Essays on School Duration and Cohort Size: Implications from the 1999 Ontario High School Reform

Xian Zhang

Department of Economics
McGill University
Montreal, Quebec, Canada

April, 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Ph.D in Economics

Abstract

This dissertation is composed of two essays which provide a comprehensive analysis of the educational reform policy implemented in Ontario, Canada in 1999. This reform eliminated the fifth year of high school and established a new four-year curriculum, enabling students to obtain high school graduation certificates after a total of twelve rather than thirteen years of schooling. This quasi-experimental feature of the policy, together with rich data sources, enables me to investigate the policy effect along multiple dimensions and understand its underlying mechanisms. The first chapter of my thesis investigates the policy effect on students' academic performance during high school, their graduation decision, students' postsecondary education choice and earnings. The second chapter focuses on a special subgroup of students: the Ontario double cohorts who have university degrees. It aims to examine the impact of cohort size on labor market outcomes.

Abrégé

Cette dissertation est composée de deux essais qui fournissent une analyse compréhensive de la réforme d'éducation implémentée dans la province d'Ontario, Canada, en 1999. Cette réforme a éliminé la cinquième année d'école secondaire et a établi une nouvelle programme scolaire, permettant aux étudiants d'obtenir un diplôme d'études secondaires après douze plutôt que treize années d'études. La nature quasi-expérimentale de cette réforme, couplée à de riches sources de données, me permet d'investiguer ses conséquences en plusieurs dimensions et de comprendre ses mécanismes sous-jacents. Le premier chapitre analyse les effets de la réforme sur la performance académique des étudiants lors de l'école secondaire, leurs décisions à la graduation, leur études postsecondaires et leurs salaires. Le deuxième chapitre est dédié à un sous-groupe d'étudiants : la double cohorte Ontarienne qui détient un diplôme universitaire. Ce chapitre vise à examiner l'impact de la taille de la cohorte sur les résultats des diplômés sur le marché du travail.

Contribution of Authors

This thesis contains two essays. The first chapter is joint work with Andrei Munteanu, a doctoral candidate at Department of Economics, McGill University. He contributed to discussion of the structural model, model calibration and writing the draft. I contributed to coming up with the research idea, performing the empirical analysis, discussion of the model, model calibration and writing the draft.

Acknowledgements

First of all, I am deeply grateful to my supervisor, Prof. Fabian Lange and my committee members Prof. Markus Poschke and Prof. Saraswata Chaudhuri for extensive support. They have been generous with their knowledge, advice and willing to help me overcome all the difficulties that I came across during my research.

Besides my committee, I would like to thank Prof. Laura Lasio, Prof. Francisco Alvarez-Cuadrado, Prof. Francesco Amodio, Prof. Theodore Papageorgiou and Prof. Erin Strumpf for their encouragement, advice, revisions and consultations. I thank my friends Qian Sun, Ailin He, Alexander Amundsen, Masaya Takano, Mathieu T-Blais for their help both in research and also in making my six years in Montreal pleasant and unforgettable.

In addition, I would like to thank my co-author and best friend Andrei Munteanu for his insightful contribution to our research project, and also his indispensable help in polishing my thesis. Last but not least, I could never over emphasize the unconditional love and support from my family: Kun Zhang, Guiying Shen and Peiwen Zhang.

I thank the following for generous financial support throughout the past six years: Gordon Whitehorne Felw-00101AR, STP-Davidson-FQRSC 2011-SE-, Grad Excellence Award-00101, STP-Lange-SSHRC 435-2015-04, the Quebec Interuniversity Centre for Social Statistics and the Department of Economics, McGill University.

The analysis presented in this paper was conducted at the Quebec Interuniversity Centre for Social Statistics (QICSS) which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the QICSS are made possible by the financial or in-kind support of the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institutes of Health Research (CIHR), the Canada Foundation for Innovation (CFI), Statistics Canada, the Fonds de recherche du Québec - Société et culture (FRQSC), the Fonds de recherche du Québec - Santé (FRQS) and the Quebec universities. The views expressed in this paper are those of the authors, and not necessarily those of the CRDCN or its partners.

Contents

List of Figures

List of Tables

1

Introduction

The theory of human capital suggests that investments in schooling can increase labor market productivity and earnings. However, the effect of different educational inputs, such as student-teacher ratio, instruction time, class size on academic performance and wages is empirically unclear. In my dissertation, I focus on one input specifically: the school duration. The impact of school duration on students' academic achievement has not been considered by the literature (Krashinsky, 2014). This may be due to the fact that the quantity of years of schooling provided to students for a given program is endogenously determined in general. As a result, an exogenous variation in school duration is required to gauge any causal impact.

The natural experiment that provides causal identification is an educational reform in Ontario, Canada. In 1999, the Ontario government eliminated the fifth year of high school and implemented a new four-year curriculum, enabling students to obtain high school graduation certificates after a total of 12 rather than 13 years of schooling. This quasi-experimental feature of the policy, together with rich data sources, enables me to investigate the policy effect along multiple dimensions and understand its underlying mechanisms. Thus the first chapter of my thesis investigates the effect of high school duration on students' academic performance during high school, their graduation decision, students' post-secondary education choice and earnings. In addition to human capital investment, the labor market entrance condition also plays an important role affecting individuals' wage profiles. As a byproduct, the implementation of the Ontario educational reform generated the famous Ontario double cohort: the last five-year cohort and the first four-year cohort, both of whom graduated from high school simultaneously in 2003. Not surprisingly, a drastic increase in the number of new university graduates in Ontario occurred four years later in 2007. Thus the second chapter explores this exogenous increase in the number of new university graduates in Ontario to analyze the cohort size effect on their earnings.

The first chapter, "The Effect of High School Duration on Educational Attainment and Labor Market Outcomes: Evidence from Ontario", explores the policy effect along various dimensions. More specifically, this chapter studies how high school duration affects students' academic progress within high school, post-secondary educational choices and their labor market outcomes. Comparing to the previous literature that studies the same policy (Krashinsky (2009), Krashinsky (2014), Morin (2013)), this paper extended the treatment group to include all the affected cohorts and provided a more comprehensive evaluation of the policy change. This is achieved not only through covering more extensive outcomes, but also by exploring and disentangling the underlying channels through which the policy affects these outcomes.

Using the Survey of Labor Income Dynamic (SLID), the National Longitudinal Survey of Children and Youths (NLSCY) and the Labour Force Survey (LFS), I compare the difference between pre- and post-reform cohorts in Ontario with those in the rest of Canada to identify the policy effect. I empirically investigate the effect of the policy on high school duration, course failure rate, graduation rates, students' cognitive performance, their post-secondary education choices and earnings. The estimates demonstrate that, on average, students spend 0.4 fewer years in high school, as some still stay for an additional non-mandatory year. Although the overall learning process seems to be more efficient, as measured by school length, there is a polarization of outcomes. The portion of low ability students , defined as those whose overall academic performance was below average immediately before entering high school, failing at least one math course increases by 11% and the high school graduation rate is reduced by 1.4%. Moreover, the negative impact on students with lower ability persists. When measured at age 21 and 22, their numeracy skill is 0.2 standard deviations lower compared to cohorts educated in the old system. At the same time, the numeracy skills of higher ability students increased by 0.2 standard deviations. Conditional on graduating from high school, the policy resulted in an increase in post-secondary enrollment. Nonetheless, this increase was fully compensated by an increase in post-secondary drop-out rates, leaving the rate of graduation from post-secondary institutions unchanged when measured at ages $28 - 36$. Once students enter the labor market, our results suggest that there is approximately a 5% to 10% wage penalty within education level for those who have graduated from high school. Two possible explanations are proposed for these empirical findings: a composition effect due to misallocation of individuals into educational levels, brought on by a decrease in how informed students are about their ability ("orientation effect") and a wage rate decrease due to a decrease in human capital accumulation ("performance effect").

In order to disentangle and measure the relative importance of each channel, we propose and estimate a dynamic discrete choice model where students learn their own ability gradually by attending school and receiving grades. Thus, the grades do not only help accumulate credits, but they relay information regarding ability. These two functions correspond to the performance and the orientation channels discussed previously.

Our structural model helps us separate the two effects. 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 across educational categories. The rest of the difference can be explained by the performance effect. As high school is shortened, the orientation effect predicts an increase in college enrollment. This occurs because individuals tend to insure themselves against uncertainty in their abilities by choosing the relatively safe option of college. This effect is further distorted by wage rate changes. Specifically, the decrease in college graduates' wage premia compared with those of university graduates contributes to an overall increase in university enrollment in the new cohorts. This is the first paper to quantify the orientation effect and disentangle it from the performance effect in the literature. The result also sheds light on the role of school duration as an important input in the educational production function.

The second chapter of the dissertation, "The Misfortune of One's Birth: Cohort Size and Youth Earnings", analyzes the impact of labor supply shock on earnings. A result of the Ontario educational reform to reduce high school length was the creation of the Ontario double cohort: the last cohort educated in the old system and the first cohort who studied in the new secondary school structure. Though starting school at different time, both cohorts graduated from high school in 2003. Four years later, Ontario experienced a 28% increase in new university graduates. By comparing the pre- and post-wage gap between new entrants and seasoned workers with a Bachelor's degree in Ontario versus in the rest of Canada, the impact of an excess labor supply shock on earnings is identified. The triple difference-in-difference estimation using the Canadian Labour Force Survey (LFS) data suggests that the labor supply shock decreased the hourly wage rate by 8.7% to 13.5% among recent university graduates, which corroborates the result obtained using the Canadian Census data. Moreover, the impact of the Ontario double cohort is not uniform along the wage distribution. Both the bottom and top quartiles are negatively affected, leaving the median unchanged. I also find that the labor supply shock decreases the proportion of university graduates taking a full-time job in the first year of their career. Moreover, I observe an increase in the proportion of recent university graduates working in a "small" firm, and landing in a low-paid occupation and industry. Finally, the depressing effect on wage persists. The negative impact on wages rate exists at least for the first ten years after graduating from university.

To summarize, although the two chapters use the same educational reform in Ontario and study its impact on earnings, each chapter attempts to answer a different question and uses different sources of variation in the data. The first chapter explores how high school duration affects students' academic progress within high school, post-secondary educational choices and their labor market outcomes. It uses the exogenous decrease in high school duration as targeted by the reform. Specifically, the treated group includes individuals born between 1985 and 1988 who attended secondary school in Ontario with the new curriculum. The estimated wage penalty within educational level due to shortened high school duration is identified by comparing the wage difference between pre- and post-reform cohorts in Ontario with those in the rest of Canada.

On the other hand, the second chapter analyzes the impact of labor supply shock on earnings. This is achieved by investigating the exogenous increase in the number of new university graduates, which happened only in the first year when the reform was implemented. As a result, the treated group in the second chapter only includes cohorts born in 1984 and 1985 who attended secondary school in Ontario and graduated from the university in 2007. This time, the observed wage drop comparing to previous university graduates who belong to a normal cohort is mainly attributed to the congestion in the labor market resulting from the increased search intensity.

We have established that the Ontario educational reform of 1999 adversely affects the double cohort students by increasing the labor supply at their graduation and affects post-reform cohorts by reducing their years of schooling. Both of the aforementioned adverse effects are present in a special case: the Ontarian 1985-cohort. Within the Ontario double cohort who finished university in 2007, those born in 1985 who spent four years in high school on average faced a wage penalty comparing to their 1984-counterparts. This is explained by both "orientation effect" and "performance effect" as high school duration was shortened. Meanwhile, when comparing to individuals who graduated from the university before 2007, the 1985-cohort from Ontario experienced further wage decrease simply because they entered the labor market in the same year as the 1984-cohort did. The sudden inflow of extra university graduates increased the search intensity, and thus imposed adverse effects on their subsequent labor market outcomes.

2

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

2.1 Introduction

The choice of optimal duration of schooling is a fundamental issue for education policy (Meyer et al., 2015). On one hand, time spent in school, as one type of input for academic outcomes, 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, and also the indirect cost from straining tax and social security systems, increased shortage of skilled workers as well as delayed intergenerational transfers of know-how (Meyer and Thomsen, 2018).

Due to these concerns, 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, which enables us to analyze the causal impact of school length on students' educational choices and their labor market outcomes.

In this article, we will show that time spent in school is important not only in enabling students to accumulate human capital, but also because it orients students. Staying in school helps students understand their abilities and interests and make better informed decisions about their future academic and job market paths. We discuss the human capital and orientation channels in more detail throughout the paper, as we try to measure and disentangle their effects. We claim that policymakers cannot ignore the orientation effect and its long-term impacts on educational attainment, wages and, ultimately, output, when deciding on an optimal school program length.

We first examine how policy reform affected high school duration, high school course failure rates, students' cognitive performance, both during high school and after they graduate and high school graduation rates. In addition, we also look at the policy impact on students' post secondary educational decisions and ultimately, wages. By employing data from various sources, we exploit the variation in the length of high school 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 students on average spend 0.4 years less in high school. Though the overall learning process seems more efficient, students' performance suggests polarization. The portion of low ability students ever failing a math course increased by 11%. The negative impact on students with low ability persisted. When measured at age 20 to 21, their numeracy skill was on average 0.2 standard deviations lower compared to the old counterparts. On the other hand, the numeracy skill of high ability students increased by 0.2 standard deviations compared to cohorts educated in the old system. Over all, shortened high school length led to a 1.4% reduction in high school graduation rate. Conditional on graduating from high school, the fraction of individuals enrolled in post-secondary education increased. More specifically, it promotes immediate enrollment into university by 3% and college enrollment with one year delay by 5%. The increased enrollment in post-secondary education is offset by increases in dropout rates, leaving the educational attainments distribution unchanged when measured at age $28-36$. When it comes to wages, there is 5% to 10% wage penalty conditional on education for those who have graduated from high school under the new system.

Although this reduced-form analysis helps us measure the overall policy effect, it raises questions about the mechanism that induces these changes. For example, how does shortened high school duration boost post-secondary education enrollment while lowering the wage premium? Moreover, why does the post-secondary dropout rate increase? We hypothesize that there are two channels through which reducing high school duration affects further educational choices and labor market outcomes.

The first channel is the "performance effect". If we believe that years of schooling affect human capital accumulation, then the reduction in high school length would directly affect the overall human capital level of post-secondary bound students adversely. This would lead to a reduction in productivity and thus, wages and may explain the increase in the post-secondary dropout rate as well.

The second channel is the "orientation effect". First proposed by Meyer and Thomsen (2016), the orientation effect refers to the impact of high school length on students "comprehension" about their own ability and further educational choices. Students do not perfectly know their academic ability and each year spent in school helps students learn this ability; a reduction in high school length means that students, on average, make worse decisions as to the optimal education level they should pursue. Thus, shortened high school duration results in more low ability students enrolling in university but eventually dropping out as well as higher ability students never enrolling in post-secondary education at all. Given that innate ability also influences wages, then the aforementioned misalloaction of students brought on by the shortening of high school could also explain the decrease in wage premia for higher education.

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 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 structural model in which schooling decisions are made sequentially by individuals as they learn their innate ability. Our structural model is a dynamic discrete choice model in which students learn their ability by going to school and receiving grades. In this way, grades do not only help accumulate credits, but they relay information regarding ability. These two functions correspond to the performance and the orientation channels discussed previously. Our structural model helps us separate these two effects. 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 remainder of the difference can be explained by the performance effect.

Regarding the policy effect on the education distribution, the orientation effect predicts an increase in college enrollment so that students insure themselves against uncertainty. Their post-secondary education decision is further distorted by wage rate changes. Specifically, the decreases in college graduates' wage premia compared with those 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. There are studies which focused on double cohorts[2] given their educational choices: Krashinsky (2014) and Morin (2013) restrict their sample to be double cohorts who were admitted into university, 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 sample of study, could also be affected by the policy reform. Our paper extends the analysis along the following dimensions. First, not only did we analyze the policy effect on students' labor market performance conditional on educational attainment, but also 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 measured in their later life cycle. All these aspects together help support a comprehensive evaluation of the policy change.

[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 as well as the first new cohort, who were the first one to take the new curriculum after the reform. Notice that the reform was implemented in a way that in year 2003 the double cohorts graduate from high school in the same year, though there was a one year difference at the time 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.

Secondly, this is the first paper to quantify the orientation effect and disentangle it from the performance effect in high school context. The estimated orientation effect is not only essential in understanding the policy reform, but also provides insights into future improvements in high school "curriculum" design. After all, compressing high school duration without changing the curriculum will not simply shift individuals' schooling decision by one year as the policy maker would expect. Its impact on students' belief regarding to their own ability cannot be ignored as it induces long run effect on students' post-secondary education decisions.

Last but not least, 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 rarely considered by the literature on academic achievement, possibly because school length is generally endogenous (Krashinsky, 2009). This paper takes advantage of the natural experiment to evaluate the randomized change in school length on students' human capital accumulation, and proves its importance as an input in educational production function.

The rest of the paper is structured as follows. Section 2.2 reviews related literature; Section 2.3 provides additional information regarding the Ontarian educational system and the policy change in 1999, Section 2.4 and 2.5 summarize the data and empirical results, Section 2.6 presents the the stylized structural model, Section 2.7 includes the calibrated results and some counterfactual analysis. Section 2.8 shows the results of a policy exercise we conduct using the structural model. Finally, Section 2.9 concludes.

2.2 Lessons Learned: the Literature on the German and Ontario School Reforms

2.2.1 The German Experience

A similar policy to the one studied in this paper was more recently implemented in Germany and provides a quasi-experimental event to help understand the mechanism. Since 2001, most federal states in Germany shortened high school length from 13 to 12 years to promote earlier high school graduation as well as earlier labor market entry. While it decreased the school duration in terms of school years, the curriculum remained unchanged. As a result, students spent on average more than 33 hours per week in school after the reform, compared to less than 30 hrs before 2011.

There is a large literature exploring this reform and investigating how learning intensity and school length affect students' progress. Two papers find negative effect on students' school performance as measured by math grades (Büttner and Thomsen, 2015) and increased grade repetition rates and lower final grade point average (Huebener and Marcus, 2017). In terms of non-cognitive skill development, the authors do not seem to have reached a consensus. Dahmann and Anger (2013) investigates the short-term effects on students' personality traits, showing that students on average are more extroverted and less emotionally stable as a result. In contrast, Thiel et al. (2014) claims they could not discover any significant effects of increased learning intensity on the personality dimensions. Hofmann et al. (2017) studies the policy effect on overall mental health status, and finds significant negative impact. Meyer and Thomsen (2016) analyzes the impact on students' further educational choices. They suggest evidence for delayed, not necessarily lower enrollment into university, especially among female students. Among university students, Meyer and Thomsen (2018) concludes that the increased learning intensity affects students' perceptions of learning abilities without changing the droping out probability, nor the final grade[4].

Although this German policy is in many ways similar to the 1999 Ontarian policy,

[4]Eble and Hu (2017) studies a Chinese policy which extended primary school from five to six years without changing the curriculum. They find little effect on post-primary educational attainments, but the policy raises average monthly wage by 2.6%.

no article analyzes how labor market outcomes are affected by it. Moreover, the policy is slightly different in that the curriculum was unchanged, which greatly increased students' learning intensity. In the Ontarian policy, as we will see in more detail, the curriculum was modified.

Overall, although there is some evidence that reducing school length can increase the difficulty of school, adversely affect student academic outcomes and lead to changes in university enrollment, there is very little in the way of an analysis of school length on labor market outcomes. Moreover, the channels through which school length affects various educational and job market outcomes are unclear, which is why structural estimation is necessary.

2.2.2 The Ontario Education Reform

There is already literature studying the Ontarian reform considered in this paper. Krashinsky (2014) studies how one extra year of high school affect students' academic performance in university. In this paper, the author uses a policy exposure dummy to instrument for the number of years the double cohorts spent in high school, and estimate the Local Average Treatment Effect (LATE) of one extra year of high school education on performance in an individual course (an introductory level management course), as well as grade point average under the monotonic assumption of treatment. The author found that both in specific course and grade point average, four-year graduates perform five to ten percentage points worse than their five-year counterparts. Similarly, Morin (2013) measures the value-added of an extra year of high school mathematics for university-bound students. Taking advantage of the fact that different subjects of the curriculum were affected differently, the author used subjects unaffected by the reform (e.g. Biology) to control for potential ability differences between the two cohorts, and concluded that the extra year of mathematics has a small positive effect on their university grades – 2.2 points on a 100 point scale or 17 percent of a standard deviation in university grades. The estimated effect was much smaller compared to the one documented in Krashinsky (2014), which is mainly due to the difference in students' ability distribution in the two samples, as well as the heterogeneity in the treatment effect.

Krashinsky (2009) studies the same Ontarian policy as we do. The author checks the causal effect of one year high school education on individuals' wage using double cohorts who entered the labor market with high school diploma. He records that students who graduated after four years in high school were paid about ten percent lower than their counterparts one year after high school graduation, and these difference persisted even two years after high school graduation for less capable individuals but converged among more capable ones.

Overall, there is some evidence suggesting that the Ontarian policy negatively impacted the human capital accumulation of students and adversely affected their wages. However, it is unclear if these effects persist during and after post-secondary education. Moreover, it is unclear how post-secondary enrollment decisions are affected by the policy and its' underlying mechanisms.

2.3 Policy Description in Ontario

In this section, we discuss the salient characteristics of Ontario secondary education institutions. 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. In the following subsections, we will outline and compare the Ontarian high school system before and after 1999.

2.3.1 Old structure

Grade 13 has been part of Ontario secondary education for over 7 decades, from 1921 to 1988, until the Ontario Academic Credit (OAC) system was introduced. OAC refers to a series of courses offered at the senior level of high school and designed to cater to the needs of students who would pursue post-secondary education.

After 1988, Ontarian secondary school courses from Grade 9 to 12 were offered at one of the following levels of difficulty: the basic level, the general level and the advanced level. Basic level courses mainly served the needs of students who would not participate in post-secondary education and provided a good preparation for direct entry into employment. General level courses were designed to prepare students for employment, careers, or further education in certain programs in the colleges of applied arts and technology or other non-degree-granting post secondary educational institutions. Advanced level courses focused on the development of academic skills and prepared students' post-secondary attendance. Students who intended to enroll in university or college would normally take most of the compulsory credits at the advanced level of difficulty in order to prepare themselves for OACs, which were offered at the advanced level of difficulty only. Students who spend five years in high school typically took eight courses per year in the first three years, and then at least 6 OACs per year in each of the final two years (Krashinsky, 2014).

