Cohort Size and Youth Earnings: Evidence from a Quasi-Experiment*

Louis-Philippe Morin†
University of Ottawa

May, 2013

Abstract

In this paper, I use data from the Canadian Labour Force Surveys (LFS), and the 2001 and 2006 Canadian Censuses to estimate the impact of an important labour supply shock on the earnings of young high-school graduates. The abolition of Ontario's Grade 13 generated a very large cohort of high-school graduates that simultaneously entered the Ontario labour market, generating a sudden increase in the labour supply. This provides a rare occasion to measure the impact of cohort size on earnings without the supply shock being possibly confounded with unobserved trends—a recurring problem in the literature. The Census findings suggest that the effect of the supply shock is statistically and economically important, depressing weekly earnings by 5 to 9 percent. The findings from the Census are supported by the LFS results that suggest that the immediate impact of the supply shock—measured about six months after high-school graduation—is also important.

Keywords: Labour Supply Shock, Youth.

JEL classification: J10, J20, J21.

*I would like to thank Michael Baker, Jean-Thomas Bernard, Pierre Brochu, Rose Anne Devlin, Eric Edmonds, Daniel Hamermesh, Diane Macunovich, Rob McMillan, Philip Oreopoulos, Solomon Polachek, Craig Riddell, Ben Sand, Aloysius Siow, and conference participants at the 2011 CLSRN/SOLE Meeting and the 2011 IZA/SOLE Transatlantic Meeting for helpful comments. I also thank Nyamekye Asare and Sergei Shibaev for their excellent research assistance. The data used in this study were accessed on site at the Carleton, Ottawa, Outaouais Local Research Data Centre (COOL RDC). This study reflects the views of the author, not of Statistics Canada.

†Contact info: Louis-Philippe Morin, Department of Economics, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5. Telephone: (613) 562-5800 ext. 4867. Email: lmorin@uottawa.ca.

1 Introduction

Economists have studied the effects of cohort size on youth economic outcomes extensively following the entrance of baby boomers onto the labour market and the associated worsening of the youth labour market situation. Since cohort size does not vary substantially from one year to the next, studies (e.g., Welch (1979); Berger (1985, 1989); Macunovich (1999); and Korenman and Neumark (2000)) have focused on long term (typically 8–25 years) variations in cohort size.[1] One problem with this strategy is that it is hard to isolate cohort size effects from other unobserved trends that are unrelated to demographics. This could explain why, for instance, in the 1980s the situation of youth in the United States worsened while demographic conditions should have improved it (Korenman and Neumark 2000).

The 1997 Ontario secondary school reform allows me to shed light on how well the labour market can absorb a sudden influx of workers. In particular, this reform provides a rare occasion to measure the impact of cohort size on youth earnings without having to worry about the supply shock being confounded with unobserved trends. Following the abolition of Grade 13, two cohorts of high school graduates simultaneously entered the labour market in 2003, creating a large and sudden youth labour supply increase. Compared to 2001, the number of high school graduates increased by more than 30 percent in 2003.

The Ontario supply shock can, in terms of its intensity, be compared to an immigration shock. Since Card's seminal 1990 paper, a series of studies (e.g., Hunt (1992); Carrington and de Lima (1996); Friedberg (2001); Glitz (2012)) have used important political changes as quasi-experiments to measure the impact of immigration supply shocks on local labour markets. Overall, the findings from these studies suggest that immigration supply shocks have, at most, a modest impact on natives (Friedberg and Hunt 1995).[2]

One advantage of using such quasi-experiments (over the use of cross-section analysis) is that it can mitigate, at least in part, self-selection issues such as the possibility that immigrants settle in booming labour markets if immigration is politically motivated, as opposed to economically motivated (Friedberg and Hunt 1995). But, although helpful in understanding the effect of immigration inflows on local labour markets, these studies can only shed limited light on the potential

[1] See Korenman and Neumark (2000) and Brunello (2010) for extensive reviews of the literature on cohort size and youth labour markets.

[2] One exception is Glitz (2012) who looks at the impact of the important inflow of immigrants to Germany following the fall of the Berlin Wall. Although he does not find evidence of a negative impact on wages, he does find that the immigration inflow affected the employment/labour force rate.

effects of exogenous increases of local workers, particularly if local workers and immigrants are poor substitutes.[3] One advantage of the supply shock studied in this paper is that it is composed of potential workers *almost identical* to what would be referred to in the immigration literature as 'native workers'. This study can therefore inform us on the capacity of the labour market to absorb supply shocks without having skills or preferences playing any confounding role in the determination of the outcome of interest. Another critical aspect of the supply shock studied in this paper is that it is less likely to be contaminated by a simultaneous labour demand shock than, say, an immigration supply shock. Ontario workers were already living and consuming in the local area, and as such, their entry into the workforce probably had a lesser impact on product demand (and thus on labour demand) as the arrival of new immigrants.

I take advantage of two sources of information to estimate the impact of the double cohort on youth earnings. First, I use the 2001 and 2006 Canadian Census master files, which allow me to observe youth earnings shortly before and after the double cohort. The Canadian Censuses are very useful to estimate the effect of cohort size for at least two reasons: 1) the richness of the data renders it possible to get a measure of weekly earnings—something that is crucial if we are interested in the effect of cohort size on the price of labour—, and 2) it is the largest Canadian data set available to researchers. The large sample size makes precise estimations possible, even for very small subsamples of the Canadian population (like Ontario high-school graduates born in 1984). The second source of data used in this paper consists of the 2002 and 2004 Labour Force Survey (LFS) master files. The LFS contains rich information on individuals' labour market conditions (e.g., hourly and weekly wages), and by observing individual earnings over an even shorter period surrounding the double cohort, it allows me to further mitigate the potential impacts of other unrelated shocks that could affect youth earnings.

