	73523	73523
R^2	0.087	0.031	0.021	0.125
HS grad age +1				
Policy	-0.055***	0.004**	0.054***	-0.000
	(0.009)	(0.002)	(0.008)	(0.012)
N	46410	46410	46410	46410
R^2	0.09	0.002	0.015	0.116

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is a school attendance indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort and province of residence fixed effect. Standard errors are between parentheses. Variance is clustered within province × birth year.

Although there is no way to track individuals longitudinally in LFS, we can explore the variation in the fraction of each educational activity measured at different times after high school graduation.

Four categories are documented in LFS: "Not registered as student", "Attending secondary education", "Attending college" and "Attending university"[11]. Here, "college" stands for the 2-year college in the U.S.

[10]In the sample, we do observe respondents who participated in college training without graduating from high school. They account for less than 5% of the observations and are therefore omitted here.

[11]There is the fifth category as "Attending other type of educational institutions". We omit this because it represents no more than 2% of the population and the specific type of education remains unclear

context, including community college or vocational school and "university" means 4-year college.

We assume that high school graduation age in other provinces and the new system in Ontario is 18 (i.e., without repeating years and enrolled in Grade 1 at age 6), and 19 in Ontario in the old structure. Conditional on high school graduation,[12] the fraction of the population not enrolled in any additional schooling has decreased significantly. In response to the reform, university enrollment increased by 3%. When it comes to college, enrollment increased with some delay. Finally, when measured one year after high school graduation, college enrollment increased by 5%.

In short, the policy increases the fraction of individuals enrolled in post-secondary education conditional on graduating from high school. Moreover, the policy boosts enrollment into college and university within one year after high school graduation. We can rationalize this in the following way: the less time they spend in high school, the less informed students are regarding their ability level. The increased uncertainty about whether one could finish college or university, together with the considerable wage premium associated with obtaining the degree, encourages more students to try post-secondary education.

4.3 Long-run effect on educational attainment

Among the two data sources, only LFS could shed light on the long-run effects on educational attainments in the population[13]. As we mentioned in Section 3, we corrected the measurement error in LFS by comparing the educational distributions in LFS with those in SLID[14]. For the outcome variable, the correction could only be applied on the aggregate level, so we use both the original and the corrected aggregate educational distribution for each province $\times$ cohort to investigate the policy effect.

The first panel in Table 3 presents the policy effect estimated using original data from LFS and the second panel includes the corrected version. Quantitatively, the policy effect is larger using the corrected the data, yet the qualitative conclusion is the same. From Table 3, shortening high school duration increased high school dropouts significantly and decreased the proportion of people holding a college degree in the long run. For other categories, though imprecisely estimated, the economic meaning is consistent with what we previously observed; the overall fraction of people receiving post-secondary education upon graduating from high school increased. Among them, it seems there are more individuals who obtained university degrees and fewer obtained college certificates. We see an increase in the number of post-secondary dropouts as well.

To sum up, shortening high school length affects individuals' educational choice in high school and their

[12]Only data collected in months 9-12 were included at graduation year because students would not report attending school in summer months, regardless of their status during school year period

[13]The NLSCY ends when the treatment group subjects were in their early twenties.

[14]The detailed process is presented in the appendix.

Table 3: Policy effect on educational attainments at age $28 - 36$

	(1) HS dropouts	(2) HS grads	(3) Some psec	(4) College	(5) University
Original					
Policy	0.013**	-0.007	0.003	-0.021*	0.011
	(0.005)	(0.008)	(0.005)	(0.011)	(0.010)
N	126	126	126	126	126
R^2	0.762	0.807	0.840	0.869	0.872
Corrected					
Policy	0.013**	-0.009	0.005	-0.023*	0.012
	(0.005)	(0.009)	(0.008)	(0.012)	(0.011)
N	126	126	126	126	126
R^2	0.762	0.758	0.840	0.869	0.871

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: LFS 2000-2017. Dependent variable is the proportion of corresponding education level of each province × birth year group when measured at age $28 - 36$. Cohorts born between 1975 and 1988 are included. Quebec is excluded from the sample due to its unique educational structure. "Corrected" means the distribution is corrected by the misreport matrix derived previously.

further post-secondary educational decision. Overall, the policy promoted post-secondary enrollment. More individuals attended university immediately after high school graduation. However, the increase in college enrollment occurred with a one year delay. In the long run, the increase in post-secondary enrollment did not persist. When measured at ages $28-36$, there are more high school dropouts than before. Moreover, the boost in college and university enrollment did not generate more individuals with a post-secondary certificate in the population.

4.4 Effect on labor market compensation

Among all the labor market outcomes, we first examine the policy effect on the labor market participation rate, unemployment rate, and self-employment rate. None of those indicators are affected by the policy, which suggests that the overall selection of employees is not affected.

We now turn to the policy effect on wages. We estimate Equation (1) with log of hourly wage as the dependent variable. The sample includes cohorts born between 1981 and 1988 who worked as employees measured between the ages of 28 and 32 and were not attending school during the reference period. We further exclude individuals whose real hourly wage rate is below two dollars. For individual characteristics variables, we include only gender and age dummies to ensure a sufficiently large sample size.

The policy induces a 4.4% drop in hourly wage rate (Column (2) Table 4). Notice that the policy effect estimate here captures both the change in overall educational distribution and the change in wage rate within educational levels.

Next, we will present the regression result of heterogeneous policy effect on log wage rate across educational levels, as indicated in Equation (2),

$$\begin{aligned} Y_{ijk} =& Cons + X'_{ijk}\gamma + \\ & S_{ijk}\beta + \text{Policy}_{ijk}\theta + \alpha_j + \beta_k + \\ & S_{ijk} \cdot \text{Policy}_{ijk}\beta_s + S_{ijk} \cdot Prov_j \alpha_{sj} + S_i \cdot cohort_k \beta_{sk} + \varepsilon_{ijk} \end{aligned} \tag{2}$$

where S_{ijk} represents the highest educational level individual i obtained, α_j and β_k indicate province and cohort fixed effect, and Policy_{ijk} is the treatment dummy that equals one if the individual i was born after 1985 and came from Ontario. Compared to Equation (1), we further include the educational attainment and its interaction with the policy as well as province and cohort dummies. Therefore, the coefficients β and β_s's together could capture how much wage is affected by the policy reform within each educational level.

When estimating Mincer wage equation (Column 1) what we obtain from LFS is consistent with the literature. The high school certificate premium is around 8.4%, and the university diploma provides a premium as high as 44%. Column (3) documents the heterogeneity in the policy effect across educational levels. Though we reject heterogeneity, all the educational levels higher than high school graduates were negatively affected, and the effect is statistically significant.

Keep in mind that all the results presented above use the educational attainment variable with non-random measurement errors. Following Savoca (2000) and together with the misreport matrix we estimated[15], Figures (1) and (2) show the corrected return to education and policy effect.

As shown in Figure 1, the average return to high school grads and college degree is lower after correcting for measurement error. The premium for obtaining a university degree is higher due to the exclusion of university dropouts who report holding university degrees by mistake. The return for post-secondary dropouts and college graduates is not affected much quantitatively in the end.

Figure 2 presents the policy effects on log hourly wage rate across different educational levels. The average wage for high school dropouts was not affected by the policy. The wages of those who graduated successfully, conditional on their educational attainment, are negatively affected. Among people who obtained a high school diploma, shortened high school length decreases their average wages by 5% to 10%, depending on educational attainment. This is consistent with the findings in Krashinsky (2009), who studied the double

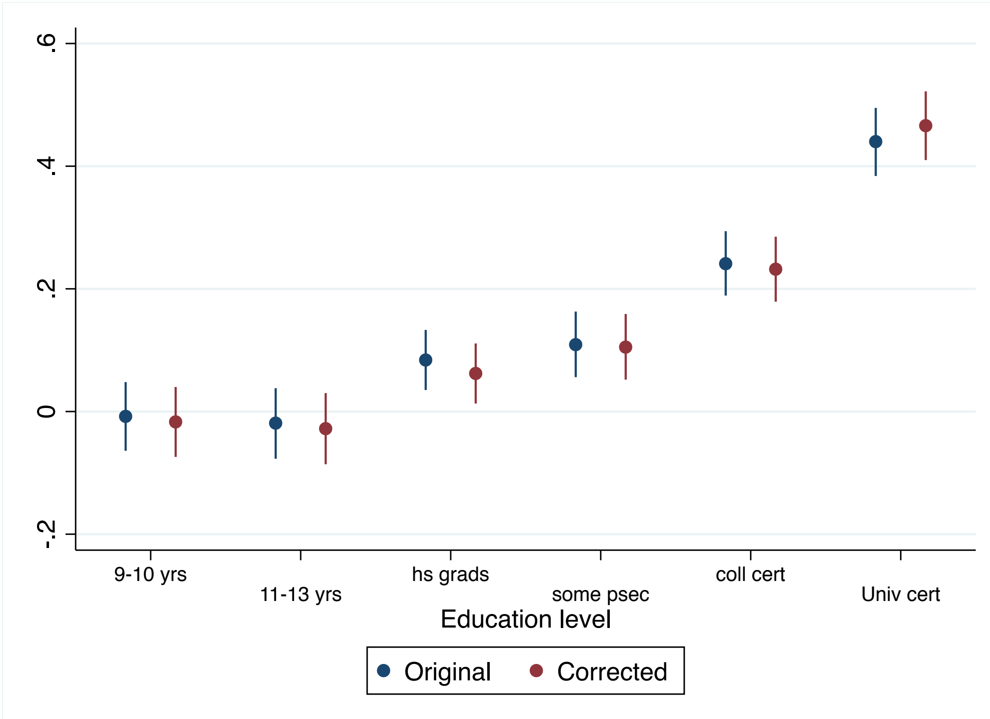
[15]See section 9.4

Table 4: Policy effect on log hourly wage rate

	(1) $log(w)$	(2) $log(w)$	(3) $log(w)$
Policy		-0.044***	0.046
		(0.006)	(0.095)
Policy · 9-10 yrs of edu			-0.032
			(0.112)
Policy · 11-13 yrs of edu			-0.053
			(0.095)
Policy · HS grads			-0.111
			(0.091)
Policy · Some psec			-0.08
			(0.097)
Policy · College			-0.067
			(0.094)
Policy · University			-0.105
			(0.099)
9-10 yrs of edu	-0.008		0.038
	(0.029)		(0.113)
11-13 yrs of edu	-0.019		0.016
	(0.029)		(0.107)
HS grads	0.084***		0.13
	(0.025)		(0.105)
Some psec	0.109***		0.147
	(0.027)		(0.123)
College	0.241***		0.278**
	(0.027)		(0.111)
University	0.440***		0.524***
	(0.028)		(0.109)
N	203000	203000	203000
R^2	0.164	0.05	0.17

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is log hourly wage. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age $28 - 32$, who are not attending school in ref month. Quebec is excluded from the sample due to its unique educational structure. Omitted educational level is "less than 8 years of edu". Other control variables include age dummy, gender, provincial dummy, year dummy and month dummy. Standard errors are in the parentheses. Variance is clustered within province × year.

Figure 1: Return to education: log hourly wage rate



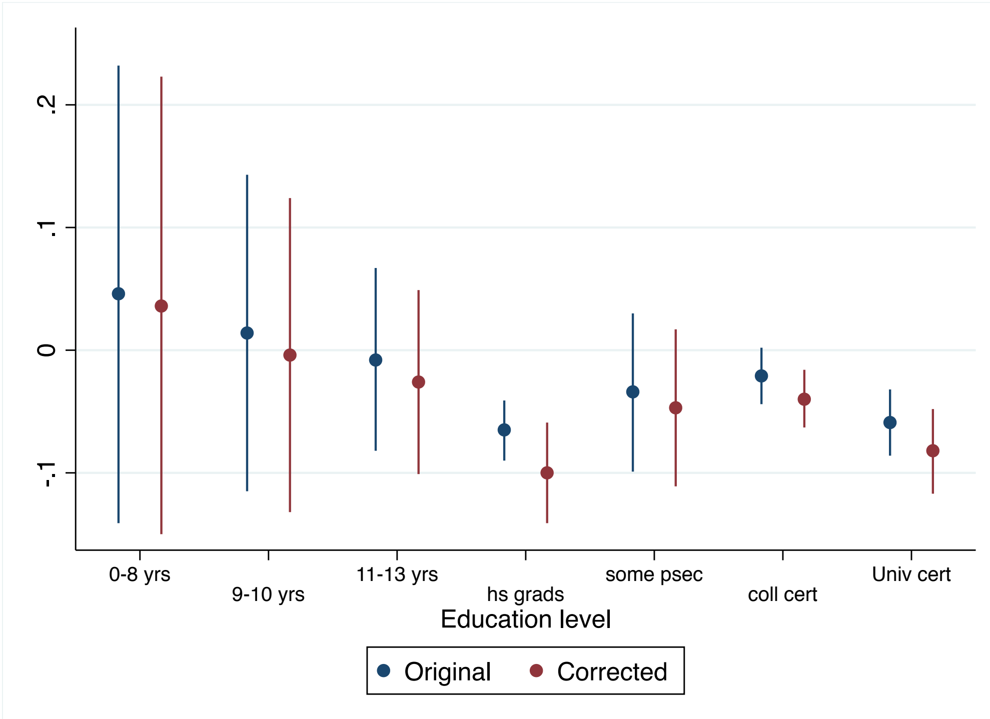
Source: LFS 2000-2017. The dependent variable is log (real) hourly wage rate. The reference level is "0-8 years of education". The sample includes individuals born after 1980, at age 28-32 yrs old, and who work as employees. Individuals whose hourly wage is below 2$ or who are attending school during ref period are excluded. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, age and year dummies. Standard errors are clustered at the province-year level. "Original" shows the estimates using original data in the LFS. "Corrected" shows the estimates after correcting for measurement error.

cohorts who entered the labor market after high school graduation. He documents that four-year graduates earn around 10% less than the five-year ones upon graduation, and this effect persisted for less able graduates two years after graduation.