Indeed, between 20% to 25% (Casas and Meaghan, 1996) of students chose to repeat one or more OACs in order to boost their grades, which is known as "OAC grazing" (Brady and Allingham, 2010). One of the reasons of OAC repeating is that only the highest grade for a given OAC course would appear on students' transcript for university application . In short, though the introduction of OAC system was intended to bring Ontario in line with other English-speaking areas in North America by abolishing Grade 13 without actually eliminating the fifth year of secondary school, it failed to break the 13-year-secondary-education tradition in Ontario (Brady and Allingham, 2010).

The Ontario Secondary School Diploma (OSSD) required in total a completion of 30 credits. While OACs were not necessary to obtain an OSSD, completing six OACs serves as the (only) prerequisite for admission in universities in Ontario[5].

2.3.2 New structure

In 1997, Ontario's newly elected conservative provincial government announced that it would compress the secondary school curriculum from five to four years starting in 1999. In other words, students who entered high school after 1999 were expected to graduate within four years instead of five after completing 30 credits[6]. The new system basically left the educational curriculum unchanged for the first three years of high school, but decreased the number of courses available for students who wish to continue post-secondary education (Krashinsky, 2014).

This is in contrast to the previously-mentioned German policy, which increased the learning intensity by shortening secondary school duration without changing the curriculum at all. In Ontario, some changes were imposed on the curriculum besides the shortening of school length. In the reorganized program, three types of Grade 9 and 10 courses are offered: Academic, Applied and Essential types. The main difference in these three types lies

[5]This is different from U.S. university admission. These rely on standard test scores in their admission process. Ontario universities rank and select graduates of the province's secondary schools exclusively on their school marks. More specifically, the marks of six OACs of each applicant (Brady and Allingham, 2010).

[6]In order to earn an Ontario Secondary School Diploma(OSSD), a student must earn a minimum of 30 credits, including 18 compulsory credits and 12 optional credits. Student must also complete 40 hours of community involvement activities and pass the provincial secondary school literacy test.

A credit is granted in recognition of the successful completion of a course that has been scheduled for a minimum of 110 hours. Scheduled time is defined as the time during which students participate in planned learning activities designed to lead to the achievement of the curriculum expectations of a course. Planned learning activities include interaction between the teacher and the student and assigned individual or group work, other than homework related to the achievement of the learning expectations in the course. Therefore, there is no change in instruction time in terms of the minimum requirement to obtain OSSD.

in the focus rather than the level of difficulty. More specifically, Academic courses emphasize theory and abstract problem solving, Applied courses focus on practical applications and provide students with more hands-on experience. The Essential type is designed to focus on core knowledge and skills to meet individual student needs. Later in Grades 11 and 12, students take destination-based courses. University preparation courses are developed in collaboration with universities and designed to equip students with the knowledge and skills needed to meet the entrance requirements for university programs. The same logic applies to other tracks. Workplace preparation courses are designed to equip students with the knowledge and skills for direct entry into labor market or apprenticeship programs[7].

In grades 9 and 10, students select a combination of Academic, Applied and Essential courses in order to explore their interest, add to their knowledge and skill base and determine the type of educational program best suited for grades 11 and 12. In grades 11 and 12, students will focus more on their individual interests and identify and prepare for their post-secondary goals. A typical progress would be to finish eight courses each year and finish high school after grade 12.

In summary, under the new regime Ontario Secondary School Diploma (OSSD) would require students to complete 30 credits, 40 hours of community involvement activities and to pass some provincial secondary school literacy tests. In order to progress to post-secondary education, six University and/or University/College preparation courses are needed[8].

2.3.3 Comparison

Since we are interested in how shortening high school duration affects individuals' educational attainment and labor market outcomes, it is necessary to compare the old and new high school programs.

Let us first look at the high school graduation requirements. In the old system, students

[7] Open courses are also available, which are designed to prepare students for further study in certain subjects and to generally enrich their education. Open courses in Grade 9 and 10 are offered in all subjects other than those offered in Academic and Applied. These courses are designed to provide students with a broad educational base. Open courses in Grade 11 and 12 allow students to broaden their knowledge and skills but are not designed to meet specific requirements of university or college programs or workplace. (Ontario Secondary Education: Grade 9 – 12)

[8] Both types of courses can be considered as part of the 30 credits for OSSD

needed to complete 30 credits. In the new regime, students are also required to complete 40 hours of community service and to pass a provincial secondary school literacy test. According to King et al. (2004), it turns out that the community service and the literacy test did not set dissuasive barriers to students' graduation. Thus, the biggest change when it comes to high school graduation is the school length.

In the old structure, students who did not need OACs for post-secondary education application could graduate after Grade 12, which was exactly what policy makers expected when introducing OACs. In reality, however, a vast majority of high school students choose to complete their degree in five years (Krashinsky, 2014)[9]. This is especially but not exclusively valid among college or university-bound students. Thus shortening high school duration could increase the learning intensity even for non university-bond students. As a result, the policy reform would affect high school graduation negatively. On the other hand, earlier entry into the labor market would decrease the opportunity cost of high school graduation. These two forces work in the opposite direction, thus the impact of the policy on high school graduation decision needs to be examined empirically.

In terms of the pursuit of post-secondary education, the policy reform affects students' post-secondary choice through the following two mechanisms as suggested by Meyer and Thomsen (2016). By shortening high school by one year, human capital accumulation tend to be negatively affected. After all, students who wish to continue education are required to spend one year less in high school in the new system. In the old structure, six OACs are prerequisite for university or college admission. These are usually obtained by students in Grade 12 and the fifth year. In the new structure, a combination of six "4U" and "4U/C" courses are required and the mixed type is not directly comparable to OAC (King et al., 2004).

Thus, when years spent in school is viewed as one input of the education production function, human capital accumulation is negatively affected. This negative impact during high school would last until post-secondary education. There is evidence documenting negative effects on individuals' academic performance in universities, in particular in Krashin-

[9]In fact, Casas and Meaghan (1996) documented that fewer than 15% of students chose the fast track, which means that students could take 8 credits per year for Grade 9, 10, 11 and then 6 university preparatory credits for Grade 12. The Ontario Ministry of Education documented that only 8.3% of first-year university students were four-year graduates prior to 2003.

sky (2014) and Morin (2013). This corresponds to the "performance effect" proposed in Meyer and Thomsen (2016).

Moreover, shortened school length would affect students' educational choices indirectly through an "orientation mechanism". As high school is shortened by one year, students are less oriented about post secondary choices and they are younger. A less precise estimate of the completion probability will increase the proportion of "overly optimistic" students who enroll into college or university, but ultimately realize, too late, that they should have entered the labor market directly[10]. Consequently, we expect that the average ability of individuals who enrolled in post-secondary education would decrease as well. This is due to the decrease in how informed students are when making post-secondary education decisions.

Once individuals enter the labor market, both the "orientation effect" and "performance effect" will jointly change the wage distribution. The "performance effect" will change individuals' wage rate conditional on educational levels, and the "orientation effect" will affect the ability composition across educational levels. Students who are affected most by this policy are those whose with beliefs about their own ability being on the margin of post-secondary enrollment. Thus, the boost in post-secondary enrollment will lower the average ability of university-bound or college-bound students as long as their beliefs are rational.

To sum up, in theory, shortening high school by one year tends to increase learning intensity and negatively affect students' progress towards high school graduation. For post-secondary-bound students, the shortened high school duration will make them less oriented and less prepared for post-secondary training, which also tends to negatively affect their further post-secondary educational attainments. The exact policy impact on students' high school graduation decision, their post-secondary education choices and earnings will be explored in the empirical section 2.5.

[10]For example, when the old cohorts apply for post secondary education, most of them have observed some OAC courses result they took in Grade 12. However, in the new structure, when Grade 12 students filed the application, they would have not received any "4U" or "4U/C" courses. Thus, this change affects individuals' perception of successful completion of post secondary education.

2.4 Data Description

This paper mainly uses the Survey of Labor Income Dynamics (SLID), 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.

2.4.1 SLID

The 1993 to 2011 waves of the Survey of Labor and Income Dynamic (SLID) data are used for analysis. The SLID provides annual panel survey data on labor market activities and income information of individuals living in Canada since 1996[11]. Each respondent will be tracked for 6 consecutive years.

The variables that are essential to model identification are the province where the respondent obtained secondary education and their birth year. Since it is impossible to perfectly observe which curriculum an individual experienced in Ontario, we assume all the individuals who were born after 1985 from Ontario were exposed to the new regime. The threshold in birth year comes from the fact that students were required to be in Grade 1 the year when they turned 6 in Ontario and the assumption of no grade repetition[12]. 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 1975 would not enter the sample neither because they were exposed to different secondary school structure before "OAC" started.

To sum up, the treatment group in SLID include individuals who obtained most of their secondary education in Ontario and were born between 1985 and 1988. Cohorts born between 1975 and 1984 from Ontario and all the cohorts born between 1975 and 1988 from other provinces in Canada[13] serve as the control group.

SLID contains comprehensive yet detailed information about educational attainment.

[11]Longitudinal estimates are available up to and including 2010

[12]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.

[13]The province of Quebec is excluded because of their unique education structure.

The key variables used in this project include total years individuals spent in high school, the indicator of graduating from high school or not, and years of post-secondary education received. Other relevant demographic variables include respondents' birth year, gender, province of residence, immigration status etc. I exclude individuals who immigrated to Canada after age 13 to make sure they receive all their secondary education within Canada.

Another group of variables used in this project is related to individuals' (current) educational activity. In each survey year, respondents were asked if they were enrolled in any of the following institutions: secondary schools, trade schools, community colleges, business schools or universities. We grouped these indicators into whether or not attending college or university, and use the longitudinal aspect of SLID to study the policy effect on students' transition from high school to post-secondary education.

2.4.2 LFS

The Labour Force Survey (LFS) is an ongoing monthly survey, which is used by the Canadian government to produce unemployment as well as other standard labor market indicators such as the participation rate and employment rate. LFS data covering years 2000-2017 (the latest wave available) are used in this project.

Comparing to SLID, LFS does not have information as to where the individual received their secondary education. Thus we use province of residence instead[14].

LFS does not contain rich information on educational attainments, yet the large sample size and the fact that it collects data on the treated cohorts till now make it possible to measure the policy effect in the long run. Other than the educational attainment, LFS provides detailed information on labor market outcomes, which helps bridge the gap between human capital accumulation and labor market outcomes. As to educational attainment variables, LFS contains relatively limited information, including an indicator if the individual graduated from high school or not, their highest grade obtained in high school if they are dropouts, and also their highest level of education category. Besides, LFS also asks respondents' current educational activities, which proves to be essential to understanding the

[14] 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.

policy effect on individuals' post-secondary educational choices. Last but not least, the labor market relevant variables used in LFS include their hourly wage rate, labor market status etc.

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"[15]. 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[16].

The presence of measurement error in one of the key outcomes of interest is not ideal.

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

[16]A more detailed and complete analysis is illustrated in Appendix.

Unfortunately, there is no other data sets that could provide more precise information. Other than LFS, the only Canadian data source that is up to date and has a reasonable sample size is the 2016 Census. However, respondents only report the highest certificate they ever obtained, leaving it impossible to identify post-secondary dropouts.

2.4.3 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's 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[17], 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.

Relevant variables in NLSCY include individual socioeconomic status when they were 15, which was derived based on the parents' education level, the household income category, the parents' occupation. The dataset also includes their school performance at age 14 or 15, right before entering high school, as well as their high school attendance status, whether they have ever failed a course or not, their post-secondary choice. Besides school performance, NLSCY provides direct assessment of students' ability. Among others, the one used in this paper is the math cognitive skill measure among the 16 and 17 years-olds and the numeracy test conducted when they are 21-22 years old.

The advantages of NLSCY data are threefold. First of all, it documents rich historical information including children's family background, school performance history. This is extremely useful when analyzing factors that shape students' educational decisions made in the past. Secondly, NLSCY records not only respondents' education attainment (like in SLID), but also students' school performance and their aspirations. These could be used to

[17] 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).

study the dynamics of students' educational choices. Last but not least, NLSCY provides a variety of direct assessment of individuals at various ages. These data provide an objective, comparable and standardized measure of respondents' ability in specific fields. Despite these advantages, NLSCY is only conducted every two years. This makes it more difficult to accurately track individuals' decisions over time.

2.5 Empirical Results

We start by estimating Eq. (2.1):

$$Y_{ijk} = Cons + \text{Policy}_{ijk}\theta + \alpha_j + \beta_k + X'_{ijk}\gamma + \varepsilon_{ijk} \tag{2.1}$$

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 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. (2.1) include the years individuals spent in high school, indicators of whether they ever failed a course in high school, their cognitive ability measure at age $16-17$ and $21-22$, the students' high school graduation status, their post-secondary education attendance dummy and their highest educational attainments measured at age $28-36$. These results will be presented in detail in the following section.

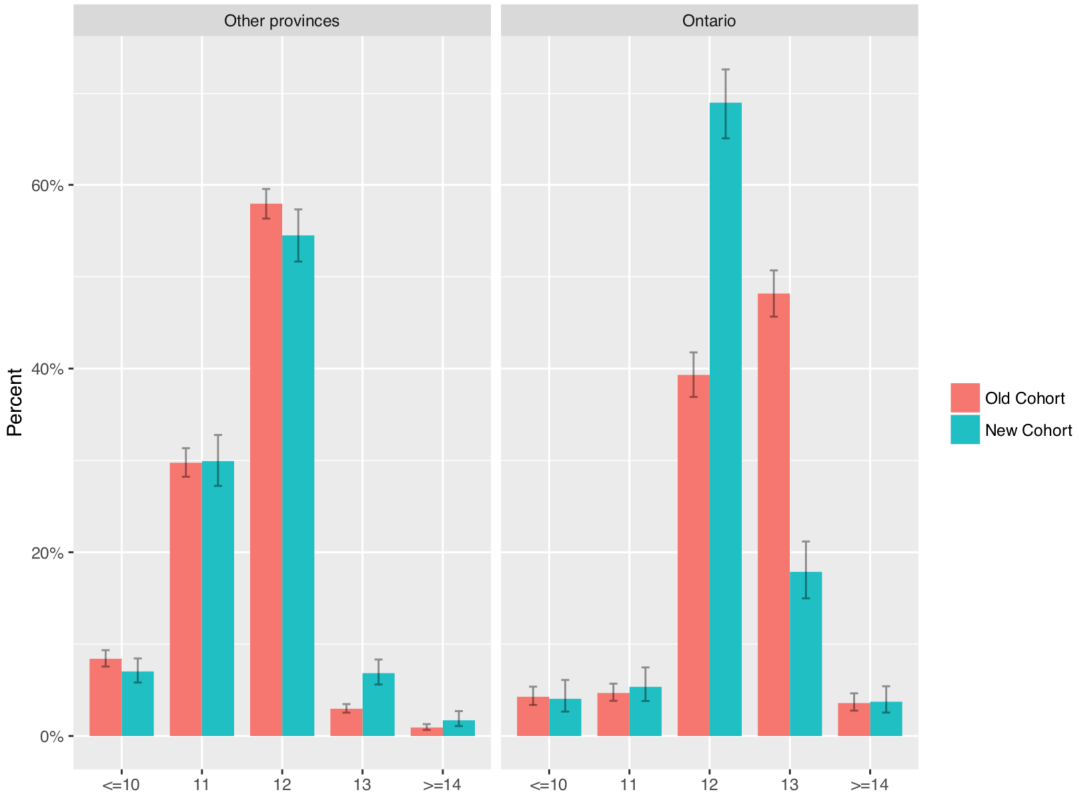
2.5.1 Number of years spent in high school

The reform's principal objective was to shorten high school duration. Thus, the very first analysis investigates whether the reform has been effective in this regard.

Before jumping into regression estimation results, we first present some descriptive statistics. Fig.2.1 shows the distribution of years spent in elementary and high school using SLID data[18]. This graph includes all the individuals born between 1975 to 1988 at age ranging between 22 and 24, who are no longer attending high school during the survey year, and weighted using the cross-sectional weight based on the 2011 Canadian population. Individuals who attended high school from all the Canadian provinces other than Ontario

[18] Although the exact length of secondary education varies across provinces, the total length of pre-secondary education, i.e. elementary and high school, is always 12 years all the Canadian provinces other than Ontario and Quebec. In addition, the length of pre-secondary education in all the Canadian provinces except Ontario and Quebec did not change over the period of interest

Figure 2.1: Years spent in elementary and high school



Source: SLID 1993-2011. Including all the cohorts born between 1975 and 1988 who attended high school in provinces other than Quebec and Quebec are shown on the left panel and the right one only includes those from Ontario. The red bar represents cohorts who were born before 1985 and the green bar represent new cohorts born between (including) 1985 and 1988. It is obvious from the right panel that the fraction of individuals who spent 13 years in elementary and high school dropped nearly $2/3$. At the same time, the number of individuals who stayed elementary and high school for 12 years roughly doubled after the policy reform. In comparison, slightly more Non-Ontarian new cohorts spend more than 13 years in pre-secondary education. The overall change is not as significant.

In fact, the effect is heterogeneous across students' ability distribution. If we split the sample according to whether they received any post-secondary education or not, as shown in Fig. 2.2, it is obvious that the reduction mainly comes from the university- or college-bound students.

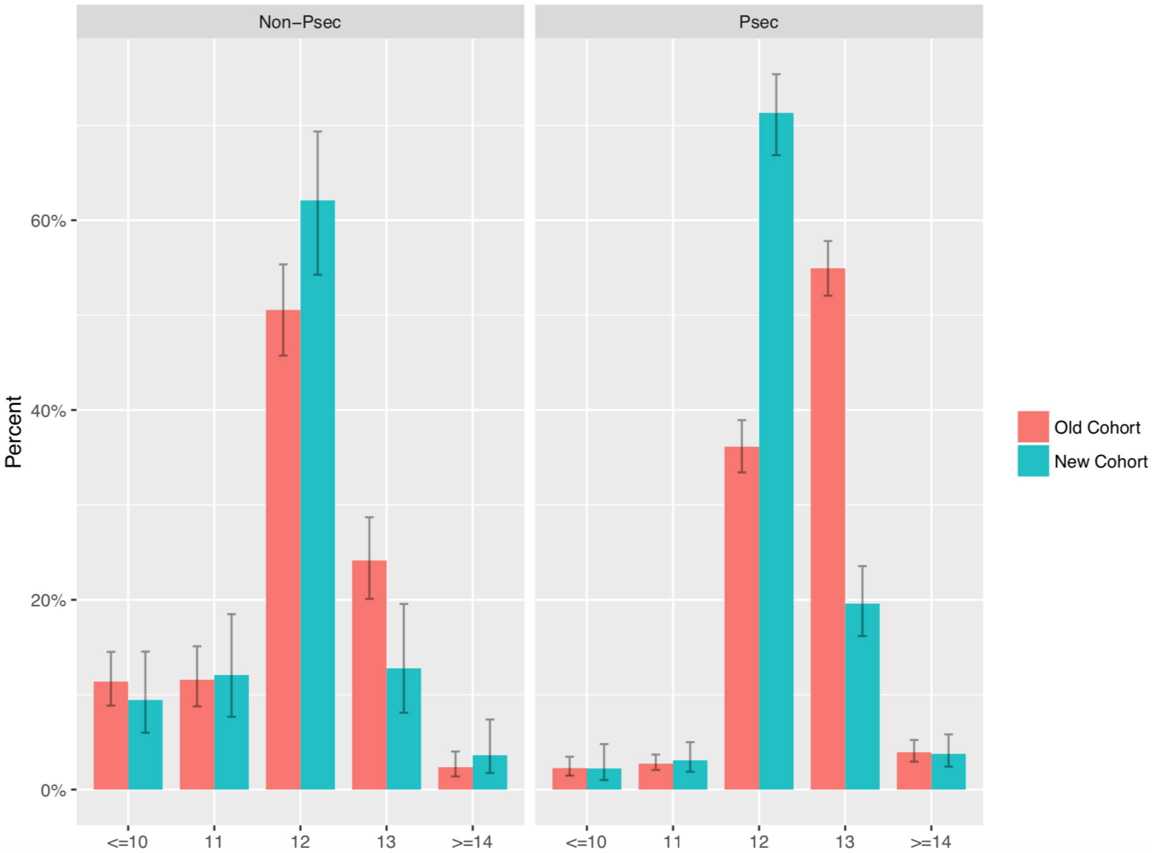
Regression estimates of the policy effect is presented in Table 2.1. The first column uses the years spent in elementary and high school as dependent variable and columns $2-8$ uses dummies indicating if individuals spend the corresponding years. For example, the dependent variable in column 7 equals to 1 if the individual spent 13 years in elementary and high school and 0 otherwise.

Column 1 suggests that on average students spend 0.4 years less in high school as a result of the reform, and this decrease is mainly driven by the drop in the proportion of people leaving high school after 13 years, as targeted by the policy. To be more specific, the proportion of students spending 13 years decreased by nearly 37% while the proportion of individuals who spent 12 years increased by 33%. Surprisingly, the proportion of 11-year high school drop-outs also increased in response.

Similar results are obtained using LFS data, as summarized in Table 2.2. Again, we notice that the policy has reduced the fraction of individuals who spend 11 years or more in pre-secondary education, while the proportion of individuals who spend 10 years or less in elementary and high school has increased among the new cohort. Indeed, the policy increased the likelihood of dropping out of high school by 0.7% and decreased the proportion who completed at least Grade 11 by 1.4%.

NLSCY data does not include the exact number of years students spend in high school,

Figure 2.2: years spent in elementary and high school, Ontario by psec status



Source: SLID 1993-2011. Including all the cohorts born between 1975 and 1988 who attended high school in Ontario

Table 2.1: Policy effect on years spent in high school, SLID

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Indicator			
	S_{hs}	$S_{hs} \leqslant 8$	$S_{hs} = 9$	$S_{hs} = 10$	$S_{hs} = 11$	$S_{hs} = 12$	$S_{hs} = 13$	$S_{hs} \geqslant 14$
Policy	-0.437***	0.002	0.004	0.008	0.034***	0.325***	-0.370***	-0.004
	(0.043)	(0.004)	(0.004)	(0.005)	(0.010)	(0.018)	(0.017)	(0.006)
N	24397	24397	24397	24397	24397	24397	24397	24397
R^2	0.081	0.004	0.005	0.006	0.018	0.181	0.237	0.019

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SLID 1993-2011. Dependent variable of first column is years spent in elementary and high school, those for columns $(2) - (8)$ uses dummies indicating years spent in elementary and high school. Sample includes all the individuals born between 1975 and 1988, with age $22 - 24$ and who were not attending high school in ref year. Other control variables include gender, immigrant status, year, cohort, province of high school fixed effect. Standard errors are in the parentheses. Variance is clustered within province $\times$ birth year.