To summarize, the policy boosts enrollment into university and college within one year after high school graduation. Unfortunately, the increase in post-secondary enrollment is canceled out by an increase in the dropout rate, leaving the educational attainment distribution unchanged when measured at ages $28-36$. When it comes to wages, the average hourly wage rate decreased by around 5% to 10% for individuals who have graduated from high school under the new system.

As analyzed previously, reduced high school duration affects students' educational decisions and labor market compensation through two channels: "Orientation effect" and "Performance effect". These reduced-form estimates reflect combined effects. However, the exact quantitative decomposition of these two channels

Figure 2: Policy effect on log hourly wage



Source: LFS 2000-2017. The dependent variable is log (real) hourly wage rate. The reference level is "0-8 years of education". The sample includes individuals born after 1980, at age 28-32 yrs old, and employees. Individuals whose hourly wage is below 2$ or who are attending school during the reference period are excluded. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, age and year dummies. Standard errors are clustered at province-year level. "Original" shows the estimates using original data in the LFS. "Corrected" shows the estimates after correcting for measurement error.

is beyond the empirical identification strategy discussed so far. In the next section, we will present a stylized model where individuals make schooling decisions sequentially as they learn their ability to further evaluate the underlying mechanism through which high school duration influence students.

5 Model

This section presents a dynamic model where agents learn their abilities as they make schooling decisions.[16] Individual ability can be either high or low. Let $\alpha \in \{0, 1\}$ denote the ability level, with $\alpha = 0$ denoting low type. Ability is not observable by the individuals. Instead, they start grade ten with some initial belief of their ability and observe signals through school experience. They then update their belief

[16] We employ a model similar to Trachter (2015). We incorporate the ability learning process in high school and omit the possibility of transferring from colleges into universities.

accordingly, denoted by $p = Pr(\alpha = 1) \in [0, 1]$.

There are three school programs in the model: high schools, colleges and universities. The general model structure is as follows. At the beginning of each time period, individuals who are in school j could choose whether to stop schooling, join the labor force, or continue their education. If they accumulate the necessary amount of credits T^j, students receive a diploma and graduate. For each period spent in school, students observe one signal g, accumulate credits s and then update their belief p. After finishing high school, students enroll in university or college or start working. If they continue post-secondary education, they pay the tuition cost τ.

Once they enter the labor market, the individual earns wages corresponding to his ability α and his educational level. Thus, individuals' ability type is only fully revealed once they join the labor market. This simplifying assumption makes it easier to solve the model without computing firms' beliefs of individual ability or signaling, which would severely complicate matters. Specifically, the education level is determined by the graduation status GS and the highest school level the individual is enrolled in $j \in \{hs, C, U\}$. Then the wage function can be described as follows

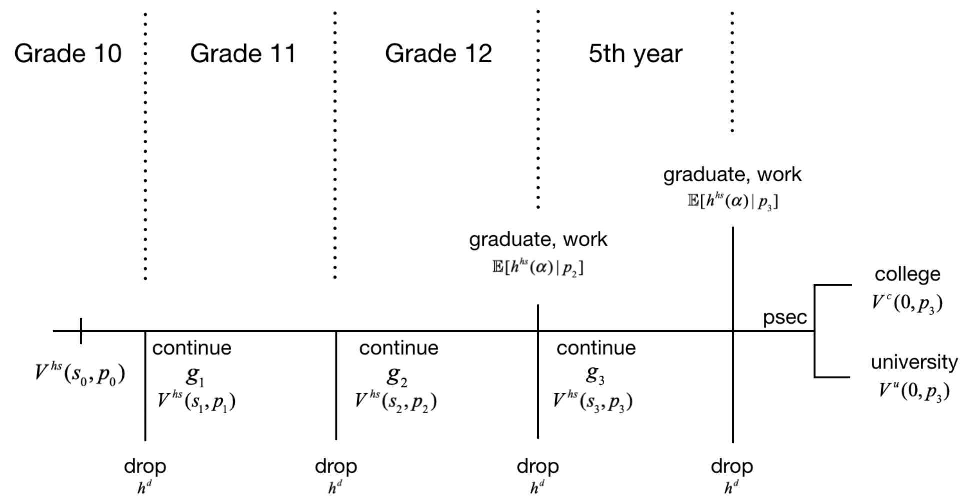
$$h(GS, j, \alpha) = \begin{cases} h^j(\alpha) & GS = 1, j = \text{hs, C, U} \\ h^d & GS = 0, j = \text{hs} \\ h^{hs}(\alpha) & GS = 0, j = \text{C, U} \end{cases} \tag{3}$$

with $h^j(1) > h^j(0) > h^d$ for all j and $h^U(\alpha) > h^C(\alpha) > h^{hs}(\alpha) > h^d$ for all α. Namely, for any given ability level, graduating from university implies higher wage profiles than graduating from college, and college graduates have higher wages than high school graduates. Furthermore, wage profiles are increasing with individuals' true ability level. We assume the post-secondary dropouts end up with the same wage profile as the high school graduates. Thus, in our model, the observed average wage premium of dropouts compared with high school graduates is fully attributed to the composition effect. We also assume that the wage profile for high school dropouts does not depend on their ability α. One justification for this is that we expect few high-ability individuals to drop out of high school. High school dropout wage profile is then determined solely by low-types mainly.

Figure 3 illustrates the Ontarian old secondary school system following the aforementioned timeline, with some complications. At the beginning of each period up to Grade 12, students choose to drop out or continue school for another year. However, after completing Grade 12, they face three possible choices: to drop out, to graduate or to continue the fifth year. Those students who have accumulated enough credits[17] could either choose to graduate immediately and enter the labor market or to stay for the fifth year for post-secondary

[17] Think of these students as those who have succeeded all the required courses up to Grade 12.

Figure 3: Old secondary education structure



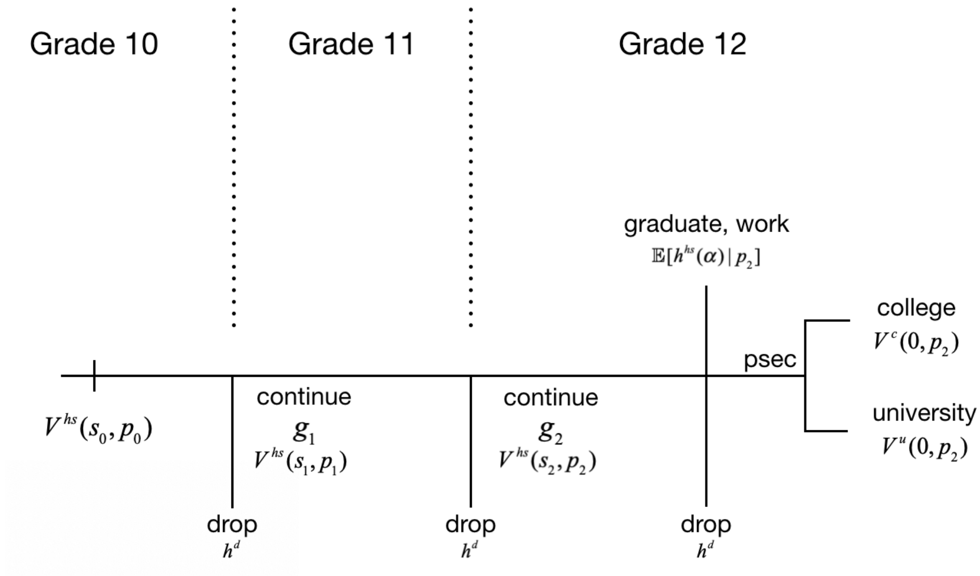
Note: $V^j(s,p)$ is the (expected) value of continuing school program j with s accumulated credits and belief p, where $j \in \{hs, C, U\}$. g_t stands for the signal received in period t and s_t means the accumulated credits in period t. h^d is the wage profile received by high school dropouts. $\mathbb{E}[h^{hs}(\alpha)|p_t] \equiv p_t h^{hs}(1) + (1-p_t)h^{hs}(0)$ is the (expected) wage profile received by high school graduates given belief p_t.

education. Were it not for the fifth year, those who failed to accumulate enough credits after Grade 12 could only choose to drop out of high school. As a result, the OAC year prepares students for post-secondary education and serves as an extra opportunity to make up failed courses to obtain a high school degree in the model. After the fifth year, those students who accumulated enough credits could graduate from high school and start their careers. Those who obtained the fifth year credit could further choose enrollment into post-secondary education. Thus under this regime, students would spend five years in high school only for two reasons. The first one is to be qualified for post-secondary education as designed by the policymaker. The other is to use the extra year to obtain the high school graduation certificate. In this sense, the fifth year serves as a buffer year to help those struggling with graduating from high school after four years.

This extra year advantage disappeared as the new four-year system was implemented in 1999. In other words, students need to graduate within four years and failing a year leads to an automatic dropout. This modeling choice reflects the increase in learning intensity for low ability students, highlighted in Section 4. As shown in Figure 4, all the students in the new system have to accumulate enough credits to graduate after Grade 12 and make post-secondary education enrollment decisions.

Once enrolled in post-secondary education, the timeline can be illustrated recursively in Figure 5, where institution j could be either college or university. College and university are different in three dimensions:

Figure 4: New secondary education structure



Note: $V^j(s,p)$ is the (expected) value of continuing school program j with s accumulated credits and belief p, where $j \in \{hs, C, U\}$. g_t stands for the signal received in period t and s_t means the accumulated credits in period t. h^d is the wage profile received by high school dropouts. $\mathbb{E}[h^{hs}(\alpha)|p_t] \equiv p_t h^{hs}(1) + (1-p_t)h^{hs}(0)$ is the (expected) wage profile received by high school graduates given belief p_t.

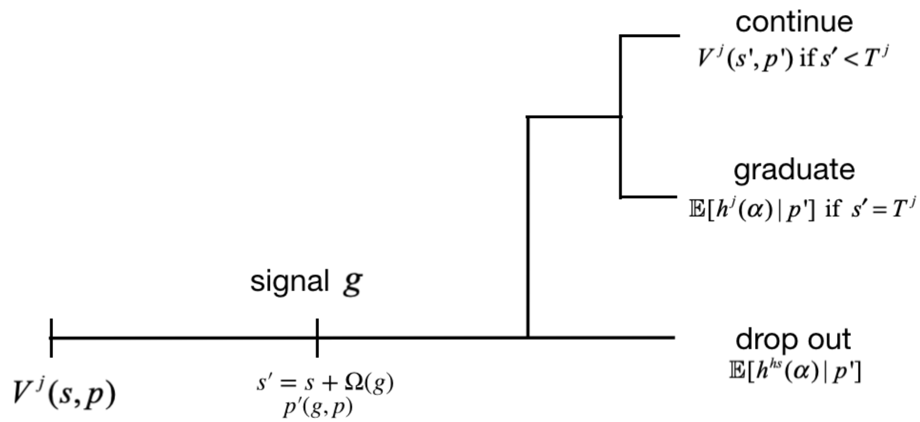
the duration, the tuition cost, and the wage premium of corresponding graduates.

Having established the timeline of individuals' decisions in different school programs, we present more model specifications. Assume the cost of attending school j is τ^j, which is constant for all the individuals, with $\tau^U > \tau^C$ and normalize $\tau^{hs} = 0$[18]. A student could graduate from institution j after accumulating required credits T^j, with $T^U = 4$ and $T^C = 2$. We also assume that students cannot be enrolled in high school longer than the required number periods. High school length T^{hs} changes from five to four years upon implementation the reform. The evolution of credit accumulation is tied to signals received during the school program, which we present shortly.