Table 2.2: Policy effect on years spent in high school, LFS

	Indicator		
	$S_{hs} \leqslant 8$	$S_{hs} = 9 - 10$	$S_{hs} \geqslant 11$
Policy	0.007***	0.007*	-0.014***
	(0.00)	(0.00)	(0.00)
N	20900	20900	20900
R^2	0.03	0.03	0.054

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: LFS 2000-2017. Dependent variable is the dummy indicating years spent in elementary and high school. Sample includes all the individuals born between 1975 and 1988, with age $22 - 24$ and who were not attending high school in ref month. Other control variables include gender, parents' education category, year, cohort, province of residence fixed effect. Standard errors are in the parentheses. Variance is clustered within province × birth year.

but it is possible to observe their school attendance status at various ages as illustrated in Table 2.3. The first two columns show the policy effect on high school attendance probability for all the individuals aged between 16 to 18 and the remaining columns further break this down into different ages.

Combining the results in odd columns, it seems that there is a slight decrease in high school attendance among all the individuals between age 16 to 18, which is mainly driven by the drop at age 18, i.e. the fifth year in high school[19]. This is consistent with the policy goal to encourage students to finish high school after four years.

The fact that the fifth year was only required for post-secondary bound students in the old system makes it less clear how the policy would affect students who prefer to work after high school graduation. In order to analyze the heterogeneity in policy effect among students with different post graduation choice, we categorize them into two ability groups based on their overall academic performance in Grade 7 or 8, right before entering high

[19]Here age is measured at the first half of school year. So age 18 corresponds to the fifth year of high school.

Table 2.3: Policy effect on high school attendance, NLSCY

	High school attendance							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age (yrs old)	16 – 18		16		17		18	
Policy	-0.079**	-0.071**	0.027	0.019	0.025	0.045**	-0.162***	-0.236***
	(0.034)	(0.034)	(0.018)	(0.025)	(0.027)	(0.018)	(0.029)	(0.030)
Policy · Low type		0.041		-0.015		0.027		0.159***
		(0.032)		(0.052)		(0.025)		(0.054)
Low type		-0.028		-0.045**		-0.087***		0.064*
		(0.017)		(0.019)		(0.022)		(0.040)
N	8625	6702	3156	2313	2999	2184	2470	2205
R^2	0.425	0.440	0.040	0.072	0.036	0.071	0.145	0.182

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: NLSCY Cycle 1 - Cycle 8. Dependent variable is the high school attendance dummy at different age. Sample includes all the individuals born between 1982 and 1988. Low type means students whose overall academic performance was "Average" or "Below average" in Grade 7 or 8. 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.

school. The reason that we do not use their post-secondary education status directly is that the post-secondary education decision is an endogenous choice made by students. On the other hand, the ability type determined before starting high school can be viewed as orthogonal to the policy change. In the regression, the high type consists of around 70% of students whose overall performance was described as "very well" or "well", and the low type include the rest 30% who performed as "average or below" in the corresponding school year.

At age 17, low type students have lower high school attendance rates on average, and the policy does not affect high type and low type students' attendance differently. The homogeneity disappeared at age 18. Individuals who were ex ante performing above average and stayed for the fifth year in high school decreased by approximately 23.6% due to the policy change. The impact among low type students is negative: individuals who were performing poorly and were willing to stay for the fifth year decreased by 7.7% as a result. This may be due to lower drop out rates or because it takes longer for low ability students to graduate from high school after the reform.

Looking at the statistics from the sample, the proportion of people attending the fifth year decreased from 65% to 42% in Ontario while the level is fixed around 20% in other provinces. Among all the Ontarian fifth year attenders, the proportion of high type dropped from 75% to 54%, and around 50% of the old cohorts would start post-secondary education within one year after high school graduation, but the proportion of fifth year students later attending college is less than 5% for new cohorts. These facts are consistent with the analysis above: The policy mainly reduces high school length for students of high ability.

In summary, the policy shortened high school duration by 0.4 years on average, and also increased proportion of high school drop outs slightly. This reduction in average high school duration mainly comes from the shift from 13 to 12 years as targeted by the policy. This is supported by evidence from all the three data sources. Beyond the average effect, heterogeneity occurs along students' ability distribution as proved by NLSCY data. The policy mainly shortened high school length for students of higher ability. Presumably, for students whose ex ante school performance was above average, a considerable proportion of them will attend the "OAC" year for post-secondary education, but in the new curriculum, they manage to finish after four years regardless of the further educational choice.

This explains the decrease in high school duration.

But why is the policy less effective among students of low ability? One possible explanation is that students' human capital accumulation progress was negatively affected by the new curriculum. This potential mechanism will be inspected next.

2.5.2 High school course failure rate

Surprisingly, as many as 20% of individuals in the sample report having failed a course in high school at least once, and this portion varies across subjects. For math courses the portion is even higher: 50% of students have failed at least one high school math course. The course failure rate measure is informative especially among students in the bottom tail of ability distribution. It is also connected to the number of years students spent in high school in the sense that it takes time to make up the lost credits. In Table 2.4, we present how the policy affected students' course failure probabilities in high school in general and also in different fields.

Overall, students of lower ability type have higher chance of failing a course in high school, which holds almost in any subject. Moreover, the policy did not impact the overall course failure percentage. However, the zero mean masks heterogeneous effects both with the fields and student ability.

More specifically, though the probability of failing a high school math course is not affected on average, students with low ability are affected significantly. In fact, low ability students are 11.2% more likely to fail a math course in high school as high school duration is shortened[20]. Conversely, the probability of failing an English course decreased by nearly 6 percentage points, and low ability students became less likely to fail science course as well.

In short, shortening high school length increased the math course failure rate among students of low ability, but decreased their chance of failing a science course. In terms of English literature, a decrease in failure rate is achieved among students of both types.

[20]The point estimate of policy effect on students with low ability is 0.112 with standard error 0.036

Table 2.4: Policy effect on course failure rate in high school, NLSCY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall		Math		English		Science	
Policy	-0.016	0.009	0.026	0.04	-0.057***	-0.046***	-0.003	0.015
	(0.033)	(0.043)	(0.024)	(0.029)	(0.015)	(0.012)	(0.022)	(0.020)
Policy · Low type		0.008		0.072		-0.054***		-0.054***
		(0.060)		(0.049)		(0.019)		(0.020)
Low type		0.207***		0.101***		0.100***		0.096***
		(0.032)		(0.020)		(0.018)		(0.015)
N	4552	4078	4539	4067	4540	4068	4539	4067
R^2	0.082	0.123	0.045	0.074	0.034	0.056	0.033	0.053

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: NLSCY Cycle 1 - Cycle 8. Dependent variable is the dummy variable indicating if the student has ever failed a certain course in high school. Sample includes all the individuals born between 1982 and 1988. Low type means students whose overall academic performance was "Average" or "Below average" in Grade 7 or 8. 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 × birth year.

2.5.3 Math skill measure

Though lacking information on students' grades in high school, NLSCY conducted direct Math skill assessment among $16-$ and $17-$ year-olds in Cycle 4 and 5. This direct assessment provides more objective measures of students' skill without the possible contamination by misreport errors from schools or teachers. The test comprises questions from the pilot of the Programme for International Student Assessment (PISA 2000) mathematics test. Unlike course failure rates, this standardized measure provides information along the entire distribution of students' math skill.

Column 1 in Table. 2.5 suggests that students' overall cognitive skill, when measured at age $16-17$, is improved as high school is shortened. The improvement in students' progress during high school is valid for both the high and low types. Note that we already control the grade fixed effect, meaning the "positive" effect on their cognitive skill is not simply because students are learning faster as high school duration is shortened. Rather, the improvement could be attributed to the curriculum change that comes with the reform. When we differentiate students according to their ex-ante math performance[21], the conclusion does not change.

However, a different story is implied when comparing students' Numeracy test results conducted later, at ages 20-21 (Columns $4-6$). The Numeracy Assessment aims to test the ability of young adults to function in society and manage mathematical demands in diverse situations. On average, cohorts exposed to the new curriculum perform no different than the old counterpart, but the distribution is more dispersed. In fact, individuals with higher ex ante math ability are performing 0.2 standard deviations higher, but lower-ability individuals are doing much worse (The policy effect on students with low ability is -0.088 with standard error 0.17. The policy effect on student with low math ability is -0.201 with standard error 0.157.).

To summarize, the policy helps students accumulate more math skill as they progress through high school. Unfortunately, the homogeneity in the advantage fails to persist. When surveyed at age $20-21$, the new cohorts' numeracy skill distribution is more dispersed than the old counterparts. Students with higher (or lower) math ability ex ante are doing even

[21] Students are grouped according to their performance in math course in Grade 7 or 8

Table 2.5: Policy effect on math and numeracy scores

	(1)	(2)	(3)	(4)	(5)	(6)
		Math			Numeracy	
Policy	0.158*	0.208**	0.275***	-0.032	0.139	0.209**
	(0.079)	(0.080)	(0.091)	(0.065)	(0.094)	(0.087)
Policy · Low type		0.093			-0.227	
		(0.117)			(0.229)	
Low type		-0.675***			-0.718***	
		(0.052)			(0.069)	
Policy · Low type $_{math}$			-0.167			-0.410**
			(0.135)			(0.203)
Low type $_{math}$			-0.729***			-0.780***
			(0.056)			(0.095)
Grade fixed effect	Y	Y	Y			
Age	16 - 17	16 - 17	16 - 17	20 - 21	20 - 21	20 - 21
N	3130	2806	2803	2487	2292	2288
R^2	0.119	0.228	0.276	0.105	0.204	0.252

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: NLSCY Cycle 1 - Cycle 8. The dependent variable is a cognitive skill measure and a numeracy ability measure. Sample includes all the individuals born between 1982 and 1988. Low type means students whose overall academic performance was "Average" or "Below average" in Grade 7 or 8. Low type$_{math}$ means students whose performance in the Math course was "Average" or "Below average" in Grade 7 or 8. 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 × birth year.

better (or worse). This looks more worrying given that the requirement for math credits of high school graduation was increased from 2 to 3, hoping to equip students with better math skills.

One possible explanation for the widened performance gap is the increased learning intensity, in which case students have to take in more knowledge for a fixed duration. Thus we checked if the shortened high school duration has any effect at all on the number of hours students' spent on school work outside of instruction time and the probability of pursuing private tutor for help. The effects are poorly estimated in general, but could provide suggestive evidence still. The first two columns in Table 2.6 show that shortened high school duration increased the number of hours students of high ability type spent on school work each week, but not among low ability students. In my sample, students spent around 5 hours per week on school work so the coefficients convey a sizable effect economically. On the other hand, students with low ability are less likely to hire a private tutor (The policy effect is -9.7% with standard error 0.038) as high school was shortened, while the effects on high ability students are not significantly different from zero.

In short, as high school was shortened from five to four years, students of low ability failed to exert as much effort, which contributes to the widened performance gap.

2.5.4 High school graduation rate

What does the theory predict regarding the effect of shortening high school length on high school graduation rate? In the human capital framework, two opposing forces play a role. On one hand, the opportunity cost of high school graduation decreased 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 decision in opposite ways.

Table 2.7 presents the policy effect on high school graduation decision using all the data sources. SLID suggests that shortening high school duration does not affect high school graduation rate, yet negative effect is documented using LFS data (The likelihood of graduating from high school decreased by 1.4%). This negative effect implies the increased learning intensity due to the curriculum change outweighs the effect of the lowered oppor-

Table 2.6: Policy effect on working hours and private tutor hiring

	(1)	(2)	(3)	(4)
	Hours		Tutor	
Policy	0.241	0.392	-0.030	-0.035
	(0.440)	(0.557)	(0.030)	(0.032)
Policy · Low type		-0.778		-0.062
		(0.841)		(0.039)
Low type		-1.459***		0.04
		(0.523)		(0.030)
Grade fixed effect	Y	Y	Y	Y
N	6800	5066	6800	6800
R^2	0.107	0.136	0.029	0.044

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: NLSCY Cycle 1 - Cycle 8. Dependent variable is hours spent in study outside of school hours per week and whether hired private tutor. Sample includes all the individuals born between 1982 and 1988. Standard errors are in the parentheses. Variance is clustered within province × birth year. All models include gender, household social economic score at age $14-15$, single parent status at age $14 - 15$, year and cohort, province of high school fixed effect. Low type means students whose overall academic performance was "Average" or "Below average" in Grade 7 or 8.

tunity cost. The policy effect estimated using NLSCY is also negative yet slightly bigger. The quantitative gap could be due to the difference in the population each data set represents[22]. 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, which would bias the estimates towards zero as classic measurement error literature predicts. Statistically, the difference in regression estimates on high school graduation rate using SLID, LFS and NLSCY is not significant.

Column 4 in Table 2.7 suggests that on average, low-type students are 12.8% less likely to graduate from high school when measured at age between 22 and 24. The estimates do not support the hypothesis that students with different ability type are affected differently when high school duration is shortened when it comes to high school graduation decision.

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. Let us now take a closer look at the distribution of the number of years individuals choose to stay in high school. The most striking phenomenon is the shift from the portion of those spending 13 years to those spending 12 years in high school, as aimed by the policy makers. This decrease is mainly driven by students with above-average ex ante academic performance before entering high school. Besides that, the proportion of high school drop outs increases. This negative effect on high school graduation rate and increase in drop out rate could be partially explained by the increase in course failure rate in high school. When measuring students' math skill, they outperform their old cohort counterpart. However, the long run numeracy effect is heterogeneous and lasting. Students of higher math ability perform better while those with lower math ability were left behind more.

[22]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 2.7: Policy effect on high school graduation rate

	(1) SLID	(2) LFS	(3) NLSCY	(4) NLSCY
Policy	0.008	-0.014**	-0.070**	-0.044
	(0.017)	(0.005)	(0.029)	(0.032)
Policy · Low type				-0.038
				(0.050)
Low type				-0.128***
				(0.024)
N	24404	20800	2801	2577
R^2	0.020	0.069	0.080	0.092

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is high school graduation indicator. Source I: SLID 1993-2011. Sample includes all the individuals born between 1975 and 1988, with age $22-24$ and who were not attending high school in ref year. Other control variables include gender, immigrant status, year, cohort, province of high school fixed effect. 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.

2.5.5 Effect on the post-secondary education choice

Conditional on graduating from high school[23], 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 individuals post high school decision-making.

First, we take advantage of the panel feature of SLID data. In SLID, each panel was followed for 6 consecutive years. The respondents were asked if they were enrolled in various types of educational institutions in each wave. Thus we construct a gap year variable using the difference between the first year when the individual enrolled in post-secondary education and the last year when he or she enrolled in high school minus one. More specifically we grouped enrollment into any of the following institution as college: trade school, business school, community college and CEGEP. If the gap equals to zero, it means immediate enrollment. If the gap is positive, it suggests delayed post-secondary enrollment. Since each panel lasts no more than 6 years, the gap is capped at value 5. In the sample, on the other hand, it's rare to see cases with gap year longer than 2.

In order to adjust for sample attrition problem, we imposed the assumption of "Missing at random" or selection on observables, including family income at age 16, whether or not living with their parents at age 16, their parents' education, the ethnicity background, gender and the province where their high school is located. Then, we combine the predicted probability of remaining in the survey (probit model) with the original cross sectional weight to correct for cycle attrition.

After adjusting for sample attrition, we regress the indicator of gap year equal to certain number on the model suggested in Eq.(2.1). We find that conditional on attending post-secondary education, there is a 11% increase in immediate enrollment into universities (i.e. gap year between high school and university equals to zero), a 12.6% decrease in university enrollment after one year, and 1% increase in delaying at least two years, which is poorly estimated. On the other hand, the policy affected the gap year for college bound students in a similar way qualitatively but none of the coefficient is statistically significant, with the

[23]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 thus are omitted here.

Table 2.8: Policy effect on gap year, SLID

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Psec			College			University	
gap year	0	1	>= 2	0	1	>= 2	0	1	>= 2
Policy	0.010	-0.010	-0.000	0.015	-0.027	0.012	0.116**	-0.126***	0.010
	(0.007)	(0.007)	(0.006)	(0.038)	(0.037)	(0.031)	(0.048)	(0.046)	(0.028)
N	17972	17972	17972	6151	6151	6151	6947	6947	6947
R^2	0.024	0.019	0.016	0.245	0.174	0.044	0.254	0.205	0.076

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: SLID 1993-2011. Dependent variables are school attendance indicators measured at various time. Sample includes all the individuals born between 1975 and 1988 and who were not attending high school in ref year. Other control variables include gender, immigrant status, year, cohort and province of high school fixed effect. Standard errors are between brackets. Variance is clustered within province × birth year.

scale being much smaller.

Table 2.9: 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 $\times$ birth year.

These results using SLID data are consistent with the findings using LFS data. 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 time after high school graduation.

Four categories are documented in LFS: "Not registered as student", "Attending secondary education", "Attending college " and "Attending university"[24]. Here "college" stands

[24]There is the fifth category as "Attending other type of educational institutions ". This is being omitted because it represents not more than 2% of the population and the specific type of education remains unclear

for the 2-year college in the U.S. context, including community college, vocational school etc. 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[25], 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, the increase happened with some delay. Finally, when measured at one year after high school graduation college enrollment increased by 5%.

NLSCY data suggests a similar story as documented in Table 2.10. The overall post-secondary education enrollment within one year after high school graduation increased by 7%. Though the heterogeneity of policy effect on students with different ability type is rejected, the policy increased the low ability students post secondary education enrollment by 11%. Conditional on enrolling into post-secondary education, a shift from college into university among students with high ability occurred. The proportion of high ability students attending college decreased by 17.5% (The policy effect on low ability type students is 0.017 with standard error 0.077), while 10% more of them chose universities instead.

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. More detailed analysis shows that the increase in post-secondary enrollment is mainly attributed to students with low ex-ante ability type, and the shift away from college into university happened among people with above average ex ante performance. This could be rationalized by the following story: with less experience in high school, students are less informed about their own ability level. The increased uncertainty about whether one could finish college or university, together with the huge wage premium with the degree, encourages more students to try post-secondary education.

[25]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

Table 2.10: Policy effect on Psec activities, NLSCY

	(1)	(2)	(3)	(4)	(5)	(6)
	Psec		College		University	
Policy	0.071**	0.070	-0.150***	-0.175***	0.064	0.105*
	(0.030)	(0.045)	(0.047)	(0.048)	(0.050)	(0.060)
Policy · Low type		0.040		0.157**		-0.075
		(0.063)		(0.064)		(0.092)
Low type		-0.117***		0.223***		-0.265***
		(0.028)		(0.030)		(0.028)
N	2704	2547	2239	2114	2239	2114
R^2	0.167	0.171	0.124	0.190	0.155	0.213

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: NLSCY Cycle 1 - Cycle 8. Dependent variables are post-secondary enrollment, college enrollment and university enrollment measured at no later than 1 year after high school graduation. Sample includes all the individuals born between 1982 and 1988 who have graduated from high school. Low type means students whose overall academic performance was "Average" or "Below average" in Grade 7 or 8. Other control variables include gender, household social economic score at age $14 - 15$, single parent status at age $14 - 15$, year, cohort, province of high school fixed effect. Standard errors are between parentheses. Variance is clustered within province × birth year.

2.5.6 Long-run effect on educational attainment

Among the three data sources, only LFS could shed light on the long run effects on educational attainments in the population[26]. As we mentioned in Section 2.4, we corrected the measurement error in LFS by comparing the educational distributions in LFS with those in SLID[27]. 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 2.11 presents the policy effect estimated using original data from LFS and the second panel includes the corrected version. Quantitatively, the policy effect is bigger using the corrected the data, yet the qualitative conclusion is the same. From Table 2.11, 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 degree, fewer obtained college certificates and there are more post-secondary dropouts as well.

To sum up, shortening high school length not only affects individuals educational choice in high school, but also their further post-secondary educational decision. Overall, the policy promoted post-secondary enrollment. More individuals attended university immediately after high school graduation, but the increase in college enrollment happened with 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 post-secondary certificate in the population.

[26]Both SLID and NLSCY ends when the treatment group were in their early twenties.
[27]The detailed process is presented in the appendix.

Table 2.11: 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 $\times$ 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.

2.5.7 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, as well as 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 (2.1) with log of hourly wage as the dependent variable. The sample includes cohorts born between 1981 and 1988 who worked as employees measured within age range $28-32$ 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, only gender and age dummies are included to ensure sufficient sample size.

The policy induces a 4.4% drop in hourly wage rate (Column (2) Table 2.12). 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 (2.2),

$$\begin{aligned} Y_{ijk} =& Cons + X'_{ijk}\gamma + \\ & S_{ijk}\beta + \text{Policy}_{ijk}\theta + \alpha_j + \beta_k + \\ & S_{ijk} \cdot \text{Policy}_{ijk}\beta_s + S_{ijk} \cdot Prov_j\alpha_{sj} + S_i \cdot cohort_k\beta_{sk} + \varepsilon_{ijk} \end{aligned} \tag{2.2}$$

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 (2.1), 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 University

Table 2.12: Policy effect on log hourly wage rate

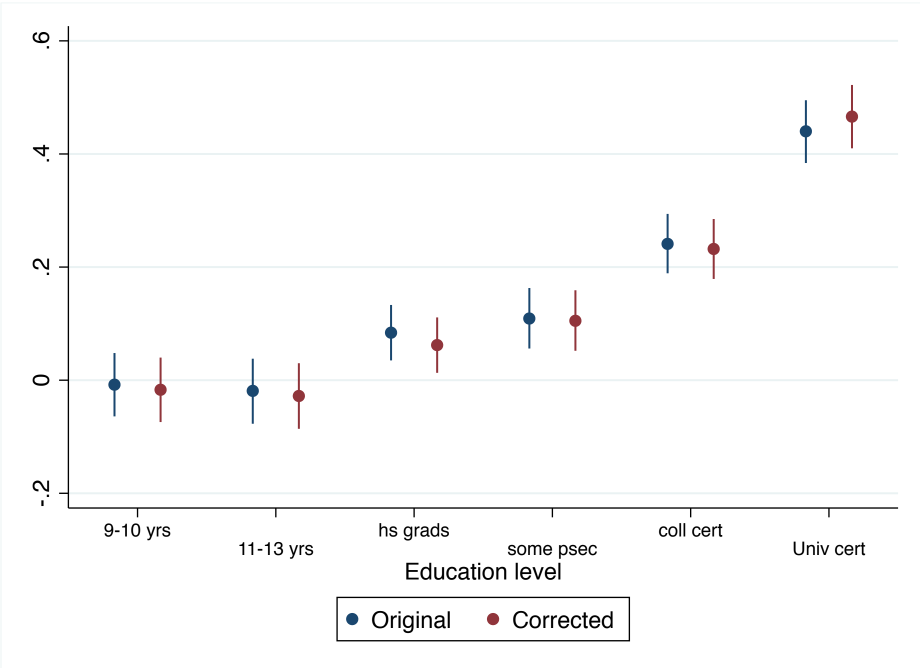
	(1) $log(w)$	(2) $log(w)$	(3) $log(w)$
Policy		-0.044***	0.046
		(0.006)	(0.095)
Policy · 9-10 yrs of edu			-0.032
			(0.112)
Policy · 11-13 yrs of edu			-0.053
			(0.095)
Policy · HS grads			-0.111
			(0.091)
Policy · Some psec			-0.08
			(0.097)
Policy · College			-0.067
			(0.094)
Policy · University			-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 $\times$ year.

diploma contributes a premium as high as 44%. Column (3) documents the heterogeneity in policy effect across educational level. Though the heterogeneity is rejected, 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[28], Figures (2.3) and (2.4) show the corrected return to education and policy effect.

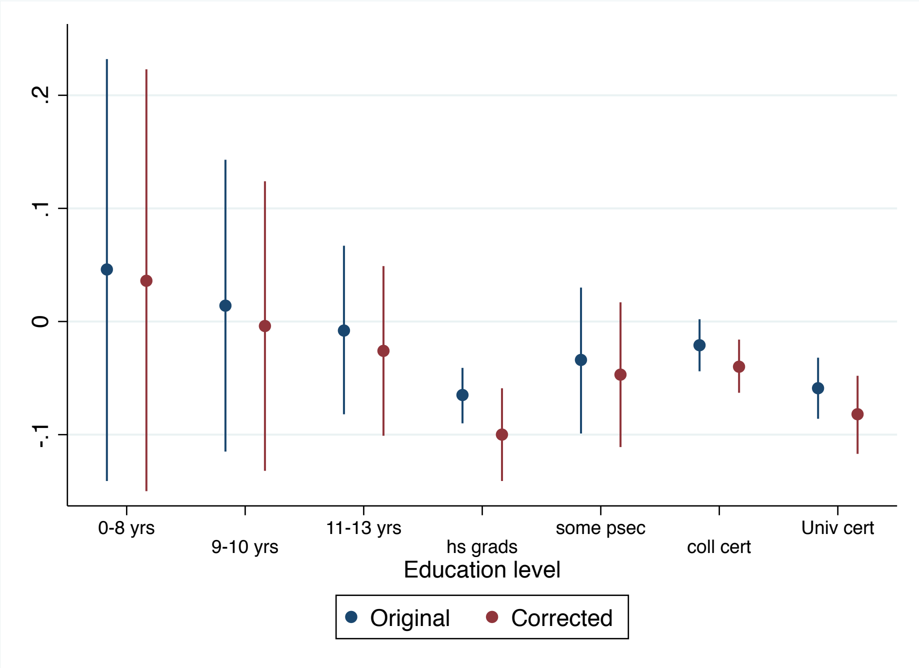
Figure 2.3: Return to education: log hourly wage rate



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. 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.

[28] See section 2.10

Figure 2.4: Policy effect on log hourly wage



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.

As shown in Figure 2.3, 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 as a result also because the exclusion of those who are university dropouts, but report holding university degrees by mistake. The return for post-secondary dropouts and college graduates is not affected much quantitatively in the end.

Figure 2.4 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 cohorts who entered the labor market after high school graduation. He documents that four-year graduates earn around 5% to 10% less than the five-year ones, and this effect persisted for less able graduates two years after graduation.

Putting all the evidence together, 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 drop out 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, shortened high school affect students' educational decision 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 is beyond the empirical identification strategy that we discussed so far. In the next section, we will present a stylized model where individuals make schooling decisions sequentially as they learn their own ability to further evaluate underlying mechanism through which high school duration influence students.

2.6 Model

In this section, we present a dynamic model where agents learn their own ability as they make schooling decisions[29]. Individuals' 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 own ability and observe signals through school experience. They then update their belief 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 and join the labor force or to continue their education. If they accumulate the necessary amount of credits T^j, students receive a diploma and graduate. In each period that students stay in school, they 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 a tuition cost τ.

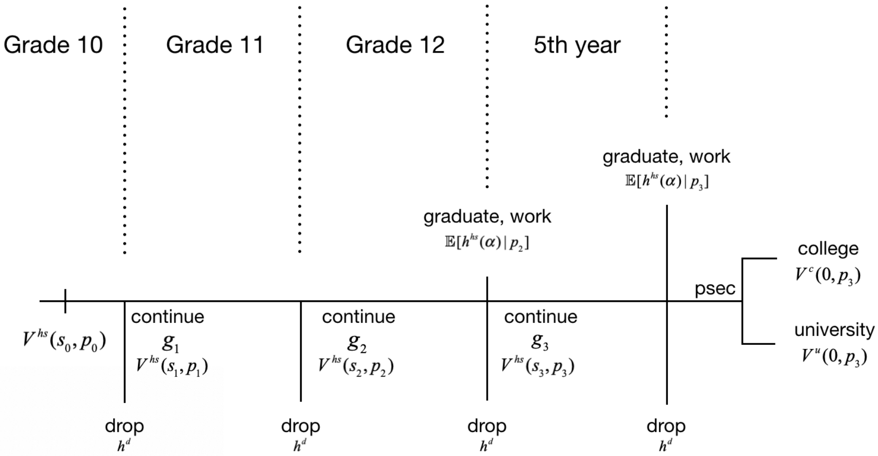
Once they enter the labor market, the individual earn wages corresponding to his own ability α and his educational level. This means that the individuals' ability type is revealed in the labor market. This simplifying assumption makes it easier to solve the model without having to compute firms' beliefs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

$$h(GS, j, \alpha) = \begin{cases} h^j(\alpha) & GS = 1, j = \text{hs, C, U} \\ h^d & GS = 0, j = \text{hs} \\ h^{hs}(\alpha) & GS = 0, j = \text{C, U} \end{cases} \tag{2.3}$$

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

[29] We modified the model in Trachter (2015) by incorporating learning process in high school and omit the possibility to transfer from colleges into university.

Figure 2.5: 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. $E[h^{hs}(\alpha)|p_t] \equiv p_t h^{hs}(1) + (1-p_t)h^{hs}(0)$ is the (expected) wage profile received by high school graduates given belief p_t.

graduating from college, and college graduates have higher wage 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, meaning the observed average wage premium of dropouts compared with high school graduates is fully attributed to the composition effect in our model. We also assume that the wage profile for high school dropouts does not depend on their ability α. This could be interpreted as there are very few individuals with high ability end up dropping out of high school, thus their wage profile is determined by low-types mainly.

Figure 2.5 illustrates the Ontarian old secondary school system following the aforementioned time line with some complications. At the beginning of each period up to Grade 12, students simply choose to drop out or to continue school for another year. After completing Grade 12, however, they face three possible choices: to drop out, to graduate or to continue the fifth year. Those students who have accumulated enough credits[30] could either

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

choose to graduate immediately and enter the labor market or to stay for the fifth year for post-secondary 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 existence of the OAC year not only prepares students for post-secondary education, but also serves as an extra opportunity to make up failed courses to obtain high school degree in the model. After the fifth year, those students who accumulated enough credits could graduate from high school and start their career. 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 policy maker. 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 who are struggling with graduating from high school after four years.

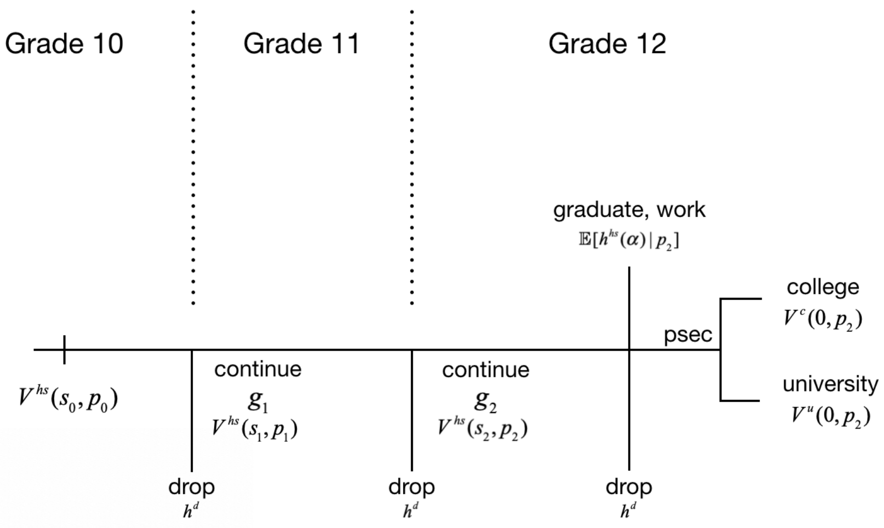
This extra year advantage disappeared as the new four-year system was implemented in 1999. In other words, students need to graduate within four years and failing a year leads to an automatic dropout. This modeling choice reflects the increase in learning intensity for low ability students, highlighted in Section 2.5. As shown in Figure 2.6, all the students in the new system have to accumulate enough credits in order to graduate after Grade 12 and make post-secondary education enrollment decisions.

Once enrolled into post-secondary education, the time line can be illustrated recursively in Figure 2.7, where institution j could be either college or university. College and university are different in three dimensions: the duration, the tuition cost as well as the wage premium of corresponding graduates.

Having established the time line of individuals' decisions in different school programs, we present more model specifications next. Assume the cost of attending school j is τ^j, which is constant for all the individuals, with $\tau^U > \tau^C$ and normalize $\tau^{hs} = 0$[31]. A student could graduate from institution j after accumulating required credits T^j, with $T^U = 4$ and $T^C = 2$. Also assume that students could not stay longer than the required number

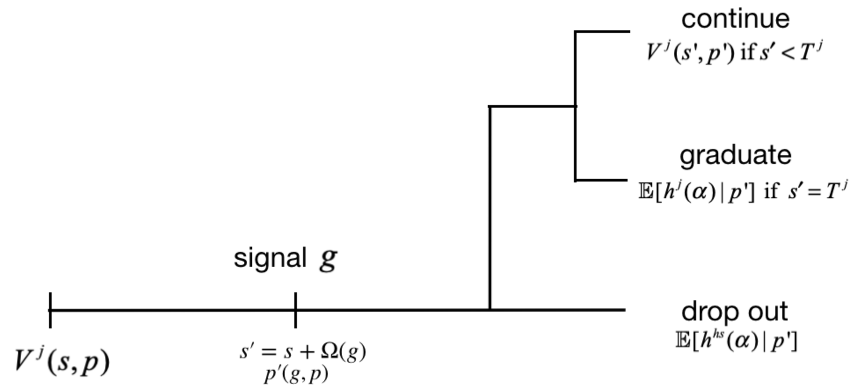
[31]There is only national level university tuition cost data available from Stats Canada. According to Vaillancourt (1995), the ratio of average annual cost of college to university is 0.38, which is used to calibrate the cost parameters in the structural model.

Figure 2.6: 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. $E[h^{hs}(\alpha)|p_t] \equiv p_t h^{hs}(1) + (1-p_t)h^{hs}(0)$ is the (expected) wage profile received by high school graduates given belief p_t.

Figure 2.7: 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. $E[h^{hs}(\alpha)|p] \equiv ph^{hs}(1) + (1-p)h^{hs}(0)$ is the (expected) wage profile received by college or university dropouts given belief p. $E[h^j(\alpha)|p] \equiv ph^j(1) + (1-p)h^j(0)$ 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.

periods. High school length T^{hs} changes from five to four years upon the implementation of the reform. The evolution of credit accumulation is tied to signals received during school program, which will be presented 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 $\frac{f(g_1|\alpha=1)}{f(g_1|\alpha=0)} \geqslant \frac{f(g_0|\alpha=1)}{f(g_0|\alpha=0)}$ for any $g_1 > g_0$. This assumption guarantees that the high ability students are more 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,

$$s^j_{t+1} = \begin{cases} s^j_t + \Omega(g_t) & \text{if } s < T^j \\ s^j_t & \text{if } s \geqslant T^j \end{cases} \tag{2.4}$$

and $\Omega(g_1) \geqslant \Omega(g_2)$ 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

$$\Omega(g) = \begin{cases} 0 & \text{if } g = F \\ 1 & \text{if } g = P \\ 1 & \text{if } g = G \end{cases} \tag{2.5}$$

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,

$$p_t \equiv pr(\alpha = 1|g_t, p_{t-1}) = \frac{1}{1 + \frac{pr(g_t|\alpha=0)}{pr(g_t|\alpha=1)}\frac{1-p_{t-1}}{p_{t-1}}} \tag{2.6}$$

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

$$H(g_{t+1}|p_t, g_t) = pr(g_t|\alpha = 0)(1 - p_t) + pr(g_t|\alpha = 1)p_t \tag{2.7}$$

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

$$pr^{hs}(g = F|\alpha = 1) = 0 \tag{2.8}$$

$$pr^{C}(g = G|\alpha = 0) = 0 \tag{2.9}$$

$$pr^{U}(g = G|\alpha = 0) = 0 \tag{2.10}$$

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 and drawing a G signal in university or college fully informs the recipient that they have high ability. The remaining learning parameters are calibrated without imposing any restrictions.

The model starts with individuals at the age of sixteen[32] 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 her expected life time 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 which reflects different labor market entry ages.

Table 2.13 lists the exogenous values used to calibrate the structural model. The risk free rate is set to be 1.8 percent, implying a yearly interest discount factor of 0.98. 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 Canadian full-time student 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 undergraduate program following Vaillancourt (1995)[33]. We assume students need to complete two school years to graduate from college and the university program lasts for four years. In order to obtain the distribution of students' initial belief, we use their math score measured at age 14-15 conducted by NLSCY.

The parameters governing the wage profile $h^i(\alpha)$ as well as the learning process $q_\alpha^j(g)$ are calibrated jointly to match the school distribution and wage distribution observed using LFS data, as well as a set of micro moments using NLSCY data.

[32] Sixteen is the minimum age to leave school in Ontario before 2008.
[33] All the monetary values are normalized by wage of high school dropouts h^d in the calibration exercise.

Table 2.13: 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

2.7 Calibration

2.7.1 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 $\mathbf{\Omega}$, where each element $\omega_{ij} = 1/\mathbf{v_{ij}}$. We then minimize the weighted distance between the vector of simulated moments $\mathbf{s}$ and the vector of observed moments $\mathbf{m}$:

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

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, post-secondary 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 post-secondary dropouts both before and after the reform[34].

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:

$$Pr(d_{10,i} = 1) = \Phi(X_i \beta) \tag{2.11}$$

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[35], 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.(2.11) to predict individuals' initial belief as indicated by Eq.2.11

$$\hat{p}_{0,i} = (1 + e^{-X_i \hat{\beta} + e_i})^{-1} \tag{2.12}$$

where $\hat{\beta}$ is the estimated coefficient from Eq.(2.11) and e_i is drawn such that the computed belief matches the observed educational decision of the individual at age 16[36]. Our model

[34] For the old cohort, we match college and university dropouts separately. For the new cohort, due to data limitations, we can only match post-secondary dropouts, without being able to distinguish between college and university dropouts.

[35] 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.

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

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, the micro moments were calculated 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 individuals in the population and the standard deviation of the initial belief. The calibrated parameters are presented in Section 2.7.3.

2.7.2 Identification

Next, we discuss how the model is identified. 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.

2.7.3 Calibration Results

Table 2.14: Calibrated wage parameter values

	HS drop	High school		College		University	
Ability α		high	low	high	low	high	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.

Table 2.14 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 educational level. When educational level is fixed, high type individuals earn a higher wage rate. The wage target used in the data are measured when individuals are $28-32$ years old, thus these estimates account 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 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 post-secondary education is four to five times smaller than the high ability counterparts. After high school was shortened, the wage premium of high ability high school graduates did not change much but those for low type ones decreased significantly. For college and university graduates, the wage rate decreased by around ten percentage points regardless of ability level. These wage penalties imply negative impacts on students' human capital accumulation as high school is shortened by one year.

Table 2.15: 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).

Table 2.15 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 post-secondary education also indicates high type. The likelihood of getting a "Fail" is a lot higher for low ability students in post-secondary education, thus failing the exam could lower their expectation regarding the innate ability. Overall the signals in post-secondary education are more informative than those in high school. Comparatively, high school is easy and more informative with respect to the lower tail while post-secondary education provides information for both ability types. Comparing college and university programs, the main difference 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 how noisy their initial beliefs are. We find that the proportion of individuals with high ability accounts roughly 46% and that the standard deviation of their initial belief is 0.34.

Table 2.16 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.

Table 2.16: 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

Given the calibrated parameters, let's take a closer look at how high school duration affect students' educational decisions. Figure 2.8 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 into 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 group. The boost in post-secondary enrollment cannot be explained by the wage rate drop, which would make post-secondary education less attractive in expectation. Thus the "orientation" channel plays a significant role in explaining the pattern as shown in Figure 2.9.

Figure 2.8: Predicted education attainment distribution

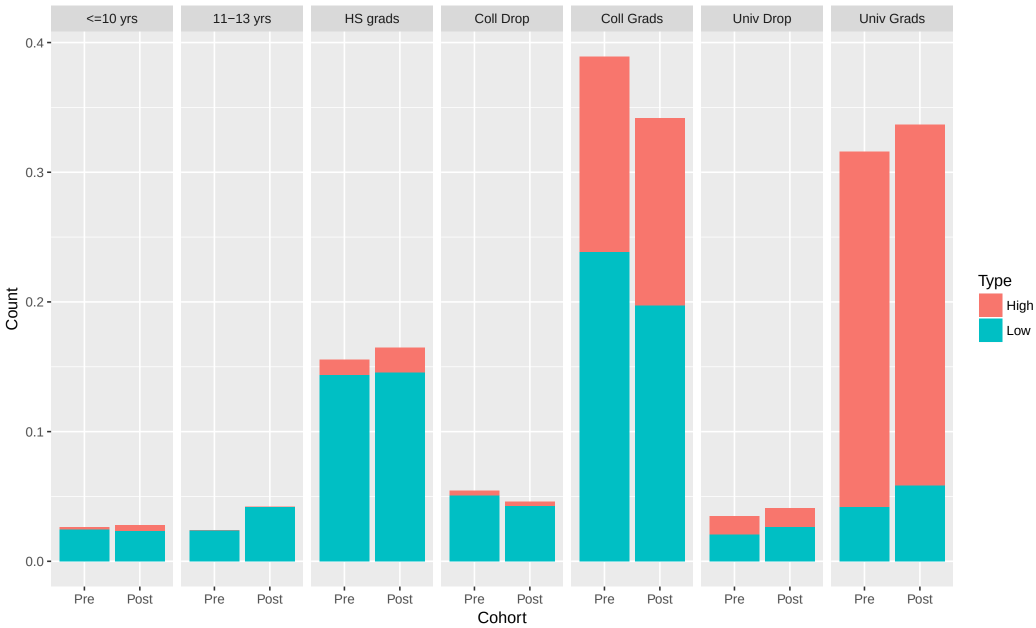
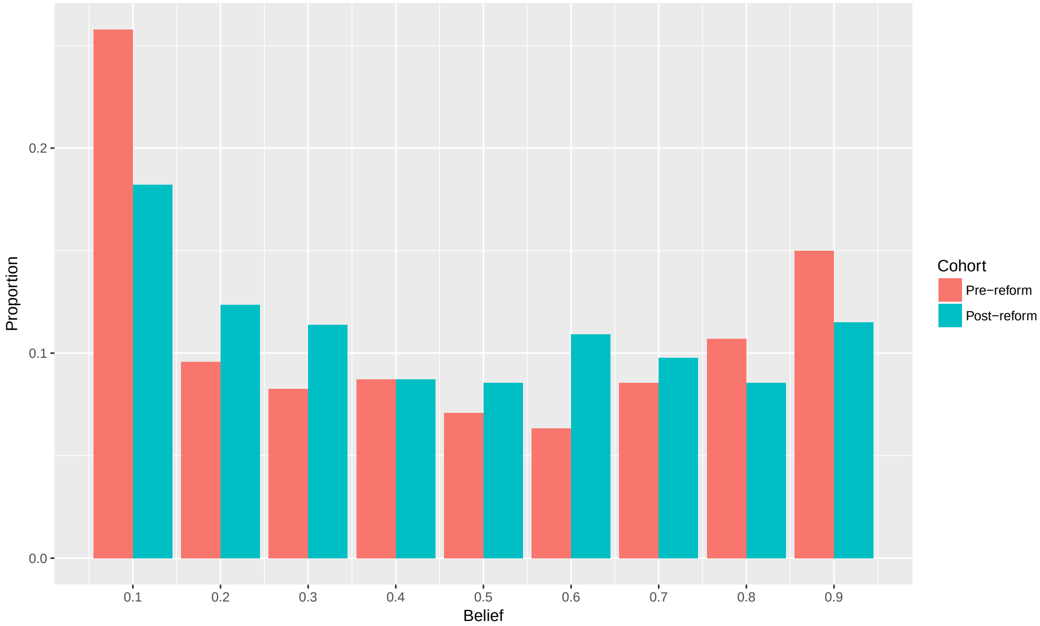


Figure 2.9 and 2.10 show the histogram of students' belief when making post-secondary education enrollment decision with the red bars representing old system and green bars indicating the new regime. It's clear that pre-reform cohorts are more informed than their new counterparts in the population. This "ambiguity" in students' belief about their true ability, together with a potential high wage premium of post-secondary degrees, encourages more post-reform students to try post-secondary programs, only to find out it's optimal to drop out later on.

2.7.4 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. As a result, any resulting changes in educational distribution and wages can be attributed solely to the orientation effect.