As to the learning parameters, students who are enrolled in school receive a signal g indicating their ability, and update their belief based on which they make further work-school decision in the following period. Let signal g have ability specific probability density function (PDF) $f(g|\alpha)$. We further assume that the ratio of densities satisfy the monotone likelihood ratio property (MLRP), i.e. $\frac{f(g|\alpha=1)}{f(g|\alpha=0)}$ is defined and $\frac{f(g_1|\alpha=1)}{f(g_1|\alpha=0)} \geq \frac{f(g_0|\alpha=1)}{f(g_0|\alpha=0)}$ for any $g_1 > g_0$. This assumption guarantees that the high ability students are more

[18]There is only national-level university tuition cost data available from Statistics Canada. According to Vaillancourt (1995), the ratio of the average annual cost of college to university is 0.38, which is used to calibrate the cost parameters in the structural model.

Figure 5: Post-secondary education structure



Note: $V^j(s,p)$ is the (expected) value of continuing school program j with s accumulated credits and belief p, where $j \in \{C, U\}$. g is the signal received and $\Omega(g)$ governs the revolution of credits given signal received. $\mathbb{E}[h^{hs}(\alpha)|p] \equiv ph^{hs}(1) + (1-p)h^{hs}(0)$ is the (expected) wage profile received by college or university dropouts given belief p. $\mathbb{E}[h^j(\alpha)|p] \equiv ph^j(1) + (1-p)h^j(0)$ is the (expected) wage profile received by graduates from school program j given belief p, where $j \in \{C, U\}$. T_j stands for the duration of program j.

prone to receiving better signals than low ability students.

The evolution of credits s_{t+1} is assumed to be a function of current signal g_t and the accumulated credits s_t,

$$s^j_{t+1} = \begin{cases} s^j_t + \Omega(g_t) & \text{if } s < T^j \\ s^j_t & \text{if } s \geq T^j \end{cases} \tag{4}$$

and $\Omega(g_1) \geq \Omega(g_2)$ for any $g_1 > g_2$. Namely, given the accumulated credits s_t, individuals receiving a high signal would accumulate at least as many credits as those receiving lower signals. The number of credits s_t in the model is not exactly the academic credits students earn in school. Rather, it is an indicator of students succeeding in progressing into the next grade level.

For simplicity, we assume that the signal evaluates the students' overall academic performance in a school year. There are three possible values of $\{F, P, G\}$, representing Fail, Pass and Good. Thus

$$\Omega(g) = \begin{cases} 0 & \text{if } g = F \\ 1 & \text{if } g = P \\ 1 & \text{if } g = G \end{cases} \tag{5}$$

meaning that students who fail in a school year cannot accumulate credits.

Having observed the signal, the students update their belief according to Bayes' rule as follows,

$$p_t \equiv pr(\alpha = 1|g_t, p_{t-1}) = \frac{1}{1 + \frac{pr(g_t|\alpha=0)}{pr(g_t|\alpha=1)} \frac{1-p_{t-1}}{p_{t-1}}} \tag{6}$$

where p_{t-1} is the prior, g_t is the signal received and $pr(g_t|\alpha)$ denotes the probability of receiving signal g_t conditional on the true ability type α. Then the PDF of future signal is

$$H(g_{t+1}|p_t, g_t) = pr(g_t|\alpha = 0)(1 - p_t) + pr(g_t|\alpha = 1)p_t \tag{7}$$

In order to reduce the number of parameters to be estimated, we make the following simplifying assumptions about the learning parameters:

$$pr^{hs}(g = F|\alpha = 1) = 0 \tag{8}$$

$$pr^{C}(g = G|\alpha = 0) = 0 \tag{9}$$

$$pr^{U}(g = G|\alpha = 0) = 0 \tag{10}$$

In other words, high ability types cannot fail in high school (although they may fail in college or university), whereas low ability types cannot excel in college nor in university (although they may excel in high school). This means that drawing an F signal in high school fully informs the recipient that they have low ability. Drawing a G signal in university or college fully informs the recipient that they have high ability. We calibrate the remaining learning parameters without imposing any restrictions.

The model starts with individuals at the age of sixteen[19] who have finished Grade 10 with belief p_0, they choose whether to drop out of high school (with ten years of schooling) or to continue Grade 11 to maximize their expected lifetime income. Let $W(t, h(GS, j, \alpha)) = \sum_{\tau=t}^{T} \frac{h(GS,j,\alpha)}{(1+r)^{\tau-t}}$ be the value of joining the labor force with wage profile $h(GS, j, \alpha)$. Here, time serves as a state variable that reflects different labor market entry ages.

Table 5 lists the exogenous values used to calibrate the structural model. We set the risk-free rate to 1.8 percent, implying a yearly interest discount factor of 0.98. The tuition cost of high school is normalized to be zero. The tuition cost of university equals the sum of the average undergraduate tuition fee of full-time Canadian students and the average accommodation cost, obtained from Statistics Canada. The tuition cost of college is not available directly and is calculated as 0.4 times the average tuition cost of the undergraduate program following Vaillancourt (1995)[20]. We assume students need to complete two school years to graduate from college and the university program lasts for four years. To obtain the distribution of students' initial belief, we use their math score measured at age 14-15 conducted by NLSCY.

[19] Sixteen is the minimum age to leave school in Ontario before 2008.

[20] We normalize all monetary values using the wages of high school dropouts h^d in the calibration exercise.

We jointly calibrate the parameters governing the wage profile $h^i(\alpha)$ as well as the learning process $q^j_\alpha(g)$ to match the school distribution and wage distribution observed using LFS data, as well as a set of micro-moments using NLSCY data.

Table 5: Parameter values taken from other sources

	Parameter	Value	Source
Tuition Cost($)	High school	0	
	College	1487.84	Vaillancourt(1995)
	University	4523.35	StatsCan
Duration(yrs)	College	2	
	University	4	
Other	r	1.8	
	Age span	$16-65$	
	p_0	Math score at age $14-15$	NLSCY

6 Calibration

6.1 Calibration Method and Moments

Before showing the calibration results, let us briefly describe the algorithm. The method employed is the simulated method of moments (SMM). Denoting $\mathbf{V}$ as the variance-covariance matrix of the data moments, we use the optimal weighing matrix $\boldsymbol{\Omega}$, where each element $\boldsymbol{\omega_{ij}} = 1/\mathbf{v_{ij}}$. We then minimize the weighted distance between the vector of simulated moments $\mathbf{s}$ and the vector of observed moments $\mathbf{m}$:

$$(\mathbf{m}-\mathbf{s})^T\boldsymbol{\Omega}(\mathbf{m}-\mathbf{s}) \tag{11}$$

Therefore, the algorithm puts more weight on minimizing the distance to moments that are observed with relatively small variances.

The moments in question are of three types: wage moments, education distribution moments and micro-moments. The wage moments we attempt to match are the average hourly wages in Ontario for high school graduates, postsecondary dropouts, college graduates and university graduates, both before and after the reform. The education distribution moments we match are the percentage of grade ten dropouts, grade eleven to thirteen dropouts, high school graduates, college graduates, university graduates and postsecondary dropouts both before and after the reform[21].

Aside from the macro moments, we further enhance our model using individual moments related to their educational attainment. Indeed, NLSCY keeps track of the respondents' progress in school every other year till age $20-21$: their educational choice, the school performance if still enrolled in school, and their household situation etc. Together with other demographic information, we integrate the micro-level moments in the following way.

First, we run a probit regression on individual initial school choice $d_{10,i}$ as follows:

$$Pr(d_{10,i} = 1) = \Phi(X_i\beta) \tag{12}$$

where $d_{10,i}$ takes value of 1 if the student continues high school at age 16 (After finishing Grade 10) and 0 otherwise, and X_i includes students overall performance in Grade 10 at school, whether they are living with their parents in at age 15, their household socio-economic status at age 15[22], and their gender. Since the school continuation decision results from comparing the students' initial belief and the threshold $\tau_{t=0}$, which is endogenously determined in the model, we use the estimated result from Eq.(12) to predict individuals' initial belief as indicated by Eq.12

$$\hat{p}_{0,i} = (1 + e^{-X_i\hat{\beta}+e_i})^{-1} \tag{13}$$

where $\hat{\beta}$ is the estimated coefficient from Eq.(12) and e_i is drawn such that the computed belief matches the observed educational decision of the individual at age 16[23]. Our model ultimately uses this initial belief as a starting point to simulate possible educational paths of each individual in our sample, based on which we compute the simulated probability of this individual ending up with one of the following educational attainments: dropping out of high school, graduating high school, enrolling into college or enrolling into university. In the end, we calculate the micro-moments as the difference between the simulated probabilities and the actual educational attainment for each individual observed at ages 20-21.

In summary, our model uses eight wage moments, thirteen educational distribution moments and 406 micro-moments to estimate twelve wage parameters, nine learning parameters, the proportion of high type

[21]For the old cohort, we match college and university dropouts separately. For the new cohort, due to data limitations, we can only match postsecondary dropouts without distinguishing between college and university dropouts .

[22]In NLSCY, Socio-Economic Status (SES) is calculated for each household assigned to each selected child in that household. It was derived from five sources: the level of education of the person most knowledgeable (PMK) and that of the spouse/partner, the prestige of the PMK and the spouse/partner's occupation and the household income.

[23]For example, if the individual in question dropped out, we make sure that e_i is drawn such that $\hat{p}_{0,i} < \tau_{t=0}$.

individuals in the population and the standard deviation of the initial belief. We present the calibrated parameters in Section 6.3.

6.2 Identification

Next, we discuss the model identification. Notice that, in our model, the learning parameters are the same for the pre- and post-reform cohorts, whereas the wage rates are particular to each of them. Therefore, data from both cohorts help us pin down the learning parameters, while data from each cohort identify the respective wage rates.

The main difference between the pre- and post-reform systems is the length of time required to complete high school. This has two effects on students. First, it reduces the opportunity cost of finishing high school by one year. Second, post-reform students receive fewer signals regarding their ability, so they are more uncertain about their ability when completing high school.

In our model, this first channel boosts high school graduation rates, particularly for low ability types. This, coupled with the added uncertainty regarding own ability types brought about by the second channel, means that more low-ability types enroll in college and universities, increasing the dropout rates. Another effect is that the ability composition of university graduates becomes more low-ability heavy. This ultimately brings down the average wage earned by university graduates and decreases the university wage premium.

Thus, we use the observed average wages for different educational categories to pin down wage rates and each educational category's ability composition. Wage rates are unobserved to us, but they, coupled with the informativeness of ability signals in high school and university, drive student decisions to either continue schooling or drop out. If university or college wage rates, conditional on student beliefs about their type are high enough, this incentivizes students to enroll in postsecondary education.

By changing wage rates and the informativeness of signals received while in school, we can shift students' distribution across educational categories by ability type. In this way, we can match the observed educational distributions and average wages before and after the reform. We provide more detailed information regarding identification in the Appendix.

6.3 Calibration Results

Table 6 presents the calibrated wage differentials of graduating from each program relative to entering the labor market as high school dropouts. Within ability type, the calibrated wage rate increases with the educational level. Fixing the educational level, high type individuals earn a higher wage rate. The wage targets used in the data are measured when individuals are $28 - 32$ years old. Thus, these estimates account

Table 6: Calibrated wage parameter values

Ability α	HS drop	High school high	High school low	College high	College low	University high	University low
Hourly wage, Pre (\$)	14.54[a]	16.79	15.55	22.74	16.60	25.31	17.05
$\frac{h^j(\alpha)-h^d}{h^d}$		0.16	0.07	0.56	0.14	0.74	0.17
Hourly wage, Post (\$)	14.94	17.49	15.24	22.24	15.57	24.67	16.31
$\frac{h^j(\alpha)-h^d}{h^d}$		0.17	0.02	0.49	0.04	0.65	0.09

a. Reference group. Note: Row 1 and 3 present calibrated pre- and post-reform hourly wage rate for individuals with different ability and education level. Row 2 and 4 show the corresponding hourly wage premium comparing to the high school dropouts.

for part of the premium due to experience in the labor market.

For high ability pre-reform cohorts, high school graduation increases the wage rate by 16% and the premium for low-ability type is only roughly half of it. Furthermore, the wage rate increase is as high as 74% for high ability individuals once graduating from university, and their college certificate wage premium is slightly lower but still considerable 56%. In contrast, the benefit for low ability students to graduate from postsecondary education is four to five times smaller than the high ability counterparts. After high school duration was reduced, high-ability high school graduates' wage premium did not change much. However, those for low type ones decreased significantly. The wage rate decreased by around ten percentage points for college and university graduates, regardless of ability level. These wage penalties imply negative impacts on students' human capital accumulation as high school duration is reduced by one year.