Figure 2.9: Histogram of belief after high school graduation



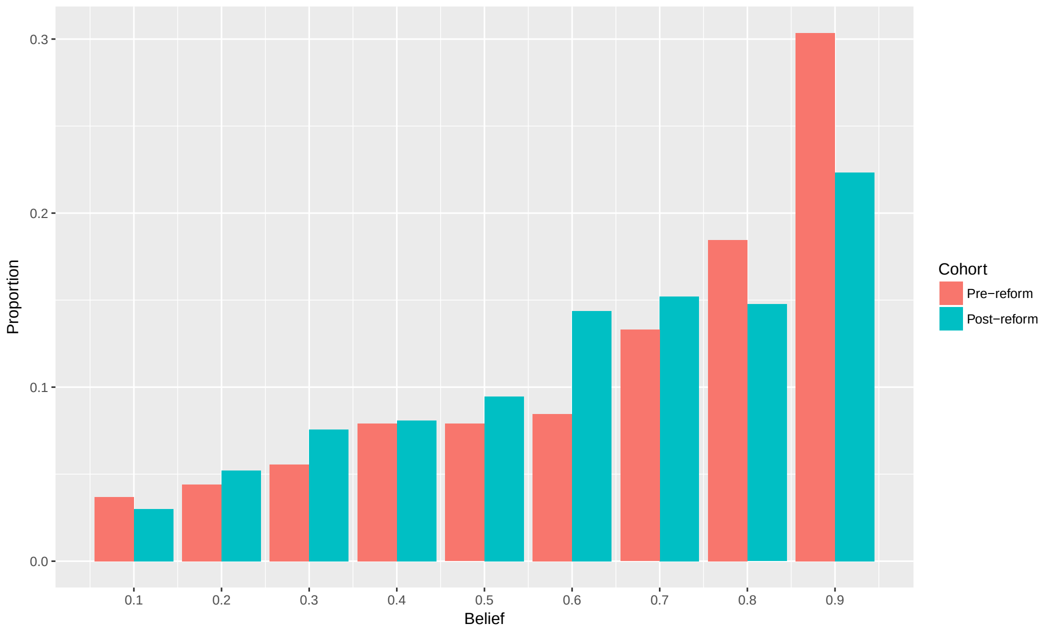
As we can see in Figure 2.11, if wage rates were fixed, shortened high school duration induces a decrease in how informed students are upon high school graduation. As a result, many more high school graduates would end up enrolling into college instead of university or instead of simply entering the labor market immediately. This is because college is especially appealing to students who have intermediate beliefs. On 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 2.17 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 college wage rate must decrease enough in order to explain the eventual drop in the fraction of college graduates (4.8 p.p.).

In terms of the wages, Table 2.18 breaks down the differences in average hourly wage

Figure 2.10: Histogram of belief after high school graduation by ability type

(a) High ability



(b) Low ability

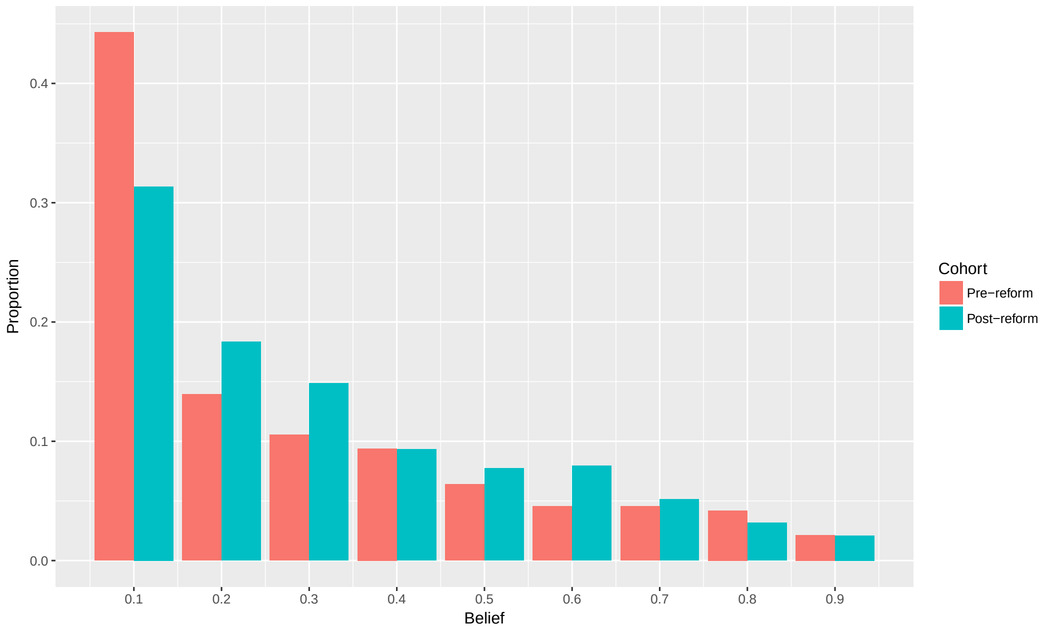
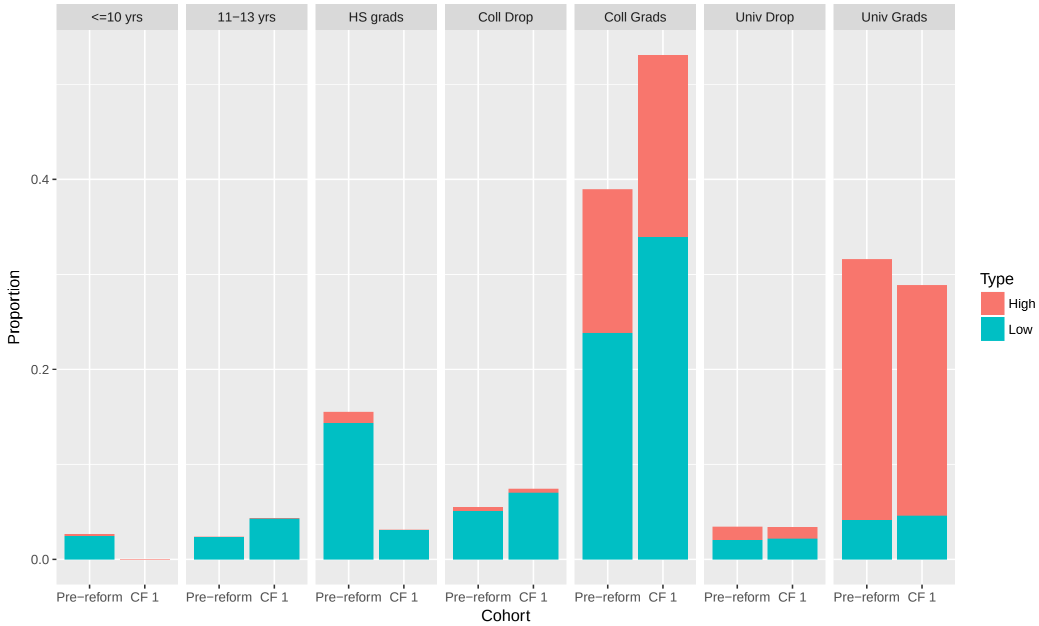


Figure 2.11: Predicted educational attainment distribution: Counterfactual I



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

Table 2.17: Decomposition of Change in Educational Distribution Due to Policy

(P.P.)	Δ Total	Δ Due to Orientation Effect	Δ Due to Performance Effect
$\leqslant$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

rate between the pre- and post-reform cohorts into the orientation and the performance component. Not surprisingly, the orientation effect negatively impacts employee earnings. This is because, as high school is 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 prevailing factor that explains the average wage decrease between the pre- and the post- reform cohort, especially among post-secondary graduates.

Table 2.18: 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, we can say that the orientation effect accounts for between 11% to 20% of the total decrease in wage premia and between 19% and 78% of the decrease in wage rates. The rest of the decreases can be explained directly by wage rate decreases that are due to lower human capital accumulation.

Table 2.19: 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

2.8 Policy Analysis

Table 2.20: 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.

In this section, we analyze a hypothetical environment in which there is an upward shift in the distribution of grades. We show that 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 simply 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 at lightening the high school curriculum, reducing the requirements or lowering the passing grade for courses in high school in order to make 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 2.20. They imply that low ability students are more heavily impacted by this grade shift than high ability types. We make this choice in order to reflect the fact that low ability students are typically more sensitive to changes in curriculum difficulty than high ability students.

Table 2.21: Educational attainment distribution after lowering requirements

Attainment (%)	Baseline	Policy	Difference
$\leqslant$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 perfectly 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 2.21, we see that grade inflation leads to a boost in high school graduation. More surprisingly, inflating grades actually 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 into university program. Moreover, as the the information content of high school grades decreases, there is more misallocation of students into educational levels, decreasing the average ability of university students slightly. In summary, we want to draw attention to the signaling role of school. Though school grade inflation can boost graduation rate, it leads to a decrease in the quality of students' further educational decisions due to the deterioration in the precision of knowing their own ability.

2.9 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 by influencing how students learn their true ability (orientation effect).

We evaluate the 1999 Ontarian policy change in order to tease out the effects of school length on multiple outcomes. Using SLID, 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 the policy decreased high school duration by 0.4 years on average. Shortened high school length led to a polarization of students' performance. Low ability students were 11% more likely to fail a math course and their numeracy skill was on average 0.2 standard deviations lower compared to the old counterparts. On the other hand, the numeracy skill of high ability students increased by 0.2 standard deviations compared to cohorts educated in the old system. Over all, 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 an approximately 5% to 10% wage penalty within educational level for people who have graduated from high school under the new system.

Since the effects of the orientation and the performance mechanisms cannot be separately identified in a reduced-form framework, we construct and estimate a dynamic discrete choice model. Given the estimation result, a counterfactual analysis is conducted to measure the relative importance of each channel. We find that the orientation effect account for between 11% and 20% of the decrease in wage premia across educational attainment categories, while the remainder is explained by the performance effect. In terms of wage rates, the orientation effect accounts for 19% to 78% across the different educational categories, with the importance 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 and especially from high school towards 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 upon entering the labor market, ability types are fully revealed to employers. 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.

2.10 Appendix

2.10.1 Estimating misreport probability in educational attainment variable in LFS

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.

2.10.2 Correcting measurement error in regression result in LFS

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

$$Y = X\beta_0 + \epsilon \tag{2.13}$$

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

$$W = X + U \tag{2.14}$$

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

$$Y = W\beta_0 + \epsilon - U\beta_0 \tag{2.15}$$

Then the OLS estimator with observables is,

$$\hat{\beta}_{OLS} = (I - (W'W)^{-1}W'U)\beta_0 + (W'W)^{-1}W'\epsilon \tag{2.16}$$

Assuming that ϵ is independent of X and U, then

$$plim(\hat{\beta}_{OLS}) = (I - \Omega)\beta_0 \tag{2.17}$$

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

$$\hat{\beta}_c = (I - \hat{\Omega})^{-1}\widehat{\beta_{OLS}} \tag{2.18}$$

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

2.10.3 Robustness check

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

Common trend assumption

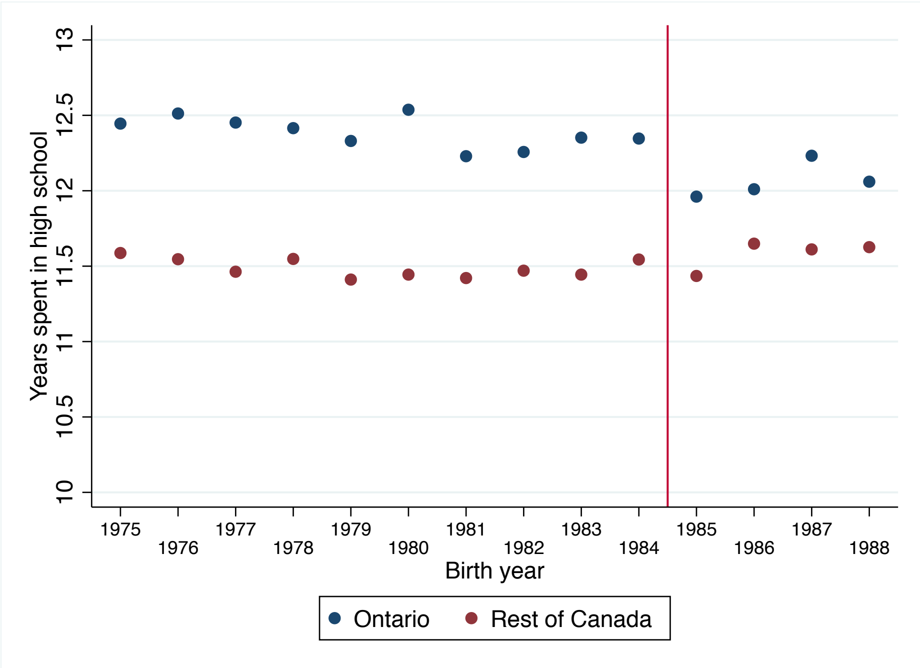
First of all, we investigate the validity of using a difference-in-difference estimation strategy. In particular, I compare the evolution of various outcomes, including high school duration, high school graduation rate, their post-secondary education attendance and their educational attainment, of the treated group (i.e. individuals from Ontario) with that of the control group (i.e. individuals from the rest of Canada) over cohorts. Figure 2.12, Figure 2.13, Figure 2.14 and Figure 2.15 provide evidence that supports the "common trend" assumption.

A more rigorous analysis is done by performing the following regression

$$Y_{ijk} = Cons + Ontario_{ijk} \cdot Cohort_j \theta_j + \alpha_j + \beta_k + X'_{ijk}\gamma + \varepsilon_{ijk} \tag{2.19}$$

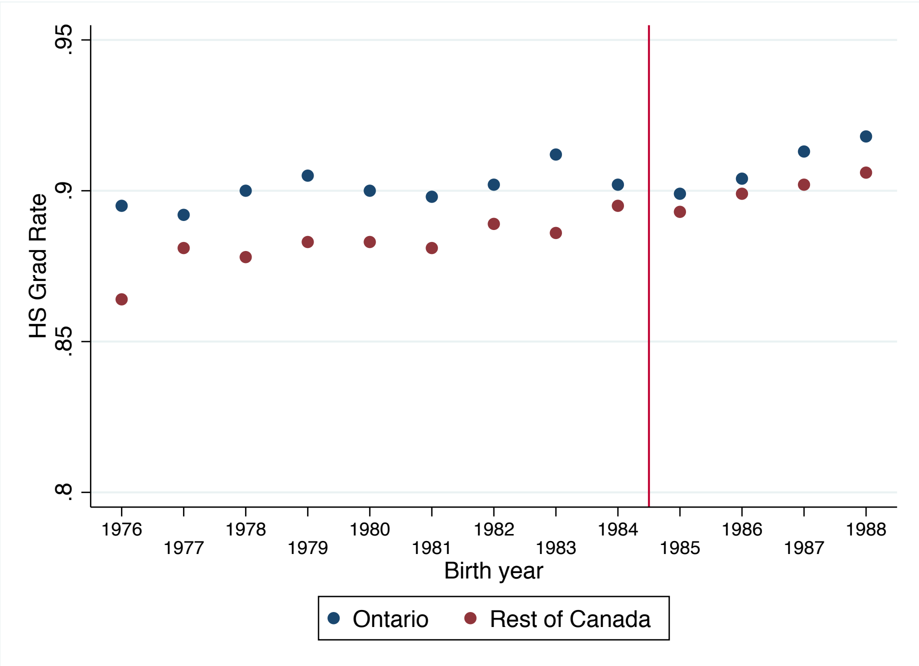
where the Y_{ijk} is the outcome variable, $Ontario_{ijk}$ is a dummy variable that equals to 1 if individual i is from Ontario and 0 otherwise, $Cohort_j$ is the cohort dummy. Other controls include cohort fixed effect α_j, province fixed effect β_k and other individual-specific characters X_{ijk}. 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. Table2.22, Table2.23, Table2.24 and Table 2.25 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 2.24), the main regression results were biased towards zero.

Figure 2.12: Trends in years spent in elementary and high school



Source: SLID 1993-2011. Sample includes all the individuals born between 1975 and 1988, with age $22-24$ and who were not attending high school in ref year. Quebec is excluded from the sample due to its unique educational structure.

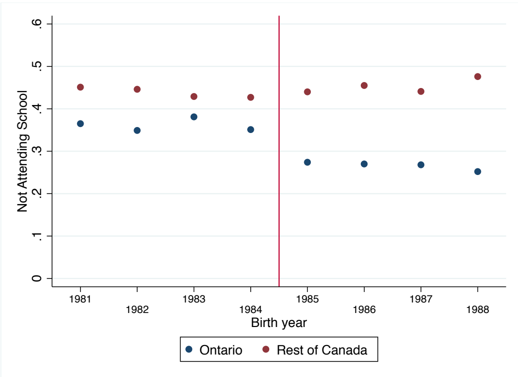
Figure 2.13: Trends in high school graduation rate



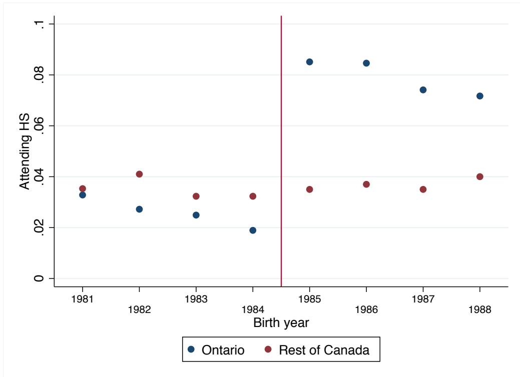
Source : LFS 2000-2017. Sample includes all the individuals born between 1976 and 1988, with age $22-24$ and who were not attending high school in ref month. Quebec is excluded from the sample due to its unique educational structure.

Figure 2.14: Trends in post-secondary education choice

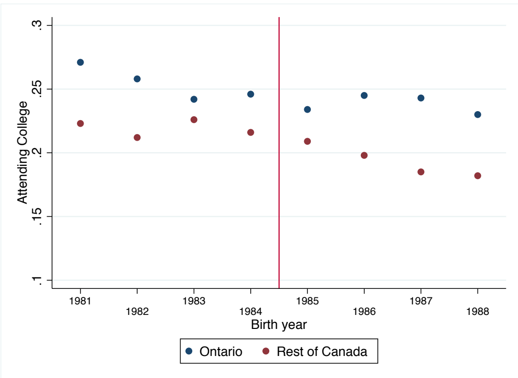
(a) Not in school



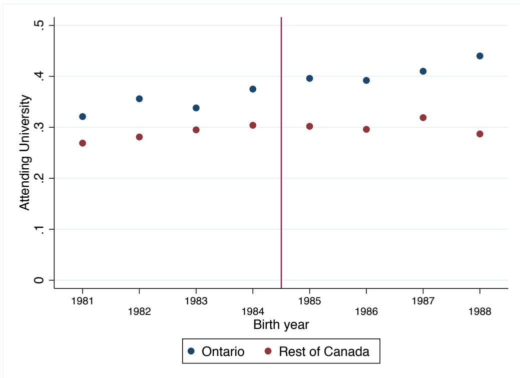
(b) Attending high school



(c) Attending college



(d) Attending university

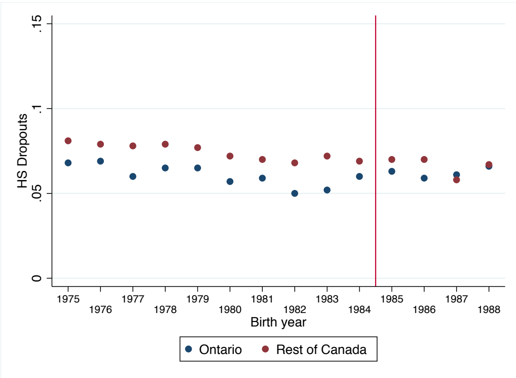


Source: LFS 2000-2017. Sample includes all the individuals born between 1981 and 1988, at age 18-21, who have graduated from high school. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure.

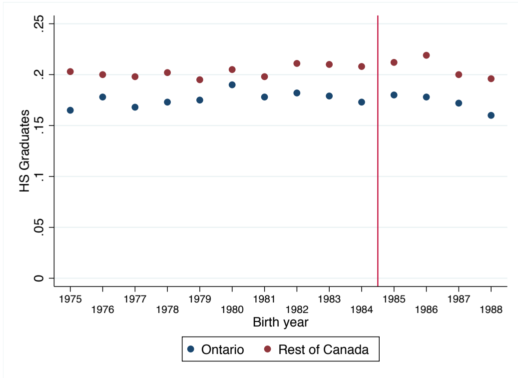
2.10 Appendix

Figure 2.15: Trends in educational attainment

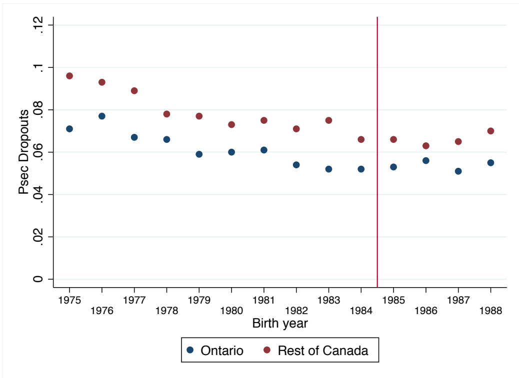
(a) HS dropouts



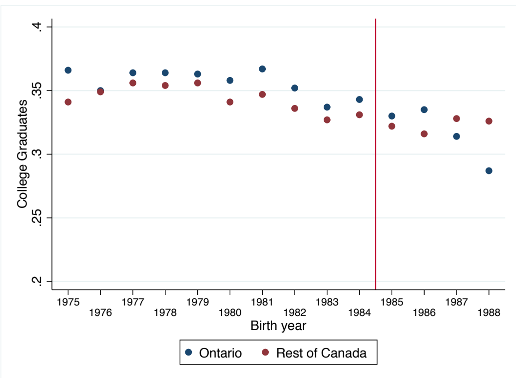
(b) HS grads



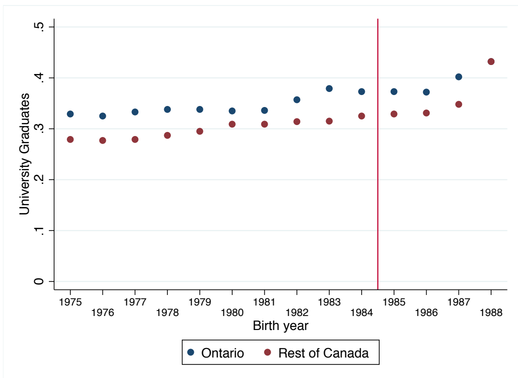
(c) Some psec



(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.

Table 2.22: 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 2.23: 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 2.24: Common trend: psec activities
HS grad age+1

	(1) Not in School	(2) School Attendance High school	(3) College	(4) 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 2.25: Common trend: educational attainment at age $28 - 36$

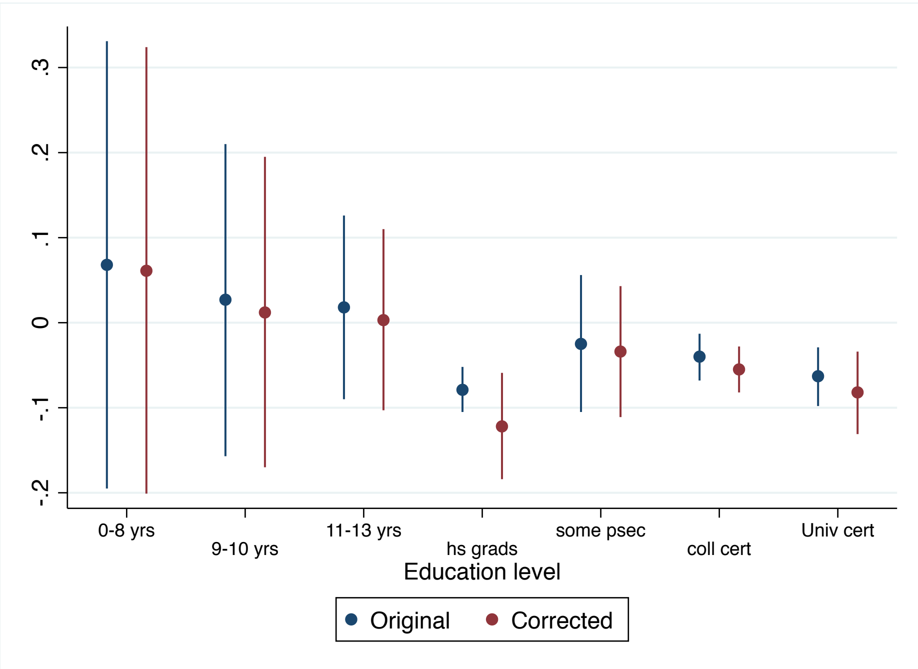
	(1) HS dropouts	(2) HS grads	(3) Some psec	(4) College	(5) University
$I_{1976} \cdot Ontario$	0.003	0.002	0.015	-0.011	-0.009
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1977} \cdot Ontario$	0.005	-0.011	0.010	-0.004	-0.001
	(0.013)	(0.020)	(0.012)	(0.026)	(0.026)
$I_{1978} \cdot Ontario$	-0.002	-0.010	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.

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, 3 and 5 in Table 2.26, Table 2.27, Table 2.28 and Figure 2.16, the policy effect estimation remains of similar magnitude after removing the double cohorts.

Figure 2.16: 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.

Table 2.26: Robustness check: high school duration, graduation rate and log hourly wage rate

	(1)	(2)	(3)	(4)	(5)	(6)
	S_{hs}		HS grad		$log(w)$	
Policy	-0.447***	-0.381***	-0.015***	-0.012**	-0.052***	-0.038***
	(0.054)	(0.042)	(0.005)	(0.005)	(0.007)	(0.007)
Double cohort	NO		NO		NO	
U		Yes		Yes		Yes
N	21280	20644	174000	208000	142000	203000
R^2	0.082	0.078	0.069	0.069	0.05	0.05

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable for first two columns is years spent in elementary and high school. Source: SLID 1993-2011. Sample includes all the individuals born between 1975 and 1988, with age $22 - 24$ and who were not attending high school in ref year. Other control variables include gender, immigrant status, year, cohort, province of high school fixed effect. Dependent variable for columns 3 and 4 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 2.27: 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 $\times$ birth year.

Table 2.28: 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.

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 condition[37]. As can be seen in columns 2, 4 and 6 in Table 2.26, Table 2.29 as well as Figure 2.17, 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.

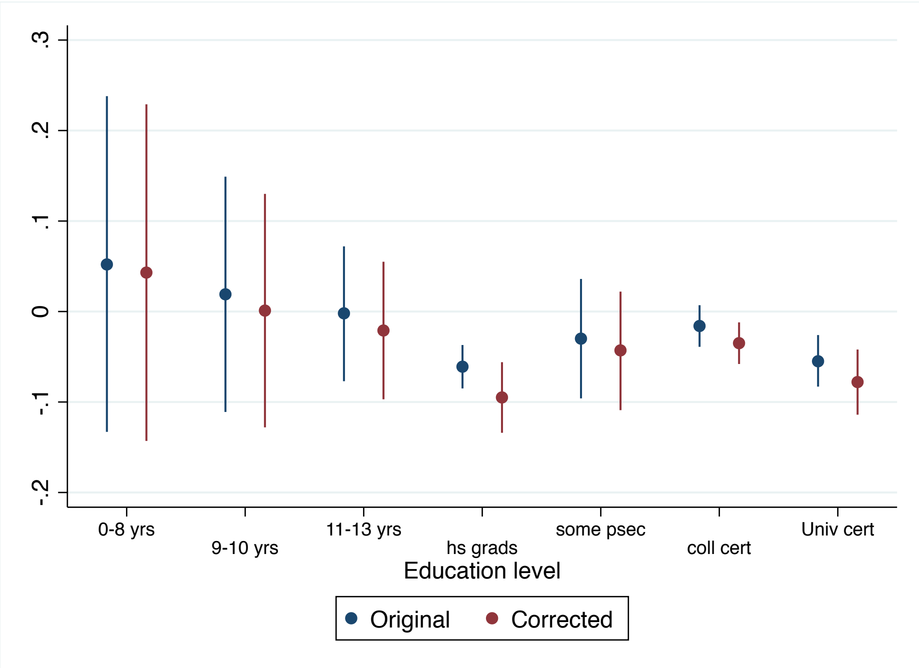
[37] We use provincial unemployment in the year when individuals graduate from high school when analyzing high school graduation rate and their post-secondary education enrollment decision

Table 2.29: 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.

Figure 2.17: 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.

3

The Misfortune of One's Birth: Cohort Size and Youth Earnings

3.1 Introduction

The relationship between the relative cohort size and earnings has been investigated in various countries (Brunello, 2010; Korenman and Neumark, 2000). Overall, the findings support the hypothesis that individuals born in relatively large cohorts face depressed real earnings (Brunello, 2010). Most of these studies use the long-run demographic changes as source of variation in the cohort size. One of the biggest challenges to this identification strategy is that it is not possible to fully rule out the fact that it is other unobserved variables other than demographic changes that drive the wage trends (Morin, 2015).

Starting in 1999, the Ontario government implemented a reform that eliminated the final year of high school while leaving the graduation requirements unchanged. The implementation of the reform generated the famous Ontario double cohorts: the last 5-year cohort and the first 4-year cohort. Both cohorts graduated simultaneously in year 2003, with the number of high school graduates increasing by more than 30% comparing to 2001 (Morin, 2015). This, however, is not the end of the story. Four years later, when most of the double cohort students graduated from university, the number of new university graduates[1] in Ontario experienced another spike. It increased from fifty-nine thousand in 2005

[1]Graduates with a Bachelor's degree

to seventy-five thousand in 2007. This 28% increase in Ontario due to the double cohort contrasts with the rest of Canada, where the number of new graduates increased only by 0.86% from 2005 to 2007. This exogenous increase in the number of new graduates with a Bachelor's degree enables me to analyze the causal impact of the relative cohort size on their earnings[2]. Moreover, an understanding of possible mechanisms leading to these effects and an exploration of whether these initial effects will persist in the long run serve as a key step in devising policy options to assist these "unlucky" cohorts.

There are two data sources which supports the empirical analysis. The main estimation relies on the Canadian Labour Force Survey (LFS), from 2004 to 2019. The LFS contains rich information on individuals' labor market outcomes, their hourly wage rate, the job characteristics etc. As a survey performed monthly, LFS could provide contemporaneous analysis of the cohort size effect, minimizing the contamination of other unobserved shocks in the labor market that might also affect recent graduates' earnings. The second data source used is the Canadian Census 2006, 2011 (National Household Survey 2011) and 2016. As one of the largest datasets available in Canada, the long-form questionnaire of the Census data includes rich information on individuals' educational attainment, their earnings, the job characteristics etc. Thus, the Census data could provide evidence to complement the main findings using the LFS. Moreover, the residence history information contained in the Census also helps provide a series of robustness checks to further consolidate the main findings.

Consistent with the existing literature, my paper supports the hypothesis that increased cohort size depresses their earnings. The LFS estimation suggests that Ontario double cohort decreased the recent university graduates' (log) hourly wage rate by 8.7% to 13.5%. The estimated result is robust across different control-group choices. Furthermore, these findings are corroborated by the Census data: the estimation result indicates their weekly wage rate decreased by 5.6%, when measured three years after their labor market entrance.

The adverse impact of the Ontario double cohort did not affect the entire wage distribution uniformly. Only the bottom and top quartiles were affected negatively. Moreover, I find that the increased cohort size decreased the proportion of recent university graduates taking

[2] Morin (2015) has already estimated that the sudden increase in high school graduates depressed their weekly earnings by $5\% - 9\%$.

a full-time job, increased the likelihood of them working for a "small" firm, and landing on a low-paid occupation and industry. All these channels help explain the observed wage decrease among the Ontario recent university graduates following the double cohorts. Finally, the depressing effect on wage rate persists several years after their graduation. The initial disadvantage of working for a "small" firm, choosing a "bad" occupation and industry is persistent, suggesting a slower recovery among those who are on the bottom tail of the wage distribution.

The rest of the paper is structured as follows. Section 3.2 reviews related literature; Section 3.3 provides additional information regarding the Ontarian educational system and the policy change in 1999, Section 3.4 and 3.5 summarize the data and the empirical results. Finally, Section 3.6 concludes.

3.2 Literature Review

This paper adds to the literature studying the impact of cohort size on labor market outcomes. The entrance of the baby-boomers into the labor market during the 1970s (Welch, 1979) initiated an extensive empirical literature on how the size of a cohort entering the labor market is related to their labor market outcomes. In this literature, a cohort normally refers to a group of individuals with similar age at given time and education attainment in some cases. Broadly speaking, the empirical findings from various countries and different time periods support the hypothesis that individuals born in large cohorts face depressed real earnings (Welch, 1979; Berger, 1985; Wright, 1991; Brunello, 2010; Korenman and Neumark, 2000; Jimeno and Rodriguez-Palenzuela, 2002; Moffat and Roth, 2016a; Card and Lemieux, 2001). When it comes to the effect on employment and unemployment, however, the literature provides conflicting evidence (Korenman and Neumark, 2000; Shimer, 2001; Jimeno and Rodriguez-Palenzuela, 2002; Biagi and Lucifora, 2005; Skans, 2005; Garloff et al., 2013; Moffat and Roth, 2016b; Roth, 2017).

Most of the aforementioned studies identify the cohort size effect using the long run demographic changes. One potential problem with this identification strategy is that it is hard to isolate cohort size effects from other unobserved trends that affect wages but are unrelated to demographic changes, especially when it comes to long-run variations (Morin, 2015).

The fundamental difference between this strand of literature and my paper is the definition of the cohort. In my paper, the term cohort is defined by the time of labor market entrance conditional on educational level, rather by the individual's birth year. The educational reform created the scenario where Ontario cohorts born in two consecutive years (cohorts born in 1984 and 1985) graduated from the university simultaneously in 2007. Without relying on long-run variations, the potential confounding effect due to unobserved demand trend is minimized in our paper. On top of that, my paper is better at identifying the congestion due to increased search competition in the labor market as opposed to the substitution in the aggregate production function as the labor supply shock I use happened among a precise cohort.

Sharing a similar nature in the definition of "cohort", my paper is also related to the literature studying the impact of immigration shocks on local labor market outcomes (Card, 1990; Hunt, 1992; Friedberg, 2001; Angrist and Kugler, 2003; Borjas and Monras, 2017). Despite the differences in the political causes that motivates the immigration flows, the economic conditions in the hosting countries' labor market, the skill level and size of the immigrants, Borjas and Monras (2017) concludes that exogenous supply shock adversely affects labor market opportunities of competing natives in the receiving countries and induces favorable impact on complementary workers.

Though helpful in understanding the effect of immigrants on the natives' labor market outcomes, these studies can only shed limited light on the potential effect of labor supply increases among local workers, as is the Ontario case. After all, part of the training that immigrants acquired prior to their arrival might be country-of-origin-specific, reducing the stock of human capital that is marketable in the post-migration period. As a result, immigrants and native worker are poor substitutes (Borjas and Monras, 2017).

Thus comparing to the immigration shock, the Ontario double cohort supply shock has the following advantages. The first one is that the excess labor supply is composed of workers who are almost identical to those who are already in the labor market, or "native" workers as referred in the immigration literature[3]. More specifically, there is no need to worry about the bias due to the "downgrading of immigrant skills" as did in the immigration literature. Another advantage of using the Ontario double cohort as labor supply shock is that the effect will not be contaminated by the simultaneous demand shock. The double cohorts were already living in Ontario before graduation, thus their entrance into the labor force had a lesser impact on the demand side comparing to the arrival of new immigrants.

While the state of the economy at labor market entrance condition plays an important role, the supply of similar qualified individuals is also an important entry condition. Morin (2015) has analyzed the labor supply shock on the earnings of young high school graduates. He concludes that the supply shock depressed young high school graduates' weekly earnings by 5% to 9%. Although I explore the labor supply shock due to the same policy change, our papers are different in the following aspects. To start with, I focus on individuals with different education levels. While Morin (2015) restricts his sample to high school graduates, I mainly investigate the labor supply increase among university graduates. Secondly, the time-window of my main analysis is the first year of their career. Instead, Morin (2015) investigated the difference in weekly wages two years after graduating from high school. Last but not least, the data sources that we use to draw the main conclusion are different. Our main analysis uses the LFS data, but Morin (2015) relies on the Census data instead. All these three aspects help explain why my estimation of the wage decrease among recent university graduates is larger than what Morin (2015) has documented.

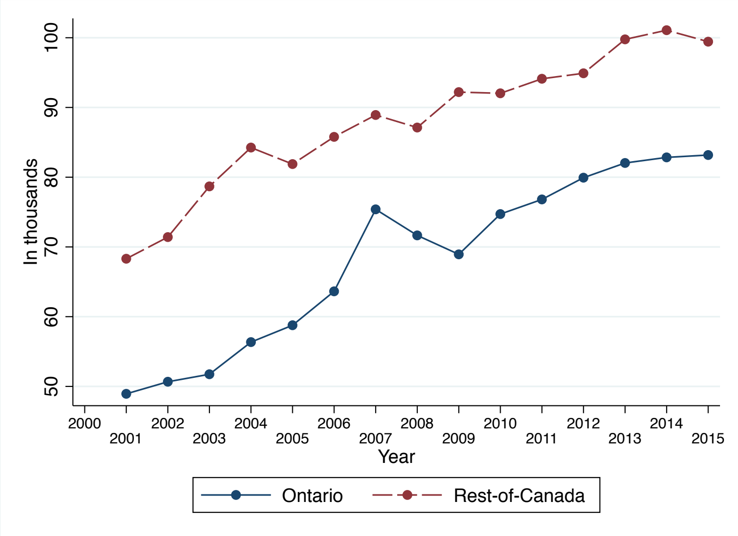
3.3 The Ontario double cohort and labor supply

Starting in 1999, the Ontario provincial government eliminated the fifth year of secondary education and mandated a new four-year system without changing the graduation requirement. As a result, in 2003, both the first 4-year cohort from the new curriculum and the last 5-year cohort who were educated in the old curriculum graduated from high school at the same time. The large cohort of high school graduates in 2003 is known as Ontario's double cohort. According to Morin (2015), the number of new high school graduates jumped by

[3]The only difference is in the number of years students spent in school given the same diploma, thus I further perform robustness check by including only the last old cohorts in the estimation to rule out changes in human capital accumulation, and the results remain similar

34.1 percent (from 91 291 to 122 406) between 2001 and 2003 in Ontario, while the rest of Canada only experienced an increase by 0.6 percent over the same period. Not surprisingly, a drastic increase in the number of university graduates from Ontario occurred four years later in 2007, as shown in Figure 3.1.

Figure 3.1: Number of university graduates per year



Source: Stats Canada.

Figure 3.1 presents the new university graduates per year from Ontario and from the rest of Canada (RoC). From year 2005 to 2007, the number of new graduates in Ontario increased by 28.3 percent, in contrast to the 0.86 percent increase in the RoC over the same period. The drastic contrast in the growth rate in new university graduates between Ontario and the rest of Canada, together with a relatively stable economic condition across Canada over this period of time, makes it possible to identify the cohort size effect on recent university graduates' wages.

3.4 Data

This paper mainly analyzes data from the Canadian Labour Force Survey (LFS). LFS is an ongoing monthly survey, which is used by the Canadian government to produce unemployment as well as other standard labor market indicators such as the participation rate and employment rate. LFS data covering years $2004 - 2019$ are used in this project.

The variables that are essential to model identification include the province where the respondents reside and the year when they obtained their Bachelor's degree and entered the labor market.

Ideally, my treatment group should include individuals who entered Ontario Labor market with a Bachelor's degree in year 2007. Unfortunately, the exact years of schooling each individual receives is not available in LFS, nor the labor market entrance year. Instead, I approximate their labor market entrance year based on the individuals' birth year and their highest educational attainment. Specifically, I assume that it takes four years to obtain a Bachelor's degree in any university programs and individuals who live in Ontario at survey time also went to high school in Ontario. Together with the fact that students start schooling at age six in Canada and the high school length in different provinces and different time, my treatment group is comprised of individuals born in 1984 and 1985 whose highest educational attainment is with a Bachelor's degree and who live in Ontario at survey time.

The main outcome variable of interest is the log of hourly wage rate. The hourly wage rate is adjusted using the provincial consumer price indices and any value below 2\$ is excluded from the sample. Other related labor market outcomes include the employment status, the full-time or part-time indicator of current job, occupation, industry and firm characteristics.

In order to exclude students who were also working at the same time, May to August survey data are not included in the analysis[4]. The LFS also contains personal characteristics, including gender, visible minority status, which enters the empirical analysis to add further precision to the estimation.

There are several advantages in using LFS as the main data source. First, since the main effect of the double cohort is concentrated on a small group of employees in the local labor market, a sufficiently big sample size provided by LFS is crucial to draw reliable conclusions. Secondly, LFS makes contemporaneous analysis available thanks to its monthly

[4]In LFS, there is a question asking if the respondent is going to school in the reference month. The proportion of people who report not going to school peaks in May, June, July and August. This suggests the earning data in May, June, July and August include lots of temporary "summer job" cases, thus were excluded.

frequency. Unlike the Census that is conducted every five years, LFS makes it possible to analyze the effect at a more flexible time window. Last but not least, LFS contains very detailed information on individuals' labor market outcomes, which enables a comprehensive portrait of the cohort size effect and its potential underlying mechanisms.

Additionally, the Canadian Census data is used to complement LFS. The 2006 and 2011 long-form questionnaires[5] target approximately 20% of Canadian households, which serves as one of the largest datasets available to researchers. Although hourly wage rate information is not available in the Census, it does include employees' average weekly wage and other labor market conditions. In addition to province of residence, the Census also asks the respondents where they obtained their highest post-secondary degree. This variable helps narrow down the definition of our treatment group: individuals born in 1984 and 1985 who obtained their Bachelor's degree in Ontario and also live in Ontario at the survey time.

One potential issue with the Census data is that the compulsory long-form questionnaire of the Census was abolished in 2010 and replaced by the NHS. Though the variables and the sample size covered in NHS is comparable to those in the previous Census, the new non-compulsory nature of the NHS makes it unclear how one should interpret the regression result combining two different sources of data. In fact, there are voices raised against the abolition of the long-form questionnaire arguing that moving to a voluntarily survey would ruin the representativity of the data (Morin, 2015). Thus the results obtained using the Census data should be interpreted with caution.

3.5 Empirical Analysis

The identification strategy is triple difference-in-difference, which is similar to that used in Morin (2015) except that the treatment group in my paper focuses on the Ontario double cohort with a Bachelor's degree[6]. By comparing the pre- and post- wage gap between recent university graduates and those with more working experience in Ontario versus in the rest of Canada, the effect of the excess labor supply shock is identified. More specifically, I

[5] The long-form questionnaire in Census 2011 is called 2011 National Household Survey (NHS)

[6] There is no clear jump for college graduates between 2003 and 2007. This could be due to the large variation in duration for college programs, which smoothed the number of new entrants into the labor market over this period of time

estimate the following regression:

$$Y_{ixpt} = \phi_{xt} + \lambda_{xp} + \eta_{pt} + \beta(\text{New Grads}_x \times \text{Ontario}_p \times \text{Post}_t) + X_{ixpt}\gamma + \epsilon_{ixpt} \quad (3.1)$$

here individuals are indexed by i, x indicates experience group, p stands for province and t means time. The outcome variable Y_{ixpt} stands for individual i's log hourly wage rate. Variable Ontario equals 1 if the individuals resides in Ontario, New Grads_x is a dummy variable equals 1 if the individual is a new university graduate in year t, and 0 if he/she graduated earlier. Dummy variables Post_t is equal to 1 if the outcome is observed after the double cohorts' entrance. Therefore, the triple interaction term New $\text{Grads}_x \times \text{Ontario}_p \times \text{Post}_t$ represents the treatment group: the Ontario double cohorts who graduated from university in previous years. The coefficient β captures the impact of cohort size on the youth log wage rate by comparing the provincial differences (Ontario versus the rest of Canada) in the changes in relative wages of recent graduates compared to seasoned workers with the same education level over time (Pre- and post the double cohort's entrance).

Two types of demand shocks are captured by ϕ_{xt} and η_{pt}. ϕ_{xt} controls the experience-related demand shocks across Canada, and η_{pt} controls the provincial demand trend. λ_{xp} allows the wage difference to vary across experience groups and provinces. Finally, X_{ixpt} further controls individual specific characteristics, including gender, race, the working industry.

As pointed out in Morin (2015), the ideal choice of the reference group to calculate the wage gap would include workers who were affected by similar demand shocks as to the recent graduates, while not being affected by the excess labor supply. Thus, I consider workers with the same education level as the treatment group but with $5-7$ years of working experience and from the same province as the reference level. The basic assumption is that workers with $5-7$ years of working experience are not likely to be substitutes to the recent graduates, but would still be affected by labor demand shocks in a similar fashion. Later on, I relax this assumption and conduct similar regressions with workers with various working experience as a robustness check. In the main empirical analysis, the treatment group includes the Ontario double cohort who graduated from the university in 2007, and the control group is composed of individuals with a bachelor's degree who entered the labor market in 2007 from other provinces, as well as all the university graduates with $5-7$ years working experience when surveyed from 2004 to 2008.