Table 7 presents the learning parameters, with each row representing different signal values and the columns indicating school and ability type. For example, the probability of individuals with high ability receiving "Good" signal in high school is 0.65. By assumption, failing a year in high school signals low ability and doing great in postsecondary education also indicates high type. The likelihood of getting a "Fail" is a lot higher for low ability students in postsecondary education, thus failing the exam could lower their expectation regarding the innate ability. Overall the signals in postsecondary education are more informative than those in high school. Comparatively, high school is easy and more informative regarding the lower tail while postsecondary education provides information for both ability types. The main difference between college abd university programs is the likelihood for high-ability students to achieve excellence.

The last set of parameters we calibrated are the proportion of high ability types in the population and

Table 7: Calibrated learning parameter values

	High school		College		University	
Ability α	high	low	high	low	high	low
Good	0.65	0.28	0.52	0^a	0.36	0^a
Pass	0.35	0.68	0.47	0.91	0.62	0.90
Fail	0^a	0.04	0.01	0.09	0.02	0.10

a. By assumption. Note: Each number stands for the probability of receiving a certain signal values (row) for students with different ability type in different school (column). For example, the probability of individuals with high ability receiving a "Good" signal in high school is 0.65 (First row, First column).

how noisy their initial beliefs are. We find that the proportion of individuals with high ability accounts for roughly 46% and that the standard deviation of their initial belief is 0.34.

Table 8: Average wage and educational distribution

		Old		New	
		Model	Target	Model	Target
Wage ($ hourly)	HS grads	15.64	16.14	15.50	15.52
	Psec dropouts	15.77	17.07	15.67	16.38
	Coll grads	18.96	18.48	18.38	18.58
	Univ grads	24.19	24.66	23.19	23.45
Edu (%)	≤10 yrs	2.6	2.6	2.8	3.0
	11-13 yrs	2.4	3.0	4.2	3.2
	HS grads	15.5	14.6	16.5	14.4
	Psec dropouts	-	-	8.7	8.9
	Coll dropouts	5.5	6.0	-	-
	Coll grads	38.9	38.5	34.2	35.4
	Univ dropouts	3.5	3.1	-	-
	Univ grads	31.6	32.0	33.7	35.1

Table 8 illustrates the educational distribution and wage distribution predicted by the model and compares them with the data counterpart. The model does a good job fitting most of the moments.

Given the calibrated parameters, let us take a closer look at how high school duration affects students' educational decisions. Figure 6 shows the predicted educational distribution in both old and new high school system with the ability composition. There are more post-reform cohorts choosing to enroll in university rather than college. Moreover, those who obtain university education due to the policy change are mainly of low ability type, lowering the average ability level for all the educational groups. The boost in postsecondary enrollment cannot be explained by the wage rate drop, making postsecondary education less attractive in expectation. Thus the "orientation" channel plays a significant role in explaining the pattern as shown in Figure 7.

Figure 6: Predicted education attainment distribution

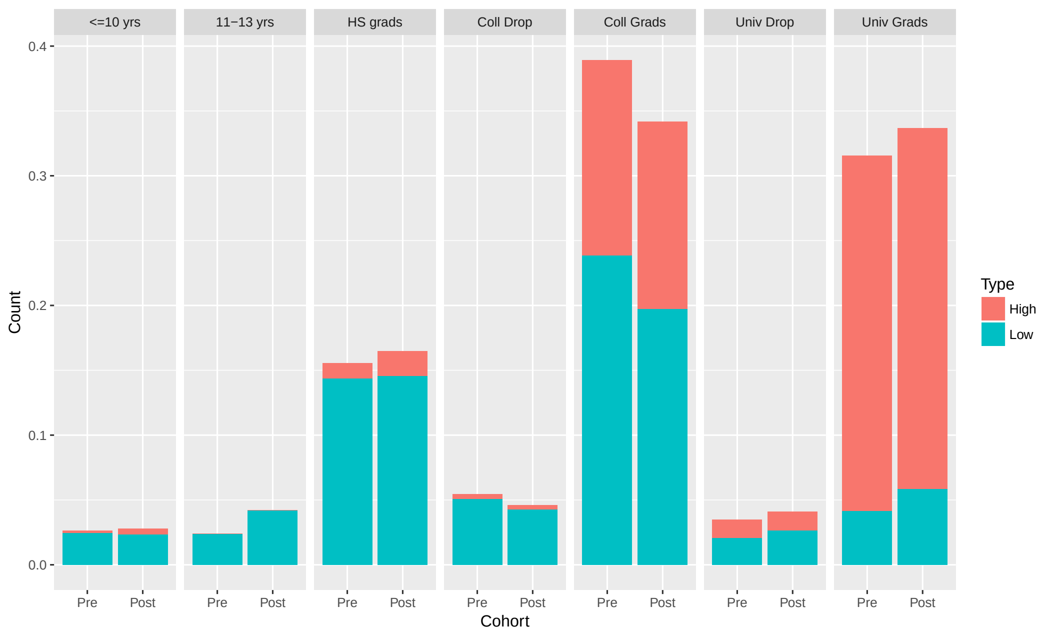
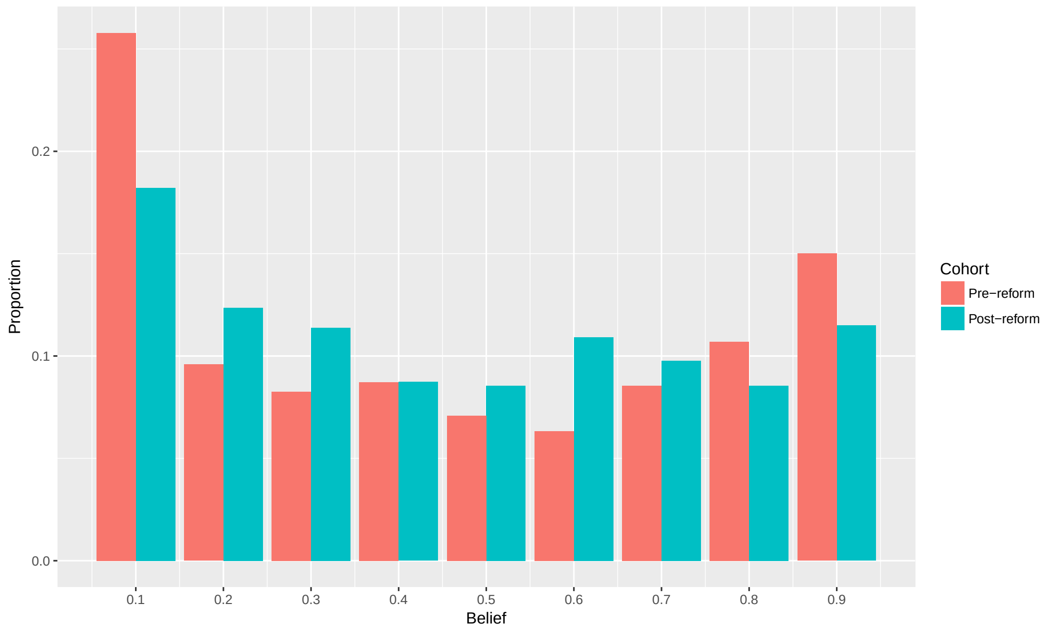


Figure 7 and 8 show the histogram of students' belief when making postsecondary education enrollment decisions with the red bars representing old system and green bars indicating the new regime. Pre-reform cohorts are more informed than their new counterparts in the population. This "ambiguity" in students' belief about their true ability, together with a potential high wage premium of postsecondary degrees, encourages more post-reform students to try postsecondary programs, only to find out it is optimal to drop out later on.

6.4 Decomposition of the policy effect

In order to separate the performance effect from the orientation mechanism, we conduct the following counterfactual simulation. By changing the high school length while holding wages fixed, we completely shut down the performance effect channel. Thus, any resulting changes in educational distribution and wages can be attributed solely to the orientation effect.

Figure 7: Histogram of belief after high school graduation



As we can see in Figure 9, fixing wage rates, shortened high school duration induces a decrease in how informed students are upon high school graduation. As a result, many more high school graduates would enroll in college instead of university or instead of entering the labor market immediately. This is because college is especially appealing to students who have intermediate beliefs. On the one hand, university education is too risky, given its length and difficulty. On the other hand, the big wage premium associated with college certificates makes entering the labor market as high school graduates less attractive.

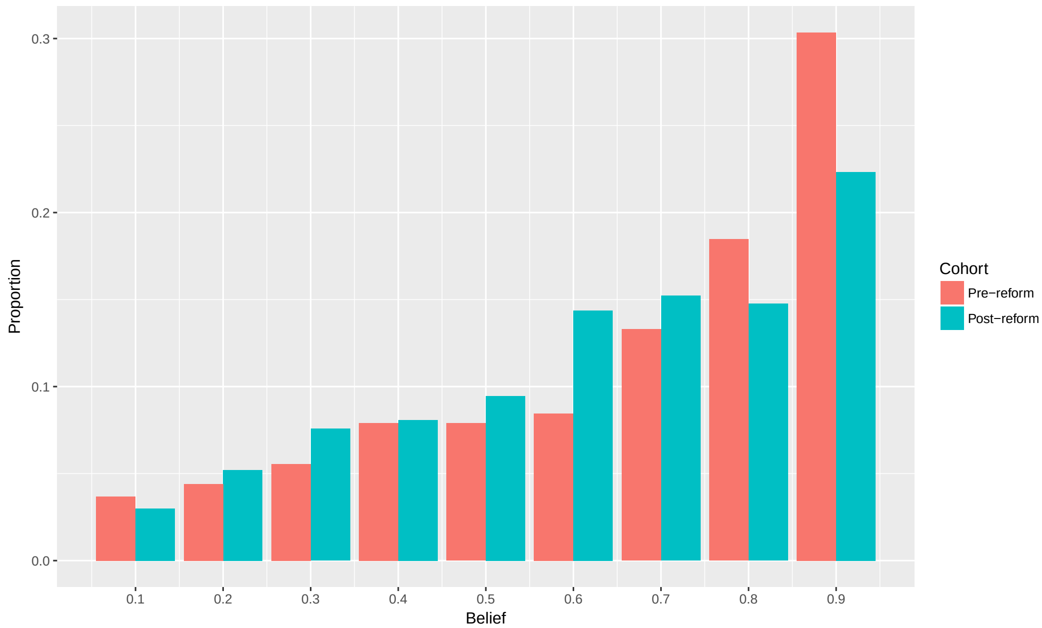
Table 9: Decomposition of Change in Educational Distribution Due to Policy

(P.P.)	Δ Total	Δ Due to Orientation Effect	Δ Due to Performance Effect
≤10 yrs	0.2	−2.6	2.8
11-13 yrs	1.8	1.9	−0.1
HS grads	1.0	−12.4	13.4
Coll Drop	−0.9	1.9	−2.8
Coll Grads	−4.8	14.2	−18.9
Univ Drop	0.6	−0.1	0.7
Univ Grads	2.1	−2.8	4.9

Table 9 shows the decomposition of the percentage point changes in distribution between the pre- and post-reform cohorts. Indeed, the orientation effect contributes to a decrease in the proportion of high school and university graduates (12.4 percentage points (p.p.) and 2.8 p.p., respectively) and an increase in college graduates (14.2 p.p.). This implies that the college wage rate must decrease enough to explain the eventual