3.5.1 Effect on log wage rate

Table 3.1: Double cohort and wage rate, LFS

	(1)	(2)	(3)
		log hourly wage	
New grads × Ontario × Post	-0.087**	-0.094***	-0.135***
	(0.035)	(0.034)	(0.033)
New grads	-0.687***	-0.450***	-0.367***
	(0.028)	(0.028)	(0.131)
Industry fixed effects	N	Y	Y
Industry × New grads fixed effects	N	N	Y
Industry × Post fixed effects	N	N	Y
N	17316	17316	17316
R^2	0.131	0.312	0.338

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables is log hourly wage rate. Source: LFS 2004-2008. Sample includes all employees who have a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Table 3.1 presents the estimation of Eq. 3.1. All the columns 1 to 3 include saturated interaction terms between year, province and experience group fixed effects, namely ϕ_{xt} , η_{pt}, λ_{xp} and other personal characteristics. Furthermore, column 2 includes industry fixed effects, column 3 includes both the industry fixed effects and its interaction with experience group dummy and that with Post-2007 dummy. Thus, in the last specification, demand shocks that are industry specific, province specific, and experience level specific could not bias the estimation of cohort size effect. Though these interaction terms could con-

Table 3.2: Double cohort and wage rate, Census

	(1)	(2)	(3)
		log weekly wage	
New grads × Ontario × Post	-0.056***	-0.039***	-0.037***
	(0.008)	(0.006)	(0.006)
New grads	-0.266***	-0.213***	-0.054***
	(0.012)	(0.005)	(0.063)
Personal Characteristics	Y	Y	Y
Industry fixed effects	N	Y	Y
Industry × New grads fixed effects	N	N	Y
Industry × Post fixed effects	N	N	Y
N	55882	55882	55882
R^2	0.126	0.213	0.217

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables is log weekly wage. Source: Census 2006 and NHS 2011. Sample includes all employees who have a bachelor's degree as their highest education level. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and immigrants fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

trol industry-specific demand shocks, they could induce bias in the estimation because the industry choice itself is endogenously affected at the same time.

As shown in the table, the Ontario double cohort university graduates earn on average 8.7% less than similar recent graduates who belong to a normal cohort measured in the first year of their career. Controlling industry specific effects further increases the estimated effect up to 13.5%, suggesting the negative effect captured in Model specification 1 is not purely driven by the composition change.

The estimated cohort size effect on log weekly wage is smaller yet still present when measured two years later using Census data, which is shown in Table 3.2. The difference in the magnitude of the estimation could be explained both by the difference in the outcome variables and in the time-window of the analysis.

Table 3.3 summarizes the estimated impact of increased cohort size using different control groups. The first two columns present benchmark result using university graduates with $5-7$ years of working experience to calculate the wage gap, and the following columns present the estimation results with different references groups: employees who have a Bachelor's degree and have $5-10$, $11-15$, and $16-20$ years of working experience respectively. It's clear that the estimated effect of cohort size on youth earnings is not purely driven by a specific choice of the reference group.

What I have found here is consistent with the findings documented in Morin (2015), where the author shows adverse effect on weekly wage ranging from 5% to 9%, when measured two years after labor market entrance.

3.5.2 Identification challenges

Consistent with the basic theoretical prediction, the regression results confirm that the supply shock due to the Ontario double cohort adversely affect recent university graduates' hourly wage rate. In this section, I further address a series of concerns that challenge our main identification method.

Table 3.3: Double cohort effect and different control groups

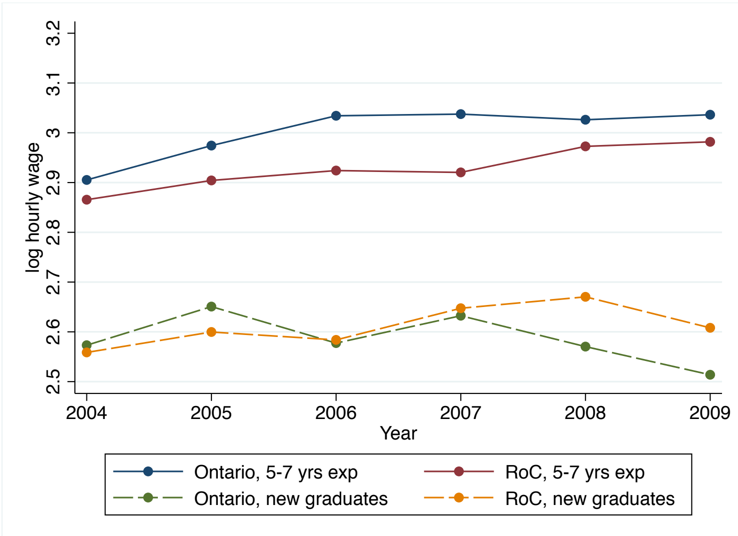
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp group	5 − 7		5 − 10		11 − 15		16 − 20	
New grads × Ontario × Post	-0.087**	-0.135***	-0.045	-0.083**	-0.059**	-0.061**	-0.125***	-0.118***
	(0.035)	(0.033)	(0.028)	(0.035)	(0.028)	(0.030)	(0.040)	(0.039)
Industry fixed effects	N	Y	N	Y	N	Y	N	Y
Industry × New grads fixed effects	N	Y	N	Y	N	Y	N	Y
Industry × Post fixed effects	N	Y	N	Y	N	Y	N	Y
N	17316	17316	32279	32279	26809	26809	25236	25236
R^2	0.131	0.338	0.104	0.289	0.176	0.339	0.185	0.352

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is log hourly wage rate. Source: LFS 2004-2008. Sample includes all employees who has a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Common trend assumption

First and foremost, I am going to present the validity of common trend assumption to justify the use of triple difference-in-difference estimation strategy. In the following figure, I depicted the evolution of log hourly wage rate of the treated and control group used in Table 3.1. More specifically, the figure shows the average log hourly wage rate of recent university graduates versus those with $5-7$ years of working experience at given year from Ontario versus from the rest of Canada.

Figure 3.2: Evolution of average hourly wage rate



Source: Labour Force Survey, $2004-2010$

It is obvious that the evolution of experienced workers' wage in Ontario mirrors those from the rest of Canada over 2004-2007. More interestingly, the parallel ceased after 2007, when the double cohort entered the labor market. Thus the estimated cohort size effect indeed captures the divergence in the wage-gap trend after 2007 of recent university graduates across Provinces.

Employment

When studying the cohort size impact on wage rate, one needs to keep in mind that only individuals who were working as employees when surveyed are qualified for the sample.

Thus the possibility that the status of being employed itself was affected by the supply shock would make the estimation result open to question. Therefore, I investigate whether the (un)employment status and the school attendance status for the first year after graduation is different among the Ontario double cohort. More specifically, I estimated Equation (3.1) using unemployment status and school attendance indicator as the dependent variable. The results is shown in Column 1 and 2 in Table 3.4

Table 3.4: Robustness check

	(1)	(2)	(3)	(4)	(5)
	U	Student	log hourly wage		Moving out
New grads × Ontario × Post	0.015	-0.005	-0.076**	-0.126***	-0.056
	(0.019)	(0.018)	(0.036)	(0.040)	(0.216)
New grads	0.326***	0.290***	-0.687***	-0.649***	-0.072
	(0.102)	(0.031)	(0.028)	(0.126)	(0.084)
Industry fixed effects	Y	Y	N	Y	N
Industry × New grads fixed effects	Y	Y	N	Y	N
Industry × Post fixed effects	Y	Y	N	Y	N
1985− cohort	Y	Y	N	N	Y
N	19216	25652	17106	17106	51234
R^2	0.092	0.082	0.125	0.334	0.182

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is unemployment status, student status, log hourly wage rate, and the dummy variable that equals 1 if the individual's current province of residence is different from the province where they obtained the bachelor's degree. Source for columns 1-4: LFS 2004-2008. Sample includes all employees who have a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. In column 3 and 4, Ontario 1985-cohort is excluded from the sample. Other control variables include gender and month fixed effect. Source for the last column: Census 2006 and NHS 2011. Sample includes all employees who have a bachelor's degree as their highest education level. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender and immigrants fixed effect. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

The estimates once again confirm our conclusion: though recent university graduates are more likely to be unemployed or to go back school on average, this trend was not affected by the double cohorts' inflow into the labor market.

School duration

Another concern comes from the fact the education reform reduces the number of years Ontarian students need to spend in high school before going to university. This means the 1985-cohort from Ontario receive one year less schooling as university graduates compared to their 1984-counterparts. Moreover, the previous chapter has demonstrated that cohorts educated in the new regime face on average 5% to 10% wage penalty. Thus the estimated adverse impact on earnings among the double cohort might be purely driven by the wage decrease of the $1985-$cohort. In that case, the estimated cohort size effect should be attributed to the change in school duration rather than the excess labor supply.

To exclude this possibility, I conduct the same analysis but only include $1984-$cohort in my treated group. In this case, the only difference between the treatment group and control group is the cohort size. The result is shown in Column 3 and 4, Table 3.4. On average, the recent graduates born in 1984, Ontario earn 7.6% to 12.6% less compared to recent university graduates who belong to a normal cohort – almost the same as the bench mark analysis where the $1985-$cohort is also included in the treated group. In summary, the adverse effect of shortened high school duration as a result of the educational reform alone cannot explain the observed wage decrease in the first year of their career.

Out-migration from Ontario

It is possible that university graduates self-select to move out of Ontario to avoid the competition due to excess labor supply, which could bias the estimate. Luckily the Census data contains not only the current province of residence, but also the province where the respondent obtained their highest university degree. Thus, I create the indicator which is equal to 1 if the respondent moved to a different province after university graduation and zero otherwise. Then I run Eq.(3.1) using the indicator of moving out as the outcome variable, and could not find evidence of moving out of Ontario in response to the entrance of double cohort graduates.

Multiple reference group choice

Table 3.5 helps justify the choice of university graduates with $5-7$ years of working experience as the reference group when calculating the wage gap in the main empirical analysis. Instead of dropping out all the university graduates with working experience neither one nor five to seven years, I conduct the following regression including all the university graduates with up to twenty years of working experience in the sample,

$$Y_{ixpt} = \phi_{xt} + \lambda_{xp} + \eta_{pt} + \beta_x(\text{Ontario}_p \times \text{Exp}_x \times \text{Post}_t) + X_{ixpt}\gamma + \epsilon_{ixpt} \tag{3.2}$$

where the working experience (indicated by subscript x) includes the following categories: recent university graduates with one year of working experience, university graduates with $2-4$, $5-7$, $8-10$, $11-15$ and $16-20$ years of working experience. Thus the coefficients β_x tells us how much the average wage of Ontarian university graduates with experience level x changed following the double cohort university graduates entering into the labor market.

The results in Table 3.5 proves that university graduates with $2-4$ years of working experience were negatively affected by the double cohorts' entrance as well, suggesting substitutability with recent university graduates. On the other hand, more experienced employees with the same education level were not affected differently from the base category, namely the university graduates with $5-7$ years of working experience. In short, university graduates with $5-7$ years of working experience serve as a reasonable reference group to estimate the double cohort effect in the triple difference-in-difference identification strategy.

3.5.3 Distributional impacts and possible channels

Distributional impacts

The aforementioned results estimate the average impact of the cohort size effect on youth earnings, which masks the potential heterogeneity of the effect across the wage distribution. In order to check the distributional impacts, I perform the quantile regressions based on Eq.(3.1) for all the quartiles of hourly wage distribution, and the result is presented in Table

Table 3.5: Multiple reference group choice

	(1)	(2)	(3)
		log hourly wage	
New grads × Ontario × Post	-0.082**	-0.077**	-0.108***
	(0.037)	(0.040)	(0.040)
$2-4$ yrs exp× Ontario × Post	-0.103***	-0.071***	-0.074***
	(0.029)	(0.027)	(0.028)
$8-10$ yrs exp× Ontario × Post	-0.050*	-0.031	-0.066**
	(0.029)	(0.028)	(0.029)
$11-15$ yrs exp× Ontario × Post	-0.014	0.002	-0.006
	(0.027)	(0.027)	(0.027)
$16-20$ yrs exp× Ontario × Post	0.040	0.041	0.011
	(0.038)	(0.036)	(0.034)
Industry fixed effects	N	Y	Y
Industry × New grads fixed effects	N	N	Y
Industry × Post fixed effects	N	N	Y
N	91445	91445	91445
R^2	0.152	0.293	0.354

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables is log hourly wage rate. University graduates with $5-7$ years of working experience is the base level. Source: LFS 2004-2008. Sample includes all employees who have a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × experience × year level.

3.6. Different columns use different quartile as the outcome variables. Table 3.6 suggests heterogeneous impact across the wage distribution, the adverse effect mainly happened at the bottom and the top quartile, with the median unaffected. I further explore potential sources of the heterogeneity in the impact of the double cohort in the next subsection.

Table 3.6: Quantile regression estimates for cohort size effect on wage rate

	(1) Bottom quartile	(2) Median	(3) Top quartile
New grads × Ontario × Post	-0.148***	-0.019	-0.129**
	(0.058)	(0.066)	(0.054)
N	17316	17316	17316

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is the log hourly wage rate. Source: LFS 2004-2008. Sample includes all employees who has a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Possible channels

I mainly focus on the following mechanisms that excess labor supply induces adverse effect on wage rates: the firm quality, the occupation and industry choice and the probability to find a full-time job.

When it comes to the firm quality, LFS asks the respondent the total number of employees their employer hire. Based on this information, I create the dummy "Small" firm that equals 1 if the total number of employees is below 20, and 0 otherwise. At the other end of the spectrum, an indicator "Big firm" is also created, which is equal to 1 if the firm has more than 500 employees and 0 otherwise. Table 3.7 presents the result of estimating Eq.(3.1) using "Small" firm and "Big" firm dummies as outcome variables. If we think firm size being an indicator of firm quality, the results are consistent with the quantile regression results. The excess supply induces more graduates to work for small, low productivity firms, and fewer of them manage to work in a high productivity firm.

Table 3.7: Possible channels: Firm size

	(1)	(2)	(3)	(4)
Firm size	"Small" firm		"Big" firm	
New grads × Ontario × Post	0.166***	0.137***	-0.060**	0.01
	(0.034)	(0.034)	(0.026)	(0.042)
Industry fixed effects	N	Y	N	Y
Industry × New grads fixed effects	N	Y	N	Y
Industry × Post fixed effects	N	Y	N	Y
N	17316	17316	17316	17316
R^2	0.019	0.189	0.023	0.236

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable include "Small" firm, which indicates if the firm has fewer than 20 employees in all locations (the first two columns) or not. "Big" firm indicates whether the employer has more than 500 employees in all locations (the last two columns). Source: LFS 2004-2008. Sample includes all employees who have a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Other than the firm quality, I investigate if recent university graduates' occupation and industry choice is affected. First, I rank the occupations based on the average wage measured in $2004-2006$ among all the university graduates who work full-time, then a dummy variable is constructed which equals 1 if the occupation profile is below the bottom quartile of the wage distribution, and zero otherwise.

As Table 3.8 suggests, the Ontario double cohort is 15.8% more likely to land in an occupation which offers wage below the median. The estimated effect is larger within the industry level. Different from the impact on firm size, the proportion of individuals working in a highly paid occupation is not affected over all, but a negative impact is documented once I control for industry-specific trends.

Apart from the occupation choice, I also examine if there is a change in the composition

Table 3.8: Possible channels: Occupation Choice

	(1)	(2)	(3)	(4)	(5)	(6)
	Bottom quartile		Below median		Top quartile	
New grads × Ontario × Post	-0.006	0.026	0.158***	0.213***	-0.051	-0.104***
	(0.050)	(0.050)	(0.045)	(0.036)	(0.036)	(0.034)
Industry fixed effects	N	Y	N	Y	N	Y
Industry × New grads fixed effects	N	Y	N	Y	N	Y
Industry × Post fixed effects	N	Y	N	Y	N	Y
N	17316	17316	17316	17316	17316	17316
R^2	0.095	0.284	0.084	0.251	0.080	0.203

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables include the indicator if the occupation is below the bottom quartile of the wage distribution (the first two columns), the indicator if the occupation is below median (columns 3 and 4) or in the top quartile of the wage distribution (the last two columns). Source: LFS 2004-2008. Sample includes all employees who has a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

of industries where recent university graduates work as well. Similarly, I rank the industry based on the average wage measured in $2004 - 2006$ among all the university graduates, then a dummy variable is constructed, which equals 1 if the industry is below the bottom quartile of the wage distribution, and zero otherwise. Another dummy variable indicating if the industry belongs to the top quartile is also constructed. Table 3.9 shows the excess supply because of the Ontario double cohort mainly pushes more individuals towards low-paid industries.

Table 3.9: Possible channels: Industry Choice

Industry	(1) Bottom quartile	(2) Top quartile
New grads × Ontario × Post	0.098**	-0.028
	(0.047)	(0.034)
N	17316	17316
R^2	0.090	0.082

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables include the indicator if the industry is below the bottom quartile of the wage distribution, and if the industry is in the top quartile of the wage distribution. Source: LFS 2004-2008. Sample includes all employees who has a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Last but not least, how the possibility of finding a full-time job after graduation was affected by the cohort size is also investigated as shown in Table 3.10. The LFS data confirm that fewer recent Ontario university graduates belonging to the double cohort were able to find a full-time job in the first year of their career. The disadvantage is even present two years later, as shown in Column 3 using the Census data.

In summary, it seems that part of the estimated negative impact raised by the Ontario double cohort is attributed a change in the distribution of firm size, occupation types, indus-

Table 3.10: Possible channels: Full-time

	(1)	(2) Full-time	(3)
New grads × Ontario × Post	-0.075**	-0.099***	-0.017***
	(0.029)	(0.030)	(0.004)
Industry fixed effects	N	Y	N
Industry × New grads fixed effects	N	Y	N
Industry × Post fixed effects	N	Y	N
Source	LFS	LFS	Census
N	17316	17316	56999
R^2	0.037	0.132	0.035

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables include the indicator if working as full-time. Source for column 1 and 2: LFS 2004-2008. Sample includes all employees who has a bachelor's degree. 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 and month fixed effect. Source of Column : Census 2006 and NHS 2011. Sample includes all employees who has a bachelor's degree as their highest education level. Quebec is excluded from the sample due to its unique educational structure. Other control variables include gender and immigrants fixed effect. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

try choices, and the job type that recent university graduates could obtain for their first year of labor market experience. In addition, the effect of increased labor supply is not uniform, individuals at the bottom wage distribution were doing worse, while those who are at the top distribution were not affected as much.

3.5.4 Spillover effects

In this section, I explore possible spillover effects from the double cohort on individuals with different education levels. In particular, one would anticipate recent college graduates or even recent high school graduates were also affected by the extra influx of new university graduates.

Thus I estimate Eq.(3.1) but on high school graduates and college graduates separately[7]. Results shown in Table 3.11 suggest the presence of the spillover effects among high school graduates. Specifically, the impact on college graduates is not significantly different from zero, while high school graduates experience a wage increase. The positive effect on high school graduates' earnings implies complementarity between high school graduates and university graduates in the Ontario labor market.

However, one needs to interpret the positive effect among high school graduates with a grain of salt. The Ontario double cohort graduated from high school in 2003. Though the data in 2003 is not included in the sample, the recent high school graduates following the double cohort could still be affected by the excess labor supply shock. If that is true, the positive estimates could simply be driven by the recovery of high school graduates' wage since 2003, rather than the influx of extra university graduates.

3.5.5 Persistence

Last but not least, I investigate whether these adverse effects on Ontario double cohorts' wage rate persist over time. In Figure 3.3, I present the evolution of the estimated effect over the following ten years. That is to say, I study how the wage gap between cohorts who graduated from universities in 2007, and those who completed their university studies $5-7$ years in advance evolves over time. For example, I compare the wage gap between

[7]Post-secondary dropouts were not included because it is difficult to estimate their working experience.

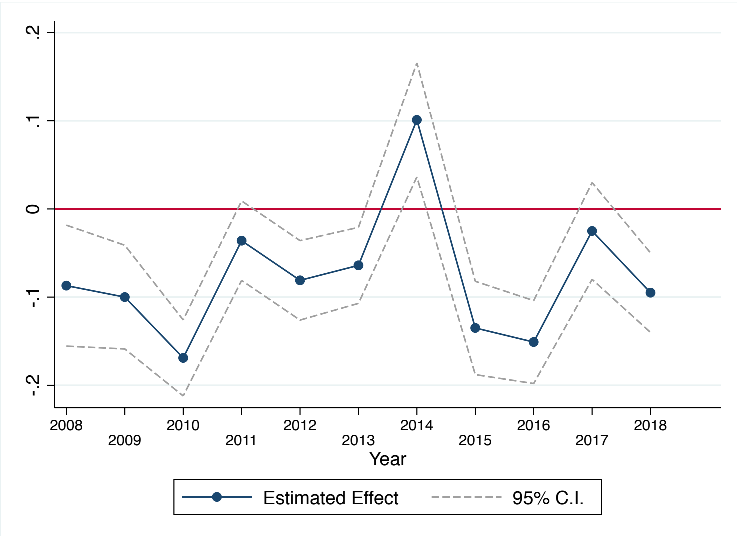
Table 3.11: Spillover effects on high school graduates and college graduates

log wage	(1)	(2)	(3)	(4)
	HS grads		Coll grads	
New grads × Ontario × Post	0.080***	0.071***	-0.027	-0.043
	(0.022)	(0.020)	(0.068)	(0.056)
Industry fixed effects	N	Y	N	Y
Industry × New grads fixed effects	N	Y	N	Y
Industry × Post fixed effects	N	Y	N	Y
N	26539	26539	8909	8909
R^2	0.281	0.451	0.248	0.448

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables is log hourly wage rate. Source: LFS 2004-2008. Sample includes all employees whose highest education level is high school diploma (first two columns) and college degree (Column 3 and 4). May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

the double cohort (who have 2 years of working experience) and seasoned workers with $6-8$ years of working experience cross provinces (Ontario versus the rest of Canada) and across time (before and after 2007) to obtain the estimates for year 2009.

Figure 3.3: Persistence on log wage rate



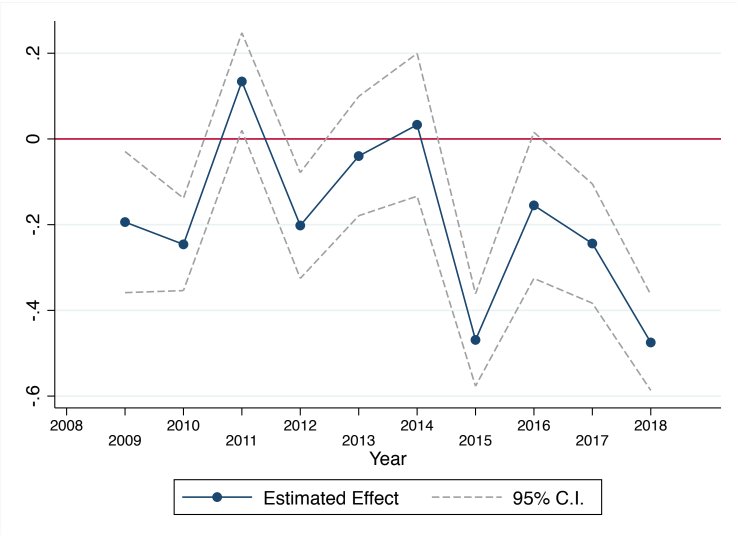
I would like to emphasize that the persistence effect is descriptive since the estimates are increasingly likely to be contaminated by unobserved factors other than cohort size as the analysis window moves forward. In general, the cohort size effect on wages rate remains negative through the first decade after their university graduation with some exceptions.

Over all, the data suggests that the adverse effect of a larger cohort size on their earnings do not disappear for the first ten years after their labor market entrance. The negative effect documented here is comparable to the negative effect of graduating from the university during recession (Altonji et al., 2016; Oreopoulos et al., 2012). Recessions lead workers to start their careers with less attractive employers, and then the less advantaged graduates move towards a better job at a much slower speed, if they recover at all (Oreopoulos et al., 2012).

In order to investigate possible mechanisms of the persistent negative effect on wage rate, I conducted similar analysis on a bunch of labor market outcomes: the job tenure, the firm size, the industry and the occupation profiles.

One of the channels that contribute to the long-lasting adverse effect of entering into labor market during recession is job mobility. Economic recessions at the labor market entrance lead workers to start with less attractive employers, and then the disadvantaged graduates move towards a better job at a much slower speeds, if they recover at all (Oreopoulos et al., 2012). Unable to observe the job mobility directly, I use the job tenure as an indirect measure of how often the individuals change employers. Specifically, I investigated how the cohort size affected the job (log) tenure evolution over time as shown in Figure 3.4. When measured two and three years after their labor market entrance, the Ontario double cohort on average have shorter tenure on their current job when surveyed. Together with the fact that the employment status was not affected, this suggests more job mobility occurred among the Ontario double cohorts than graduates from a normal cohort. Four years after their labor market entrance, the effect on log tenure gradually diminished. Unfortunately LFS data does not provide enough information to further investigate whether the impact on tenure is uniform across individuals' skill or ability.

Figure 3.4: Persistence on log tenure



There is also evidence suggesting job mobility through the (change of) employer characteristics. I analyzed the evolution of the likelihood of individuals working for firms with fewer than 20 employees, and those whose firm hire more than 500 employees. It turned out that entering the labor market with a larger cohort increased the proportion of them working with very small firms for the first couple of years. The disadvantage gradually disappears through job mobility across firms as workers accumulate experience. At the other extreme, the disadvantage of not being hired by a very "big" firm is not as significant.

Figure 3.5: Persistence on firm size: "Small" firm

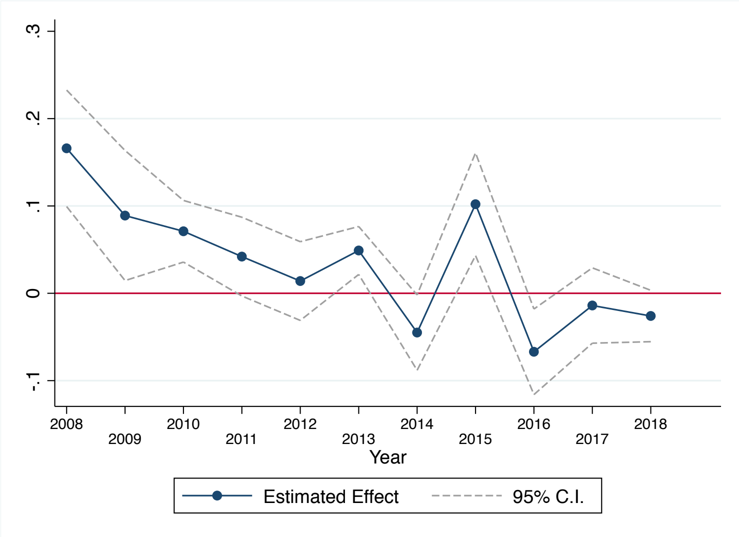
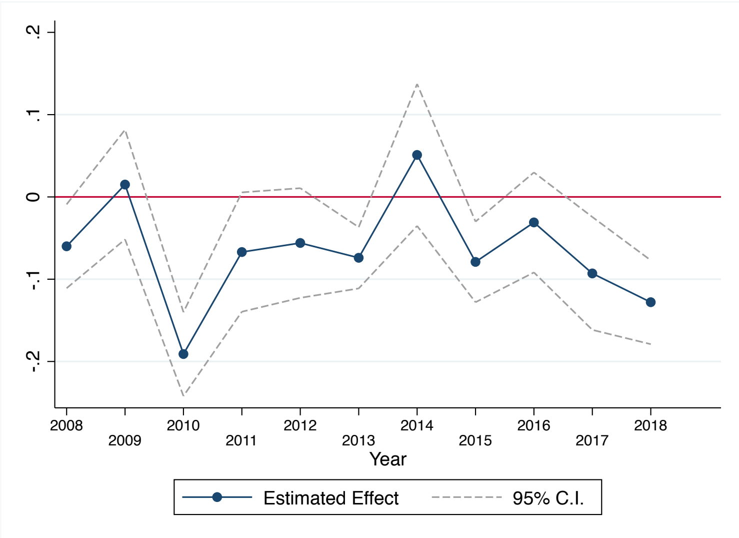


Figure 3.6: Persistence on firm size: "Big" firm



When it comes to both the occupation and industry profiles, entering the labor market with a larger cohort size induces heterogeneous effect along the wage distribution. A larger cohort size increases proportion of recent university working in a low-paid occupation and industry (i.e. an occuption or industry profile in the bottom quartile of the wage distribution). However, there is no clear indication that Ontario double cohorts were less likely to have a highly-paid job, nor work in a highly-paid industry. Thus, the impact of the Ontario double cohort was not uniformly distributed across the occupation and industry profiles.

Figure 3.7: Persistence on occupation choice: Bottom quartile

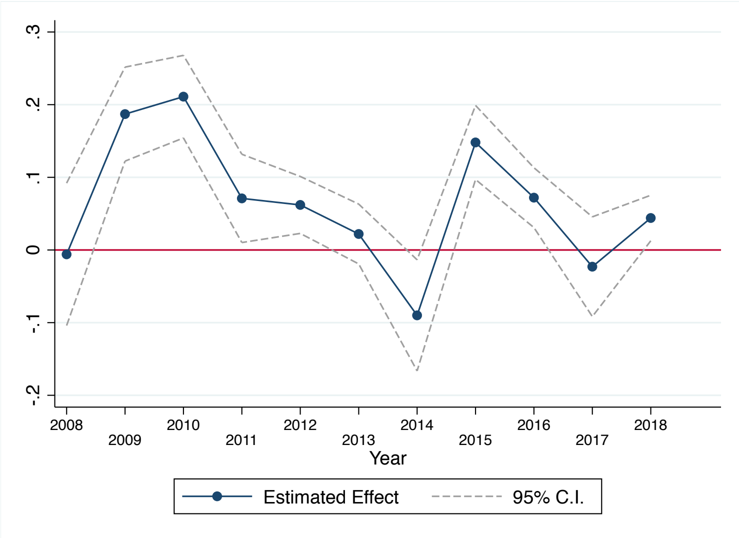


Figure 3.8: Persistence on occupation choice: Top quartile

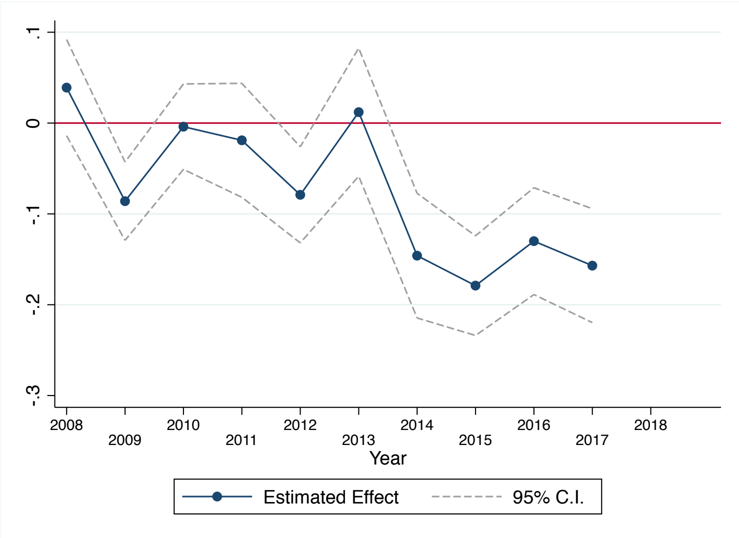


Figure 3.9: Persistence on industry choice: Bottom quartile

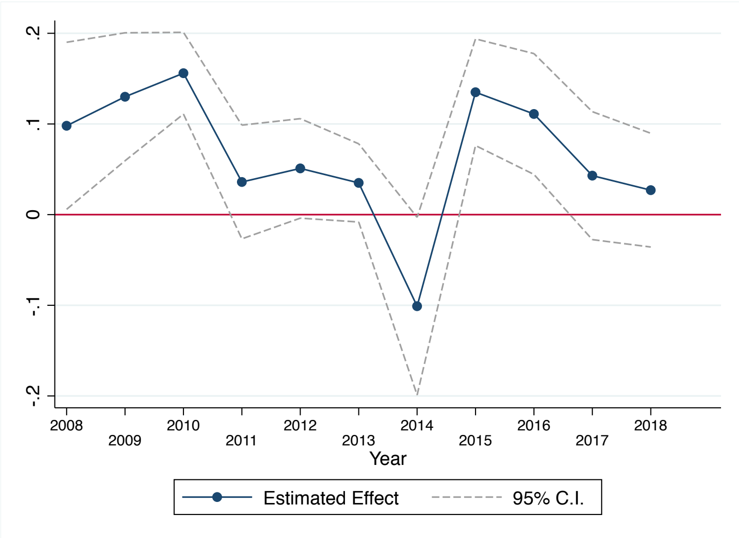
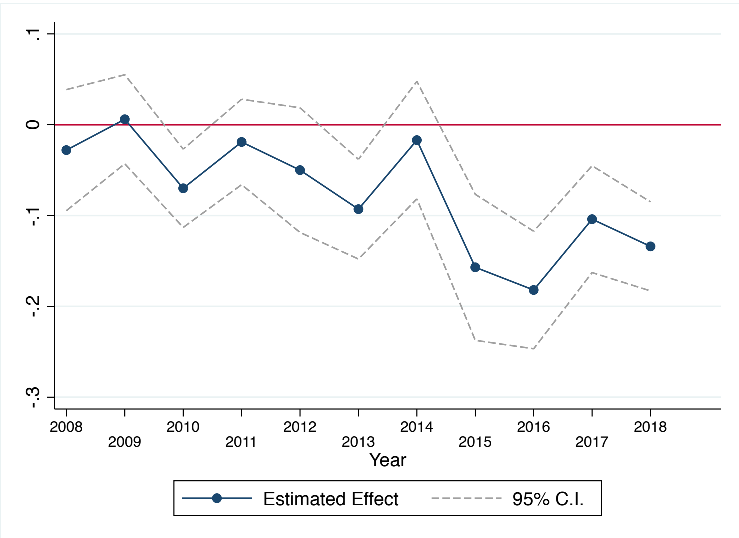


Figure 3.10: Persistence on industry choice: Top quartile



In short, these findings suggest that graduating together with a large cohort size does affect their wages negatively, possible by pushing them towards smaller firms, worse occupations, worse industries. Moreover, these adverse effects persist at several years after their graduation. On the other end of the spectrum, graduates from the top part of the wage distribution were less affected. The asymmetry of negative impact of cohort size on individuals' labor market outcome across the skill distribution is consistent with the mechanisms documented in Oreopoulos et al. (2012) and Altonji et al. (2016).

Table 3.12: Persistence of double cohort effect, I

Years	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After graduation	2	3	4	5	6	7	8	9	10	11
log wage										
New × Ont × Post	-0.100***	-0.169***	-0.036	-0.081***	-0.064***	0.101***	-0.135***	-0.151***	-0.025	-0.097***
	(0.030)	(0.022)	(0.023)	(0.023)	(0.022)	(0.033)	(0.027)	(0.024)	(0.028)	(0.021)
R^2	0.124	0.092	0.07	0.072	0.077	0.071	0.059	0.067	0.065	0.069
"Small" firm										
New × Ont × Post	0.089**	0.071***	0.042*	0.014	0.049***	-0.045**	0.102***	-0.067***	-0.014	-0.02
	(0.038)	(0.018)	(0.023)	(0.023)	(0.014)	(0.022)	(0.030)	(0.025)	(0.022)	(0.015)
R^2	0.012	0.015	0.014	0.011	0.014	0.011	0.012	0.011	0.008	0.011
"Big" firm										
New × Ont × Post	0.015	-0.191***	-0.067*	-0.056	-0.074***	0.051	-0.079***	-0.031	-0.093***	-0.115***
	(0.034)	(0.026)	(0.037)	(0.034)	(0.019)	(0.044)	(0.025)	(0.031)	(0.035)	(0.024)
R^2	0.016	0.015	0.013	0.015	0.015	0.012	0.013	0.015	0.018	0.015
N	18810	20280	20249	20187	20790	20951	20630	20235	20236	19846

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. New × Ont × Post stands for New graduates × Ontario × Post-2007. Source: LFS 2004-2018. Sample includes all employees who has a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Table 3.13: Persistence of double cohort effect, II

Years	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After graduation	2	3	4	5	6	7	8	9	10	11
Occupation: Bottom quartile										
New × Ont × Post	0.187***	0.211***	0.071**	0.062***	0.022	-0.090**	0.148***	0.072***	-0.023	0.046***
	(0.033)	(0.029)	(0.031)	(0.020)	(0.021)	(0.039)	(0.026)	(0.021)	(0.035)	(0.013)
R^2	0.084	0.062	0.034	0.032	0.032	0.026	0.021	0.028	0.024	0.028
Occupation: Top quartile										
New × Ont × Post	0.039	-0.086***	-0.004	-0.019	-0.079***	0.012	-0.146***	-0.179***	-0.130***	-0.131***
	(0.027)	(0.022)	(0.024)	(0.032)	(0.027)	(0.036)	(0.035)	(0.028)	(0.030)	(0.028)
R^2	0.083	0.074	0.066	0.06	0.062	0.049	0.042	0.047	0.043	0.045
Industry: Bottom quartile										
New × Ont × Post	0.130***	0.156***	0.036	0.051*	0.035	-0.101**	0.135***	0.111***	0.043	0.077
	(0.036)	(0.023)	(0.032)	(0.028)	(0.022)	(0.050)	(0.030)	(0.034)	(0.036)	(0.027)
R^2	0.091	0.072	0.034	0.038	0.041	0.028	0.025	0.031	0.029	0.032
Industry: Top quartile										
New × Ont × Post	0.006	-0.070***	-0.019	-0.05	-0.093***	-0.017	-0.157***	-0.182***	-0.104***	-0.117***
	(0.025)	(0.022)	(0.024)	(0.035)	(0.028)	(0.033)	(0.041)	(0.033)	(0.030)	(0.025)
R^2	0.084	0.073	0.063	0.058	0.059	0.047	0.041	0.047	0.039	0.043
N	18810	20280	20249	20187	20790	20951	20630	20235	20236	19846

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. New × Ont × Post stands for New graduates × Ontario × Post-2007. Source: LFS 2004-2018. Sample includes all employees who has a bachelor's degree. May to August data are excluded for precision of student status. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and month fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

Table 3.14: Persistence of double cohort effect, III

	(1)	(2)
	log weekly wage	
Years after graduation	3	8
New grads × Ontario × Post	-0.056***	-0.072***
	(0.008)	(0.007)
R^2	0.126	0.133
N	55882	55882

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables is log weekly wage. Source: Census 2006, NHS 2011 and Census 2016. Sample includes all employees who have a bachelor's degree as their highest education level. Quebec is excluded from the sample due to its unique educational structure. All the models include a set of fully saturated interaction terms between province dummy, year dummy, and experience group dummy. Other control variables include gender and immigrants fixed effect. Standard errors are between parentheses. Variance is clustered at province × New grads × year level.

3.6 Conclusion

In this article, I exploit the quasi-experimental nature of a 1999 Ontarian educational reform in order to study the impact of a labor supply on earnings. The policy in question abolished the fifth year of high school in this Canadian province and thus created a situation where two different cohorts graduated high school at the same time in 2003. I call these two "double cohort" throughout this paper. This situation dramatically increased the number of high school graduates that year and also induced a corresponding spike in the number of new university graduates four years later in 2007.

Focusing on university graduates, I use a triple difference-in-difference estimation to compare the wage gaps between recent graduates and experienced workers in Ontario versus the rest of Canada across time. At the baseline model, I find that the 28% increase in the new university graduates due to the double cohort is associated to an hourly wage decrease of 8.7%. Adding industry fixed effects and controlling for industry-experience-specific demand further increases this effect to 13.5%, suggesting that there is a reallocation of graduates towards industries in which the wage gap between experienced and inexperienced workers is relatively low. Similarly, part of the wage gap between new graduates and experienced workers can be explained by the composition change across industries.

I perform a series of robustness checks in order to alleviate any concerns about modeling choices that readers may have. First of all, the results are robust to different specifications of the experienced worker group. Another concern is that the results are at least partly driven by the fact that some students in the double cohort spent one less year in high school. I therefore restrict the treated group to the double cohort who spent five years in high school. The wage decrease thus obtained is virtually identical to our baseline specifications. Last but not least, I verify that employment status and migration from Ontario to other provinces are not affected by the policy neither.

In terms of distributional impacts, I conclude that the policy had heterogeneous effects across the distribution of wages. Indeed, top and bottom quartile earners were negatively affected by the cohort size, but not the median. A possible explanation may be the fact that more double cohort graduates are hired by "small" firms and there is a larger fraction of graduates employed in a low-paid occupation. Moreover, similar adverse effect is documented with industry profiles.

Next, I check if there is any evidence of spillover effects. To be more precise, it is possible that the sudden influx of university graduates affects recent university graduates and even high school graduates. I find that there is a 7.1% to 8% increase in the wages of recent high school graduates, indicating a possible complementarity between workers who possess a university degree and those with a high school diploma. Unfortunately, this result is possibly distorted by the fact that the wage rate of high school graduates in my sample are affected by the excess supply due to the double cohort high school graduates.

Lastly, I investigate the persistence of the adverse effects on Ontario's double cohort over time: wage penalties persist several years after graduation. The initial increase in the proportion of recent graduates work in a "small" firm remains present for at least six years, and then dies out gradually. Likewise, the jumps in the fractions of double cohort university graduates who are employed in a low-paid occupation and industry persist for several years as well.

Though focusing on a single cohort of students, the documented empirical findings have more general implications. In the end, the cohort size effect identifies the consequences of congestion in the labor market due to increased search competition as a result of excess labor supply. Thus the analysis not only helps us understand how the labor market functions in general, but also corroborates with the findings in the literature that study the effects of immigrants on natives' labor market outcomes, the effects of graduating during economic recessions etc.

Bibliography

Altonji, J. G., L. B. Kahn, and J. D. Speer (2016). Cashier or consultant? entry labor market conditions, field of study, and career success. *Journal of Labor Economics 34*(S1), S361–S401.

Angrist, J. D. and A. D. Kugler (2003). Protective or counter-productive? labour market institutions and the effect of immigration on eu natives. *The Economic Journal 113*(488), F302–F331.

Berger, M. C. (1985). The effect of cohort size on earnings growth: A reexamination of the evidence. *Journal of Political Economy 93*(3), 561–573.

Biagi, F. and C. Lucifora (2005). Demographic and education effects on unemployment in europe: Economic factors and labour market institutions.

Borjas, G. J. and J. Monras (2017). The labour market consequences of refugee supply shocks. *Economic Policy 32*(91), 361–413.

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.

Brunello, G. (2010). The effects of cohort size on european earnings. *Journal of Population Economics 23*(1), 273–290.

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.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review 43*(2), 245–257.

Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. *The Quarterly Journal of Economics 116*(2), 705–746.

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.

Eble, A. and F. Hu (2017). Does primary school duration matter? evaluating the consequences of a large chinese policy experiment.

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

Friedberg, R. M. (2001). The impact of mass migration on the israeli labor market. *The Quarterly Journal of Economics 116*(4), 1373–1408.

Garloff, A., C. Pohl, and N. Schanne (2013). Do small labor market entry cohorts reduce unemployment? *Demographic Research 29*, 379–406.

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.

Hunt, J. (1992). The impact of the 1962 repatriates from algeria on the french labor market. *ILR Review 45*(3), 556–572.

Jimeno, J. and D. Rodriguez-Palenzuela (2002). Youth unemployment in the oecd: demographic shifts, labour market institutions, and macroeconomic shocks.

King, A. J., W. Warren, W. Boyce, P. Chin, M. King, and J.-C. Boyer (2004). Double cohort study: Phase 4 report. *Social Program Evaluation Group, Queen's University*.

Korenman, S. and D. Neumark (2000). Cohort crowding and youth labor markets (a cross-national analysis). In *Youth Employment and Joblessness in Advanced Countries*, pp. 57–106. University of Chicago Press.

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*.

Moffat, J. and D. Roth (2016a). The cohort size-wage relationship in europe. *Labour 30*(4), 415–432.

Moffat, J. D. and D. Roth (2016b). Cohort size and youth labour-market outcomes: the role of measurement error. Technical report, IAB-Discussion Paper.

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.

Morin, L.-P. (2015). Cohort size and youth earnings: evidence from a quasi-experiment. *Labour Economics 32*, 99–111.

Oreopoulos, P., T. Von Wachter, and A. Heisz (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

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

Roth, D. (2017). Cohort size and transitions into the labour market. Technical report, IAB-Discussion Paper.

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.

Shimer, R. (2001). The impact of young workers on the aggregate labor market. *The Quarterly Journal of Economics 116*(3), 969–1007.

Skans, O. N. (2005). Age effects in swedish local labor markets. *Economics Letters 86*(3), 419–426.

Thiel, H., S. L. Thomsen, and B. Büttner (2014). Variation of learning intensity in late adolescence and the effect on personality traits. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 177*(4), 861–892.

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.

Welch, F. (1979). Effects of cohort size on earnings: the baby boom babies' financial bust. *Journal of Political Economy 87*(5, Part 2), S65–S97.

Wright, R. E. (1991). Cohort size and earnings in great britain. *Journal of Population Economics 4*(4), 295–305.