Figure 8: Histogram of belief after high school graduation by ability type

(a) High ability



(b) Low ability

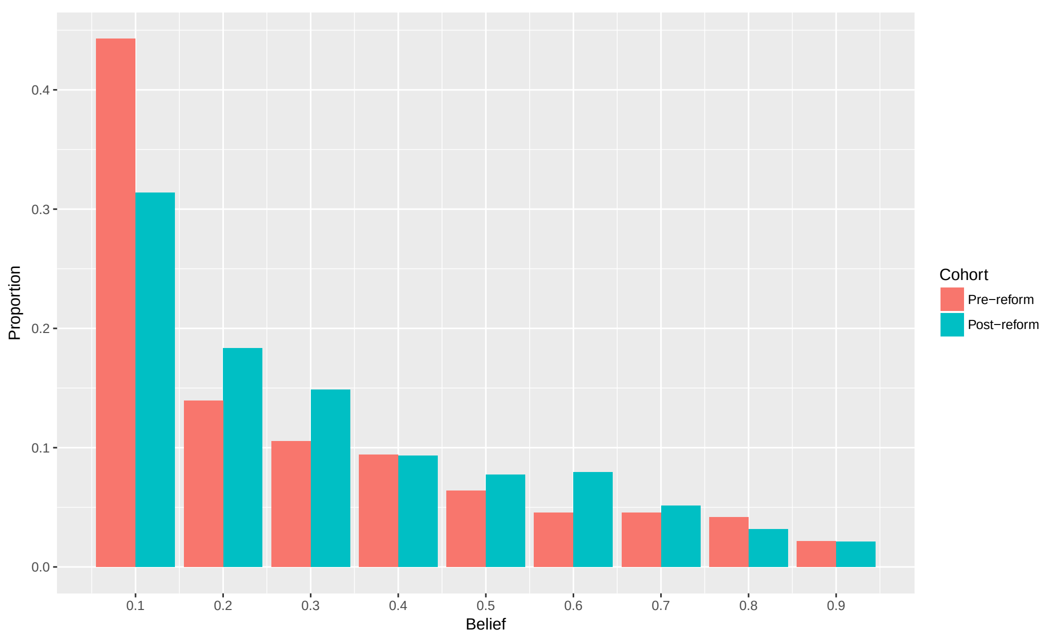
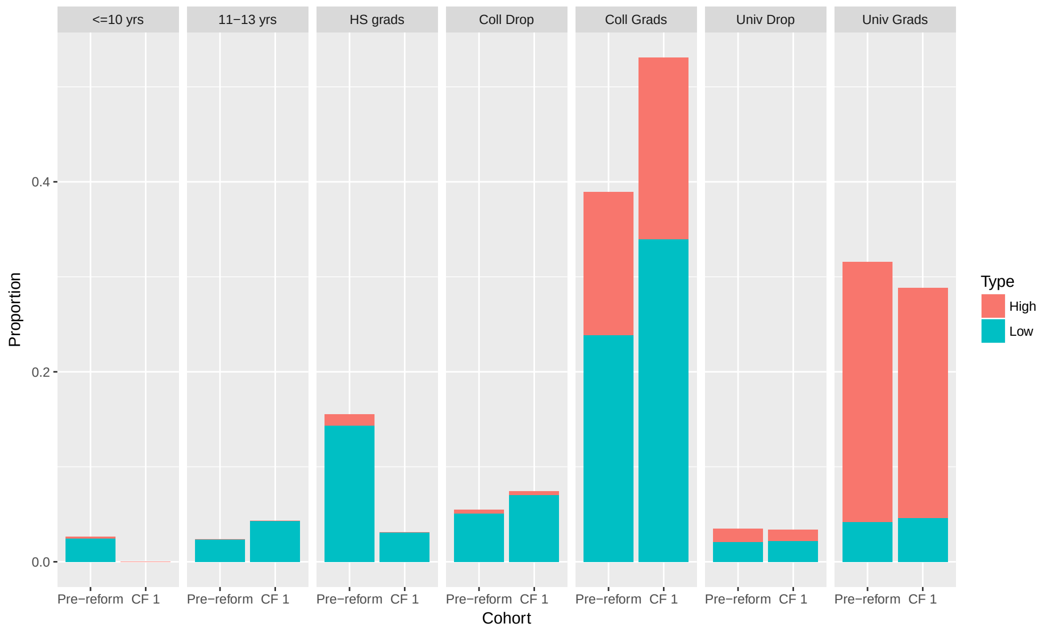


Figure 9: Predicted educational attainment distribution: Counterfactual I



Note: CF 1 stands for counter factual case 1, where old wage rate and new high school structure is used

drop in the fraction of college graduates (4.8 p.p.).

In terms of the wages, Table 10 breaks down the differences in the average hourly wage rate between the pre- and post-reform cohorts into the orientation and the performance component. Not surprisingly, the orientation effect negatively impacts employee earnings. As high school duration is reduced shortened, students are less informed about their ability level, thus making worse educational decisions. This in turn leads to more misallocation of students into educational levels, which adversely affects average wages. However, the performance effect is the dominant factor that explains the average wage decrease between the pre- and the post-reform cohort, especially among postsecondary graduates.

Table 10: Decomposition of change in average wages due to policy

Hourly Wages ($)	Δ Total	Δ Due to Learning	Δ Due to Wage Rate
HS Grad	−0.14	−0.11	−0.03
Coll/Univ Drop	−0.10	−0.06	−0.05
Coll Grad	−0.58	−0.15	−0.43
Univ Grad	−1.00	−0.19	−0.81

Let us now turn to the wage premia. The orientation effect accounts for the decrease in wage premia by between 0.4 p.p. and 1.3 p.p. Indeed, the change in the wage rates also helps explain the decrease in the wage premia from 3.1 p.p. for high school graduates to 9.9 p.p. for university graduates. In conclusion, the orientation effect accounts for 11% to 20% of the total decrease in wage premia and 19% to 78% of the decrease in wage rates. The rest of the decreases can be explained directly by wage rate decreases due to lower human capital accumulation.

Table 11: Decomposition of change in wage premia due to policy

Wage premia (P.P)	Δ Total	Δ Due to Learning	Δ Due to Wage Rate
HS Grad	-3.9	-0.8	-3.1
Coll/Univ Drop	-3.6	-0.4	-3.2
Coll Grad	-7.4	-1.0	-6.4
Univ Grad	-11.2	-1.3	-9.9

7 Policy Analysis

In this section, we analyze a hypothetical environment in which there is an upward shift in grades distribution. We show that the orientation mechanism is essential in curriculum design. The motivation for analyzing this scenario is threefold. First of all, the previous empirical results show that the 1999 Ontarian policy change adversely affected course failure rates and cognitive ability accumulation, especially for low ability students. We also presented previously that the overall high school graduation rate was negatively affected as well. These facts could be seen as a concern and one way to directly address it would be to shift grade upward by reducing school difficulty. This could be achieved either by lightening the curriculum, or by reducing course completion requirements. Secondly, high school graduation rates in Ontario have increased by more than 17 percentage points between 2004 and 2015 (Queen's Printer for Ontario, 2016). It is unclear whether this follows a secular trend, or is influenced by government programs to support struggling students, or is achieved through a formal or informal reduction in the graduation requirements.

Nevertheless, the net effect is that fewer and fewer students are failing high school. Thirdly, grade inflation has been documented in many education systems, including in Ontario (Finefter-Rosenbluh and Levinson, 2015), in particular as a response to increasing university admission standards. In any case, the net result of all these factors is that high school grades are higher than ever. As we will show, although higher grades are associated with higher high school graduation rates, they also have the unexpected perverse effect of leading to low quality post-secondary choices and inefficiencies in the labor market.

We run a counterfactual exercise in which the distribution of high school grades is shifted upwards and the type-specific wage rates remains unchanged. This could be the effect of a policy aiming to lighten the high school curriculum, reduce the requirements, or lower the passing grade for high school courses to make

Table 12: Calibrated learning parameter values

	High school		College		University	
Ability α	high	low	high	low	high	low
Good	$0.65 + 0.04$	$0.28 + 0.04$	0.52	0^a	0.36	0^a
Pass	$0.35 - 0.04$	0.68	0.47	0.91	0.62	0.9
Fail	0^a	$0.044 - 0.04$	0.01	0.09	0.02	0.10

Note: Each number stands for the probability of receiving a certain signal values (row) for students with different ability type in different school (column). The numbers in black are the calibrated results, and the numbers in red show the change used for policy analysis.

up for the increased learning intensity brought on by a reduction in the number of years in high school. In our model, this kind of policy translates into a decrease in the probability of receiving a low (failing or pass) grade and an increase in the probability of receiving a good grade. Specifically, for high types, we increase the probability of receiving a good grade by 0.04 and decrease the probability of receiving a medium grade by the same amount. For low types, this shift is between good grades and failing grades. These changes are documented in Table 12. They imply that low ability students are more heavily impacted by this grade shift than high ability types. We make this choice to reflect the fact that low-ability students are typically more sensitive to curriculum difficulty changes than high-ability students.

Table 13: Educational attainment distribution after lowering requirements

Attainment (%)	Baseline	Policy	Difference
≤10 yrs	2.8	2.6	0.0
11-13 yrs	4.2	0.0	1.4
HS grads	16.5	16.6	1.8
Coll dropouts	4.6	5.6	0.0
Coll grads	34.2	40.2	0.0
Univ dropouts	4.1	3.4	0.0
Univ grads	33.7	31.5	0.0

The grade inflation influences students' decisions in two ways. First of all, as failure rates decrease, grades become less informative. In particular, failures serve as entirely informative signals of low ability. Likewise, high grades contain less information. As a result, students' beliefs of their types are "weaker": both types can achieve better results more easily. Secondly, decreasing failure rates means that students are more likely to graduate from high school and do not need to fear the risk of dropping out as much as before.

This makes staying in high school more attractive, especially for those who have relatively pessimistic beliefs. In our counterfactual simulation results presented in Table 13, we see that grade inflation leads to a boost in high school graduation. More surprisingly, inflating grades leads to decreases in the university graduation rate. This means that although more students graduate from high school, they are less informed about their abilities. Among them, fewer are confident enough to enroll in university programs. Moreover, as the information content of high school grades decreases, there is more misallocation of students into educational levels, decreasing university students' average ability slightly. In summary, we want to draw attention to the signaling role of school. Though school grade inflation can boost the graduation rate, it leads to a decrease in the quality of students' future educational decisions due to the deterioration in the precision of knowing their ability.

8 Conclusion

This article sheds light on the effects of school length on students' educational attainments and labor market outcomes. We study two mechanisms through which school duration can impact students: by affecting human capital accumulation directly (performance effect) and influencing how students learn their true ability (orientation effect).

We evaluate the 1999 Ontarian policy change to tease out the effects of school length on multiple outcomes. Using LFS and NLSCY, we compare the difference between pre- and post-reform cohorts in Ontario to those in the rest of Canada to identify the policy effect. We conclude that shortened high school length led to a 1.4% reduction in high school graduation rate, caused a 3% increase in immediate university enrollment and a 5% increase in college with one year delay, without increasing the proportion of the population holding a post-secondary degree when measured at ages $28-36$. This hints at an increase in the post-secondary dropout rate. Meanwhile, there is a 5% to 10% wage penalty within educational levels for high school graduates under the new system.

Since the orientation and performance mechanisms' effects cannot be separately identified in a reduced-form framework, we construct and estimate a dynamic discrete choice model. Given the estimation result, we conduct a counterfactual analysis to measure the relative importance of each channel. We find that the orientation effect accounts for between 11% and 20% of the decrease in wage premia across educational attainment categories, while the performance effect explains the remainder. In terms of wage rates, the orientation effect accounts for 19% to 78% across the different educational categories, with its importance of decreasing with educational level. In monetary terms, the orientation effect predicts a decrease of \$0.06 to \$0.19 in hourly wage rates.

Furthermore, we find that the orientation effect causes a distributional shift from university, especially from high school to college. This shift is offset by an opposite and larger effect caused by a decrease in the wage rates. Overall, the evidence suggests that the policy's negative effects on wages persist and their magnitude is amplified as the educational level increases. This is intuitive, since post-secondary bound students are most affected by the policy reform. After all, high school duration for post-secondary bound students decreased by a full year after the reform.

Several avenues for further research remain. In particular, we have assumed that ability types are fully revealed to employers upon entering the labor market. It would be interesting to explore a setup where the employers cannot perfectly observe ability types, instead paying wages equal to expected productivity given their educational level. In that way, post-secondary wages may decrease due to signaling if more low ability students pursue college or university education as high school is shortened.

As a closing remark, we would like to note that although this paper is mainly focused on analyzing the reduction in high school length in Ontario, the framework we use can be applied to any policy aiming to change school duration. As a main takeaway, we emphasize the fact that policymakers should increasingly think of school not simply as a means to gain skills or knowledge, but also as a means to enable students to understand their talents, interest and abilities and efficiently match them to professions where they can succeed and be productive.

References

Brady, P. and P. Allingham (2010). Pathways to university: the victory lap phenomenon in ontario. *Canadian Journal of Educational Administration and Policy 113*, 1–25.

Büttner, B. and S. L. Thomsen (2015). Are We Spending Too Many Years in School? Causal Evidence of the Impact of Shortening Secondary School Duration. *German Economic Review 16*(1), 65–86.

Casas, F. R. and D. E. Meaghan (1996). A study of repeated courses among secondary students in ontario. *The Journal of Educational Research 90*(2), 116–127.

Dahmann, S. and S. Anger (2013). The impact of education on personality: Evidence from a german high school reform.

Finefter-Rosenbluh, I. and M. L. Levinson (2015). What is wrong with grade inflation (if anything)?

Hofmann, S., A. Mühlenweg, et al. (2017). Learning intensity effects in students' mental and physical health-evidence from a large scale natural experiment in germany. Technical report, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät.

Huebener, M. and J. Marcus (2017). Compressing instruction time into fewer years of schooling and the impact on student performance. *Economics of Education Review 58*, 1–14.

Krashinsky, H. (2009). How would one extra year of high school affect wages? evidence from an unique policy change," working paper no. 20, canadian labour market and skills researcher network.

Krashinsky, H. (2014). How would one extra year of high school affect academic performance in university? evidence from an educational policy change. *The Canadian Journal of Economics/Revue canadienne d'Economique 47*(1), 70–97.

Meyer, T. and S. L. Thomsen (2016). How important is secondary school duration for post-secondary education decisions? evidence from a natural experiment. *Journal of Human Capital 10*(1), 67–108.

Meyer, T. and S. L. Thomsen (2018). The role of high-school duration for university students' motivation, abilities and achievements. *Education Economics 26*(1), 24–45.

Meyer, T., S. L. Thomsen, and H. Schneider (2015). New evidence on the effects of the shortened school duration in the german states: An evaluation of post-secondary education decisions. *German Economic Review*.

Morin, L.-P. (2013). Estimating the benefit of high school for university-bound students: evidence of subject-specific human capital accumulation. *Canadian Journal of Economics/Revue canadienne d'économique 46*(2), 441–468.

Queen's Printer for Ontario (2016). Getting results. `http://www.edu.gov.on.ca/eng/gettingResultsGrad.html`. Accessed: 2018-09-06.

Savoca, E. (2000). Measurement errors in binary regressors: an application to measuring the effects of specific psychiatric diseases on earnings. *Health Services and Outcomes Research Methodology 1*(2), 149–164.

Trachter, N. (2015). Stepping stone and option value in a model of postsecondary education. *Quantitative Economics 6*(1), 223–256.

Vaillancourt, F. (1995). The private and total returns to education in canada, 1985. *Canadian Journal of Economics*, 532–554.

9 Appendix

9.1 Estimating misreport probability in educational attainment variable in LFS

One big challenge to model identification is the non-random measurement error in the reported highest educational levels in LFS. The variable that records individuals' highest level of education has the following categories

- $0-8$ years of education / $9-10$ years of education / $11-13$ year of education without graduating from high school
- High school graduates
- Some post-secondary education
- Trades certificate or diploma from a vocational or apprenticeship training / Non-university certificate or diploma from a community college, CEGEP, school of nursing, etc.
- University certificate below bachelors degree / Bachelors degree / University degree or certificate above bachelors degree

Thus by definition, the category "Some post-secondary education" should include both people who are still attending college or university and those who dropped out without graduation. However, we notice that there is a considerable proportion of people who are attending college or university in the reference month yet reporting their highest level of education being "high school graduates" or even "high school dropouts"[24]. Therefore we suspect that there is non-negligible, non-random measurement error contained in this variable in LFS data.

To address this issue, we estimate the distribution of measurement error by comparing the educational distribution in LFS with that in SLID, and then use the estimates to correct the coefficients in relevant regression.

[24]There exist individuals in the sample who attended postsecondary education without finishing high school. However, this proportion is very small thus excluded in the analysis

Comparing the educational distribution of cohorts born between 1965-1976 measured at age $30-35$ yrs old using LFS and SLID data, LFS predicts significantly smaller proportion of individuals with "some post-secondary education", with "college certificate", and at the same time considerably more high school graduates as well as university degree holders. The discrepancy could be due to the confusion between the "highest education level", which is asked in the LFS, and the "highest degree obtained" as interpreted by the respondents. Thus we assume the misreport only happens among post-secondary institution dropouts, either with high school certificates or college certificates (Those who drop out of university programs after finishing certain college certificate). Furthermore, we assume that the possible misreport levels depend on the true educational attainment. People with "some secondary educational" could misreport their highest educational level being "high school grads", "college certificates" or "university degree". Those who have college certificates but later on drop out from university program could claim their educational attainment being "university degree" by mistake. Notice that the category "University degrees" here includes all the degrees issued in university programs, including "degrees below Bachelor's", "Bachelor's degree" and those above Bachelor's.

The following equation system fully describes the relationship between educational attainment observed in LFS and the corresponding true levels,

$$\begin{bmatrix} Pr_L(\text{HS grad}) \\ Pr_L(\text{some psec}) \\ Pr_L(\text{coll cert}) \\ Pr_L(\text{univ cert}) \end{bmatrix} = \begin{bmatrix} 1 & p_1 & 0 & 0 \\ 0 & 1-p_1-p_2-p_3 & 0 & 0 \\ 0 & p_2 & 1-p_4 & 0 \\ 0 & p_3 & p_4 & 1 \end{bmatrix} \begin{bmatrix} Pr_T(\text{HS grad}) \\ Pr_T(\text{some psec}) \\ Pr_T(\text{coll cert}) \\ Pr_T(\text{univ cert}) \end{bmatrix}$$

where

$$p_1 = Pr(z = \text{HS grad}|\text{some psec})$$

$$p_2 = Pr(z = \text{coll cert}|\text{some psec})$$

$$p_3 = Pr(z = \text{univ cert}|\text{some psec})$$

$$p_4 = Pr(z = \text{univ cert}|\text{coll cert})$$

The left hand side represents the proportion observable in LFS with measurement error, and the proportion with subscript T on the right hand side stands for the true value. Since a lot more detailed information was documented in SLID, it's reasonable to assume that SLID data is of higher quality, thus reflecting the true educational distribution. Then the next step is to estimate probabilities p_1 to p_4 by solving the linear equations system.

The solution is as follows

$$p_1 = Pr(z = \text{HS grad}|\text{some psec}) = 0.35$$

$$p_2 = Pr(z = \text{coll cert}|\text{some psec}) = 0.00$$

$$p_3 = Pr(z = \text{univ cert}|\text{some psec}) = 0.053$$

$$p_4 = Pr(z = \text{univ cert}|\text{coll cert}) = 0.09$$

suggesting that around 40% of post-secondary dropouts misreport in LFS, among which 35% tend to report being high school graduates and 5% report having obtained university degree. On the other hand, for people who have already obtained college degree and also some university education, around 9% of them would over report having university certificate.

Further assuming that the misreport probability distribution do not change over age nor across cohorts, we could use the matrix estimated above to correct the educational level distribution for each province $\times$ birth year combination to estimate the policy effect in the long term, and also to correct estimated policy effect on return to education.

9.2 Correcting measurement error in regression result in LFS

This part is mainly taken from Savoca (2000). Consider the population regression model

$$Y = X\beta_0 + \epsilon \tag{14}$$

where Y is the outcome variable of interest and X is the independent variable vector, including the categorical regressor. In the data, however, we could only observe the regressor W with measurement error U modeled in the following way

$$W = X + U \tag{15}$$

Where both X and W are categorical variables. A non-zero correlation between the error U and X is present due to the binary feature, and could be modeled with the misreport probability estimated in the previous section. Then the regression based on observables can be rewritten as

$$Y = W\beta_0 + \epsilon - U\beta_0 \tag{16}$$

Then the OLS estimator with observables is,

$$\hat{\beta}_{OLS} = (I - (W'W)^{-1}W'U)\beta_0 + (W'W)^{-1}W'\epsilon \tag{17}$$

Assuming that ϵ is independent of X and U, then

$$plim(\hat{\beta}_{OLS}) = (I - \Omega)\beta_0 \tag{18}$$

where $\Omega = \Sigma_{ww}^{-1} \cdot \Sigma_{wu}$ and Σ_{AB} denotes the covariance matrix between variables A and B.

Though X is not available from the data, we calculate Σ_{xu} and then construct Σ_{wu} from the estimated misreport probability. For Σ_{ww}, we use sample moments to estimate it. Thus the corrected coefficient is obtained from the OLS estimates in the following way

$$\hat{\beta}_c = (I - \hat{\Omega})^{-1}\hat{\beta_{OLS}} \tag{19}$$

Where $\hat{\Omega} = \widehat{\Sigma_{ww}}^{-1} \cdot \widehat{\Sigma_{wu}}$

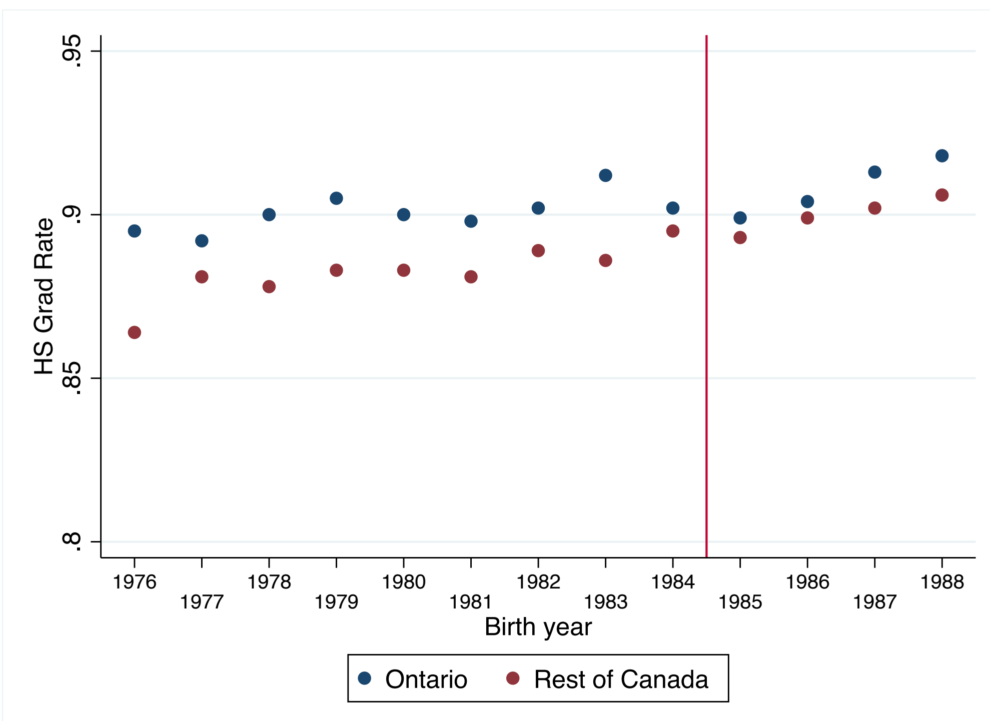
9.3 Robustness check

In this section, we perform several robustness checks that address possible confounding factors in the identification strategy used in the empirical section.

9.3.1 Common trend assumption

First of all, we investigate the validity of using a difference-in-difference estimation strategy. In particular, I compare the evolution of various outcomes, high school graduation rate, their post-secondary education attendance and their educational attainment, of the treated group (i.e. individuals from Ontario) with that of the control group (i.e. individuals from the rest of Canada) over cohorts. Figure 10, Figure 11 and Figure 12 provide evidence that supports the "common trend" assumption.

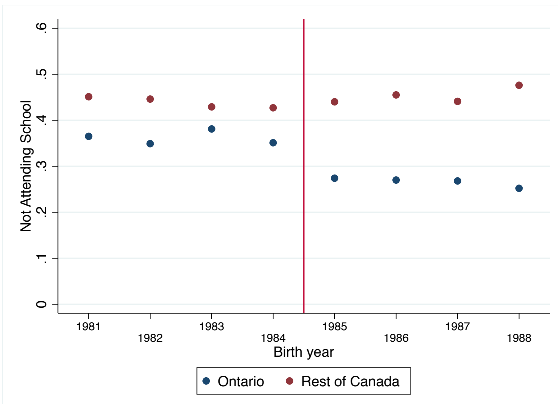
Figure 10: Trends in high school graduation rate



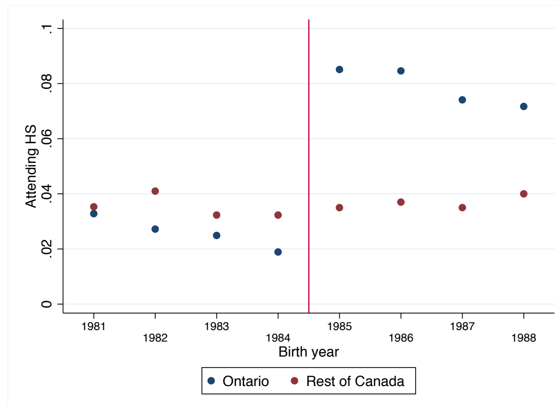
Source : LFS 2000-2017. Sample includes all the individuals born between 1976 and 1988, with age $22-24$ and who were not attending high school in ref month. Quebec is excluded from the sample due to its unique educational structure.

Figure 11: Trends in post-secondary education choice

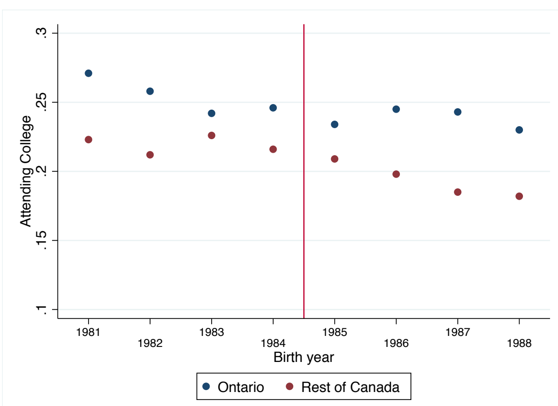
(a) Not in school



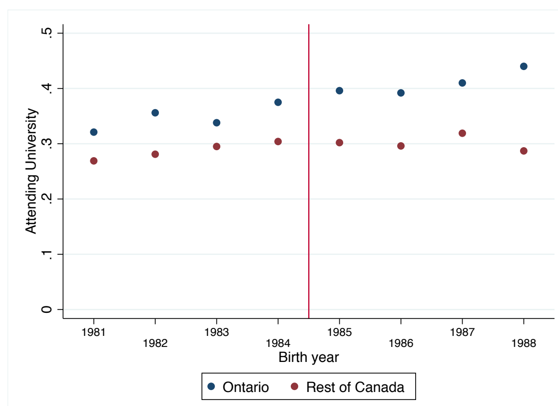
(b) Attending high school



(c) Attending college



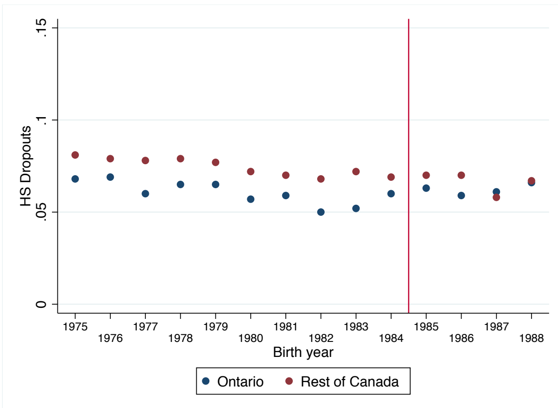
(d) Attending university



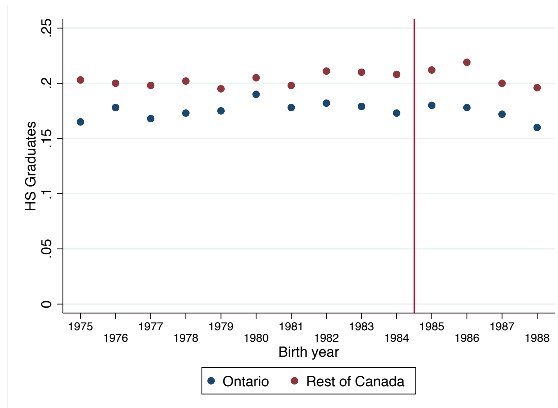
Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure.

Figure 12: Trends in educational attainment

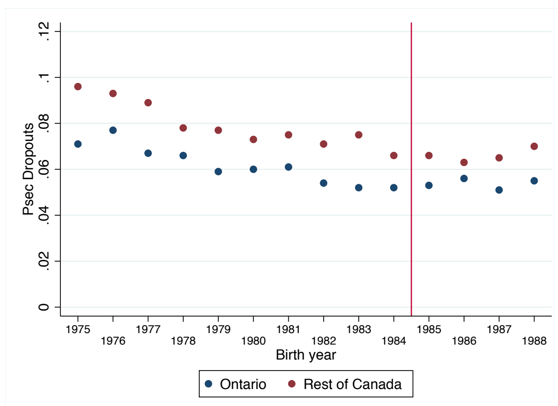
(a) HS dropouts



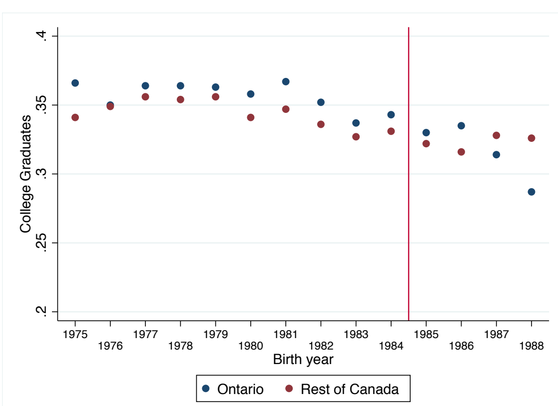
(b) HS grads



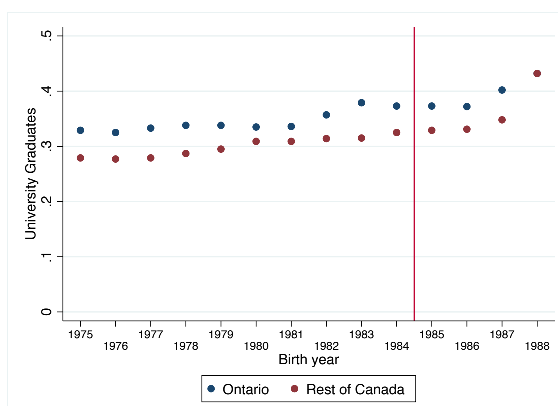
(c) Some psec



(d) College graduate



(e) University graduate



Source: LFS 2000-2017. Sample includes all the cohorts born between 1975 and 1988 at age 28 – 36. Quebec is excluded from the sample due to its unique educational structure.

A more rigorous analysis is done by performing the following regression

$$Y_{ijk} = Cons + Ontario_{ijk} \cdot Cohort_j \theta_j + \alpha_j + \beta_k + X'_{ijk}\gamma + \varepsilon_{ijk} \tag{20}$$

where the Y_{ijk} is the outcome variable, $Ontario_{ijk}$ is a dummy variable that equals to 1 if individual i is from Ontario and 0 otherwise, $Cohort_j$ is the cohort dummy. Other controls include cohort fixed effect α_j, province fixed effect β_k and other individual-specific characters X_{ijk}. When the outcome variable is log hourly wage rate, education level is further added in the regression. θ_j indicates the average wage difference between Ontario and the rest of Canada for cohort j. Thus "common assumption" is valid if θ_j is not different among all the pre-reform cohorts. Table14, Table15, Table16 and Table 17 present the estimation results and the common trend assumption holds in most of the cases. Even in scenarios where the common trend fails to hold (e.g. College and University attendance at high school graduation age +1 in Table 16), the main regression results were biased towards zero.

9.3.2 Removing the double cohorts

One might worry the observed change in students' educational decision and labor market outcome is mainly driven by the double cohorts since the increased competition might have distorted their educational choices temporarily. In that case, our difference-in-difference regression result would capture a temporary "adjustment process", over-estimating the true policy effect. Thus we perform robustness check by estimating the policy effect while excluding the double cohorts from the sample. As shown by Column 1 and 3 in Table 18, Table 19, Table 20 and Figure 13, the policy effect estimation remains of similar magnitude after removing the double cohorts.

9.3.3 Adding non-parametric trend

Another confounding factor that challenges the difference-in-difference identification is the provincial-cohort specific idiosyncratic shock. In order to rule out this possibility, we further add provincial unemployment rate to capture the provincial specific macro economic

Table 14: Common trend: high school graduation rate and log hourly wage rate

	(1) HS grad	(2) $\log(w)$
$I_{1976} \cdot Ontario$	0.011	
	(0.021)	
$I_{1977} \cdot Ontario$	-0.012	
	(0.014)	
$I_{1978} \cdot Ontario$	-0.008	
	(0.008)	
$I_{1979} \cdot Ontario$	-0.012	
	(0.009)	
$I_{1980} \cdot Ontario$	0.011	
	(0.010)	
$I_{1981} \cdot Ontario$	-0.001	-0.007
	(0.008)	(0.009)
$I_{1982} \cdot Ontario$	-0.000	-0.022
	(0.009)	(0.015)
$I_{1984} \cdot Ontario$	-0.023***	-0.037**
	(0.008)	(0.015)
$I_{1985} \cdot Ontario$	-0.023**	-0.060***
	(0.010)	(0.013)
$I_{1986} \cdot Ontario$	-0.026***	-0.054***
	(0.008)	(0.011)
$I_{1987} \cdot Ontario$	-0.013	-0.071***
	(0.011)	(0.011)
$I_{1988} \cdot Ontario$	-0.012*	-0.061***
	(0.007)	(0.018)
N	0.069	0.051
R^2	20800	20300

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable in first column is high school graduation indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1976 and 1988, with age $22-24$ and who were not attending high school in ref month. Dependent variable for the second column is log hourly wage rate. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age $28-32$, who are not attending school in ref month. Quebec is excluded from the sample due to its unique educational structure. Other control variables include age dummy, gender, provincial dummy, year dummy and month dummy. Cohorts born in 1983 is the reference level. Standard errors are in the parentheses. Variance is clustered within province $\times$ year.

Table 15: Common trend: psec activities
HS grad age

	(1)	(2)	(3)	(4)
	Not in School	School Attendance		
		High school	College	University
$I_{1981} \cdot Ontario$	0.048	-0.067***	0.007	0.009
	(0.040)	(0.014)	(0.032)	(0.017)
$I_{1982} \cdot Ontario$	0.056**	-0.031***	-0.014	-0.012
	(0.022)	(0.008)	(0.020)	(0.017)
$I_{1984} \cdot Ontario$	-0.037*	0.029***	-0.014	0.016
	(0.021)	(0.008)	(0.017)	(0.012)
$I_{1985} \cdot Ontario$	-0.095***	0.095***	-0.033	0.030
	(0.021)	(0.032)	(0.025)	(0.019)
$I_{1986} \cdot Ontario$	-0.134***	0.184***	-0.067**	0.018
	(0.032)	(0.052)	(0.030)	(0.028)
$I_{1987} \cdot Ontario$	-0.160***	0.255***	-0.109***	0.012
	(0.041)	(0.064)	(0.033)	(0.029)
$I_{1988} \cdot Ontario$	-0.145***	0.324***	-0.137***	-0.044
	(0.052)	(0.071)	(0.038)	(0.035)
N	0.088	0.032	0.021	0.126
R^2	73523	73523	73523	73523

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is a school attendance indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988 at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort and province of residence fixed effect. Cohorts born in 1983 is the reference level. Standard errors are between parentheses. Variance is clustered within province × birth year.

Table 16: Common trend: psec activities
HS grad age+1

	(1)	(2)	(3)	(4)
	Not in School	School Attendance		
		High school	College	University
$I_{1981} \cdot Ontario$	-0.042*	0.012***	-0.034	0.056*
	(0.022)	(0.003)	(0.021)	(0.032)
$I_{1982} \cdot Ontario$	-0.033*	0.001	0.027**	0.003
	(0.019)	(0.003)	(0.011)	(0.020)
$I_{1984} \cdot Ontario$	-0.028*	0.005**	0.006	-0.000
	(0.016)	(0.002)	(0.020)	(0.023)
$I_{1985} \cdot Ontario$	-0.084***	0.007***	0.079***	-0.003
	(0.019)	(0.002)	(0.012)	(0.012)
$I_{1986} \cdot Ontario$	-0.070***	0.009***	0.039***	0.018*
	(0.015)	(0.002)	(0.010)	(0.010)
$I_{1987} \cdot Ontario$	-0.097***	0.012***	0.053***	0.022
	(0.015)	(0.002)	(0.011)	(0.013)
$I_{1988} \cdot Ontario$	-0.055***	0.005	0.041***	0.004
	(0.020)	(0.004)	(0.015)	(0.025)
N	0.090	0.002	0.015	0.116
R^2	46410	46410	46410	46410

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is a school attendance indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988 at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort and province of residence fixed effect. Cohorts born in 1983 is the reference level. Standard errors are between parentheses. Variance is clustered within province × birth year.

Table 17: Common trend: educational attainment at age $28-36$

	(1) HS dropouts	(2) HS grads	(3) Some psec	(4) College	(5) University
$I_{1976} \cdot Ontario$	0.003	0.002	0.015	-0.011	-0.009
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1977} \cdot Ontario$	0.005	-0.011	0.010	-0.004	-0.001
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1978} \cdot Ontario$	-0.002	-0.010	0.013	0.002	-0.002
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1979} \cdot Ontario$	0.006	-0.000	0.004	0.001	-0.011
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1980} \cdot Ontario$	0.004	0.009	0.009	0.007	-0.029
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1981} \cdot Ontario$	0.008	0.001	0.006	0.011	-0.025
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1982} \cdot Ontario$	0.005	-0.002	0.006	0.012	-0.021
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1984} \cdot Ontario$	0.011	-0.013	0.010	0.006	-0.014
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1985} \cdot Ontario$	0.016	-0.014	0.014	-0.004	-0.012
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1986} \cdot Ontario$	0.012	-0.013	0.016	0.007	-0.023
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1987} \cdot Ontario$	0.019	-0.004	0.005	-0.025	0.005
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1988} \cdot Ontario$	0.022*	-0.013	0.009	-0.046*	0.028
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
N	0.768	0.814	0.846	0.877	0.882
R^2	126	126	126	126	126

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: LFS 2000-2017. Dependent variable is the proportion of corresponding education level of each province × birth year group when measured at age $28-36$. Cohorts born between 1976 and 1988 are included. Quebec is excluded from the sample due to its unique educational structure. Cohorts born in 1983 is the reference level.

Table 18: Robustness check: high school graduation rate and log hourly wage rate

	(1)	(2)	(3)	(4)
	HS grad		$log(w)$	
Policy	-0.015***	-0.012**	-0.052***	-0.038***
	(0.005)	(0.005)	(0.007)	(0.007)
Double cohort	NO		NO	
U		Yes		Yes
N	174000	208000	142000	203000
R^2	0.069	0.069	0.05	0.05

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable for columns 1 and 2 is high school graduation indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1975 and 1988, with age $22 - 24$ and who were not attending high school in ref month. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort, province of residence fixed effect. Dependent variable the last two columns is log hourly wage rate. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age $28 - 32$, who are not attending school in ref month. Quebec is excluded from the sample due to its unique educational structure. Other control variables include age dummy, gender, provincial dummy, year dummy and month dummy. Standard errors are in the parentheses. Double cohort means cohorts born in 1984 and 1985. U stands for provincial unemployment rate. Variance is clustered within province $\times$ year.

Table 19: Robustness check: psec activities without double cohorts

	(1)	(2)	(3)	(4)
	Not in School	School Attendance		
		High school	College	University
HS grad age				
Policy	-0.059***	0.018***	-0.002	0.043***
	(0.017)	(0.006)	(0.016)	(0.015)
Double cohort	No	No	No	No
N	54027	54027	54027	54027
R^2	0.094	0.027	0.02	0.131
HS grad age +1				
Policy	-0.049***	0.004**	0.047***	-0.006
	(0.011)	(0.002)	(0.008)	(0.015)
Double cohort	No	No	No	No
N	34881	34881	34881	34881
R^2	0.094	0.002	0.014	0.121

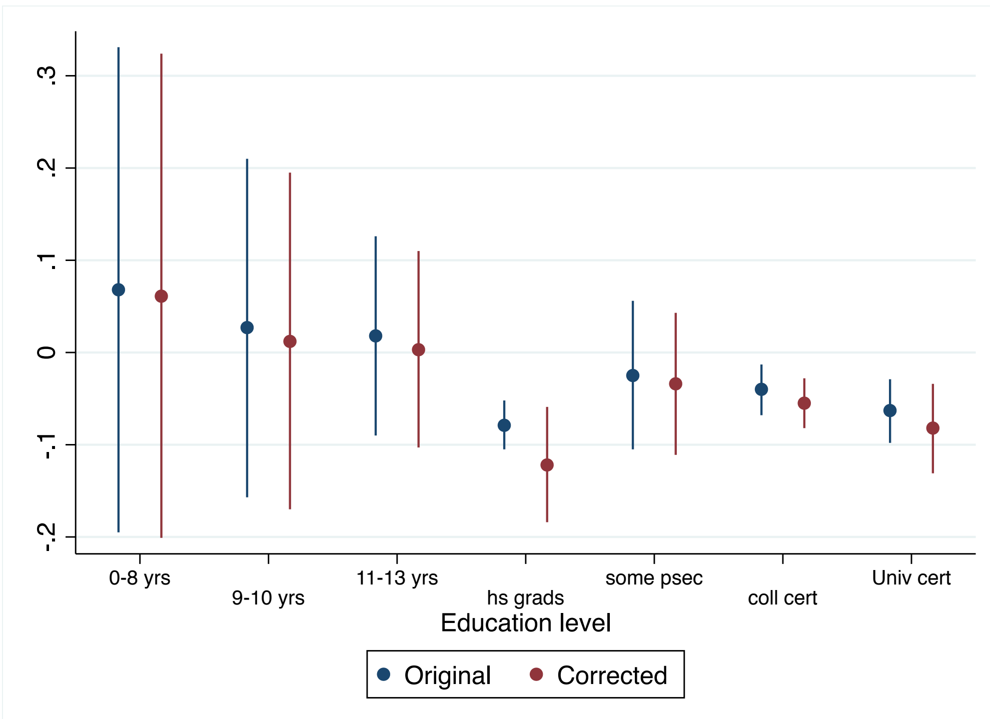
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is a school attendance indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988 except 1984 and 1985, at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort and province of residence fixed effect. Standard errors are between parentheses. Variance is clustered within province × birth year.

Table 20: Robustness check: educational attainment at age 28-36 without double cohorts

	(1) HS dropouts	(2) HS grads	(3) Some psec	(4) College	(5) University
Original					
Policy	0.015**	-0.007	0.003	-0.025*	0.014
	(0.006)	(0.009)	(0.006)	(0.012)	(0.012)
Double cohort	No	No	No	No	No
N	108	108	108	108	108
R^2	0.769	0.792	0.820	0.868	0.877
Corrected					
Policy	0.015**	-0.008	0.004	-0.027*	0.016
	(0.006)	(0.011)	(0.010)	(0.014)	(0.012)
Double cohort	No	No	No	No	No
N	108	108	108	108	108
R^2	0.769	0.733	0.820	0.868	0.876

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: LFS 2000-2017. Dependent variable is the proportion of corresponding education level of each province × birth year group when measured at age $28 - 36$. Cohorts born between 1975 and 1988, except 1984 and 1985, are included. Quebec is excluded from the sample due to its unique educational structure. "Corrected" means the distribution is corrected by the misreport matrix derived previously.

Figure 13: Robustness check: Policy effect on log hourly wage without double cohorts



Source: LFS 2000-2017. Dependent variable is log (real) hourly wage rate. Reference level is "0-8 years of education". Sample includes individuals born after 1980, at age 28-32 yrs old, and who work as employees. Individuals whose hourly wage is below 2$ or who are attending school during ref period are excluded. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, age and year dummies. Standard errors are clustered at province-year level. "Original" shows the estimates using original data in the LFS. "Corrected" shows the estimates after correcting for measurement error.

condition[25]. As can be seen in columns 2 and 4 in Table 18, Table 21 as well as Figure 14, the policy effect estimates remain similar in magnitude as those in the original model specifications. This suggest the policy effects we identified in the paper are not purely driven by idiosyncratic shocks.

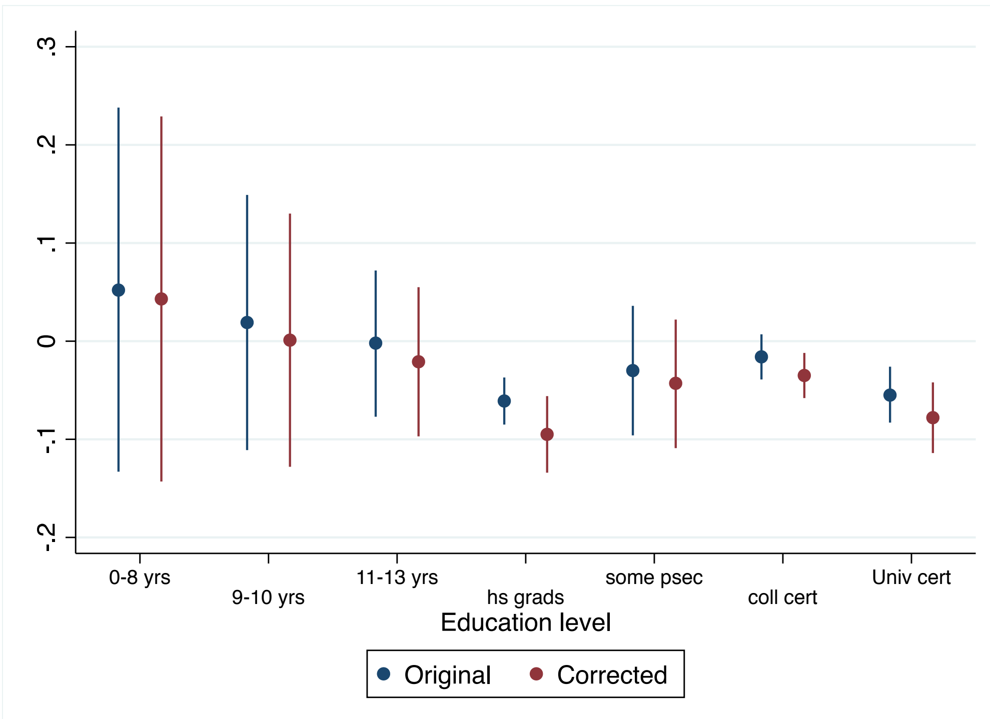
Table 21: Robustness check: psec activities with unemployment rate

	(1)	(2)	(3)	(4)
	Not in School	School Attendance		
		High school	College	University
HS grad age				
Policy	-0.048***	0.014*	0.000	0.034***
	(0.014)	(0.007)	(0.013)	(0.012)
U	Yes	Yes	Yes	Yes
N	73523	73523	73523	73523
R^2	0.087	0.031	0.021	0.125
HS grad age +1				
Policy	-0.055***	0.004**	0.054***	0.000
	(0.009)	(0.002)	(0.008)	(0.012)
U	Yes	Yes	Yes	Yes
N	46410	46410	46410	46410
R^2	0.09	0.002	0.015	0.116

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is a school attendance indicator. Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, parents' educational attainments, year, cohort and province of residence fixed effect, and provincial unemployment rate. Standard errors are between parentheses. Variance is clustered within province $\times$ birth year.

[25] We use provincial unemployment in the year when individuals graduate from high school when analyzing high school graduation rate and their post-secondary education enrollment decision

Figure 14: Robustness check: Policy effect on log hourly wage with unemployment rate



Source: LFS 2000-2017. Dependent variable is log (real) hourly wage rate. Reference level is "0-8 years of education". Sample includes individuals born after 1980, at age 28-32 yrs old, and who work as employees. Individuals whose hourly wage is below 2$ or who are attending school during ref period are excluded. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender, age and year dummies and provincial unemployment rate. Standard errors are clustered at province-year level. "Original" shows the estimates using original data in the LFS. "Corrected" shows the estimates after correcting for measurement error.

9.4 Identification

Here follows a more detailed discussion on structural model identification. Notice that, in our model, the learning parameters are the same for the pre- and post-reform cohorts, whereas the wage rates are particular to each of them. Therefore, data from both cohorts helps us pin down the learning parameters, while data from each cohort identify the respective wage rates.

Loosely speaking, the probability of failure for all types of schools is pinned down by the number of dropouts in each education category. However, students are rational in the model, so an excessive increase in the failure probabilities decreases the likelihood of graduating and can induce students not to enroll or to drop out from a school program, thus reducing both the number of graduates and the number of dropouts. Likewise, if the failure rate is too small, continuing education becomes more appealing as the probability of graduation increases. This would increase the number of enrolling students, as well as the number of graduates. However, the effect on the number of dropouts is ambiguous: more enrollment leads to more dropouts, while a smaller failure probability mechanically decreases the drop rate among enrollees as well. In summary, two opposite effects are at play: the probability of failure mechanically increases the number of dropouts, but at the same time may reduce enrollment and may actually decrease the number of dropouts through this channel.

In a similar manner, as the probabilities of receiving good grades in post-secondary education (which are non-zero only for high types in our model) increase, high ability students learn their true ability type more rapidly and therefore fewer of them mistakenly drop out. This also affects low ability students since receiving a "pass" becomes more associated to being low type. As a result, one would expect the dropping out rate for low ability students to increase as well.

This analysis does not necessarily hold for high school grades because both high and low ability students can receive good grades. In high school, as long as the probabilities of receiving grades are very similar, the informational content of grades is limited and learning is slow. This leads to a large proportion of high school graduates who have ability beliefs close to 0.5. When taking into account the potential big wage premium associated with post-secondary degree and also the difference in the risk of dropping out from college and from university, high school graduates with intermediate beliefs are more likely to enroll into college as a "safe" option. Therefore, the probabilities of receiving good versus passing grades in high school are pinned down by the amount of students who choose different educational paths. More college enrollment is indicative of low informational content in grades, while more university enrollment coupled with a higher rate of dropping out from high school is associated to high informational content of grades.

Wages work in a more straightforward manner. For any type of schooling, increasing the wage rate induces more people to enroll and complete the program. Increasing the high ability wage will mechanically increase average wages directly, but will also attract more students with higher beliefs. The average wages will increase indirectly as a result. Using a similar argument, the effect of increasing the low ability wages is somewhat ambiguous: it could directly increase the average wage of corresponding school program, it may also attract more low ability types into that type of school, decreasing the average wages at the same time. The sign of the overall effect depends on the magnitude of these two opposite effects, the ability distribution and the belief distribution.

Lastly, the education distribution helps pin down the proportion of high types individuals in our model population. Indeed, a high proportion of high types is associated to more university enrollment and graduation and lower dropout rates across all levels of education.

The important takeaway is that the model is identified and estimated jointly. We use a two-step approach. In the first step, we use the Genetic Algorithm to identify a set of unique learning and wage parameters which minimize the objective function. In the second step, we use the Genetic Algorithm again to refine the wage parameter estimates, by estimating pre- and post-cohorts separately while keeping the learning parameters fixed to the values estimated in the first step.