My results show that a supply shock like the one created by the double cohort can significantly affect labour market outcomes. The Census results, based on a triple-difference estimation strategy, suggest that the Ontario double cohort decreased weekly wages of workers who recently graduated from high school by between 5 and 9 percent. Moreover, the magnitude of the estimated impact increases as the control group is further away in age from the treatment group. This indicates that workers close in age to the double-cohort graduates may have been affected by the supply shock as well. The double cohort also affected the likelihood to be working full time and for a full year. By

[3] See Ottaviano and Peri (2008), Card (2009), and Peri (2011) for evidence of imperfect substitutability between natives and immigrants.

taking this last finding into account, I estimate the 'upper' and 'lower' bounds of the supply shock effect on wages to be -3 and -17 percent.[4] The Census findings are corroborated by the LFS results, which indicate that the immediate (six months after the shock) impact of the double cohort was to depress wages by 11 to 25 percent. This last finding should be interpreted with caution, as the analysed sample size is relatively small.

As some studies (e.g., Borjas et al. (1996, 1997), Borjas (2003, 2006) and Boustan, Fishback and Kantor (2010)) suggest that native workers might move away from regions with significant in-migration, I also investigate whether young Ontario workers moved out of the province in reaction to the double cohort. I do not find any evidence of out-migration from young Ontario workers.

Interestingly, the impact of the double cohort is concentrated at the bottom of the wage distribution. I find that the double cohort affected both increased the proportions of young workers taking 'bad' jobs and depressed the wages conditional on have a bad job.

The next section describes the Ontario double cohort and its potential consequences for the estimation of the cohort size effect. I describe the two sources of data used in this paper in Section 3. The estimation strategy is presented in Section 4. Section 5 presents the findings from the Census data followed by the findings from the LFS. In Section 6, I tackle a number of identification issues, while I investigate some potential mechanisms through which the double cohort affected wages in Section7. Section 8 concludes.

2 The Ontario Double Cohort and Labour Supply

In 1997, the provincial government of Ontario introduced an important reform to its secondary school system. The centrepiece of this reform was the compression of the curriculum from five to four years. It brought the length of Ontario's secondary school curriculum into line with most surrounding provinces. Starting in 1999, students would now be expected to graduate from high school after four years (after Grade 12) instead of five.[5] An inevitable consequence of this reform was that, in 2003, both the first cohort from the new curriculum and the last cohort from the old curriculum graduated from high school in the same year, creating a drastic increase in the number of high school graduates. This large cohort of high school graduates was known as Ontario's Double Cohort. Since students graduate from secondary school almost simultaneously across the

[4] These bounds are computed using a trimming strategy, based on the work of Lee (2009). More details are given in Section 6.4.1.

[5] See King et al. (2002, 2004, 2005) for more details about the reform.

province, one would expect the labour supply shock caused by the double cohort to be important and concentrated within a short time span.

Figure 1 shows the number of recent high school graduates aged 17 to 19 between 1998 and 2006 for Ontario and the Rest-of-Canada (henceforth RoC).[6] The number of recent graduates jumped by 34.1 percent (from 91,291 to 122,406) between 2001 and 2003 in Ontario, while only increasing by 0.6 percent in the RoC over the same period.[7] The drastic contrast in growth rates in recent high school graduates, combined with an economic climate of stability in Canada over this period, will allow me to clearly identify the effect of an increase in cohort size on youth earnings.[8]

3 Data

In order to estimate the impact of the double cohort on youth earnings, I combine information from the Canadian LFS, and the 2001 and 2006 Canadian Censuses (long-form questionnaires). Both sources of information will complement each other as the Censuses contain a very large number of observations, while the LFS contains detailed labour force information and allows one to observed graduates shortly after having graduated from high school.[9] The large sample size of the Censuses will prove to be very helpful as the population of interest (i.e., 2003 Ontario high-school graduates who did not get post-secondary education) represents a small fraction of the Canadian population.

[6] Recent high school graduates are individuals who had graduated from secondary school at the time of their first LFS interview but who were attending school full time in the previous March. Note that, in the LFS, individuals are usually interviewed for six consecutive months. Source: 1998 to 2006 August Labour Force Surveys. See Appendix A for a detailed description of the data used to construct Figure 1.

[7] Although 2003 was labelled as the 'double' cohort year, the number of high school graduates was not twice as large in 2003 compared to 2002. Many college-bound students fast-tracked high school to graduate in 2002 or took an extra year to complete high school (graduating in 2004) in order to avoid the increased competition for college admission in 2003 (Morin 2013).

[8] Between 2000 and 2005, the average real GDP growth rates for Ontario and Canada were 2.3 and 2.5 percent, respectively. Importantly, Canada, unlike the US, did not experience a recession in 2001. Source: Statistics Canada Table 384-0002. One could be concerned that the 2000-2001 'high-tech bust' might have affected earnings in Ontario more than in the rest of Canada. Although the bust is believed to have mainly affected university graduates (Bowlby and Langlois 2002) and that these graduates are excluded from the analysis, the estimation strategy is design to control for province-specific shocks and the robustness section investigates the potential impacts of the tech bust. The results from my robustness checks (Section 6) suggest that the high-tech bust does not appear to be a cause for concern in this context.

[9] This strategy has been used before by Lemieux and Milligan (2008) for estimating the effect of social assistance on a variety of labour market outcomes (e.g., employment and annual earnings). They use the LFS to complement their Census results for the exact same reasons: 1) The large sample size of the Census data allows them to study a small subsample of the Canadian population, and 2) Since the LFS is conducted monthly Lemieux and Milligan (2008) observe individuals soon before, and soon after a policy change affecting social assistance.

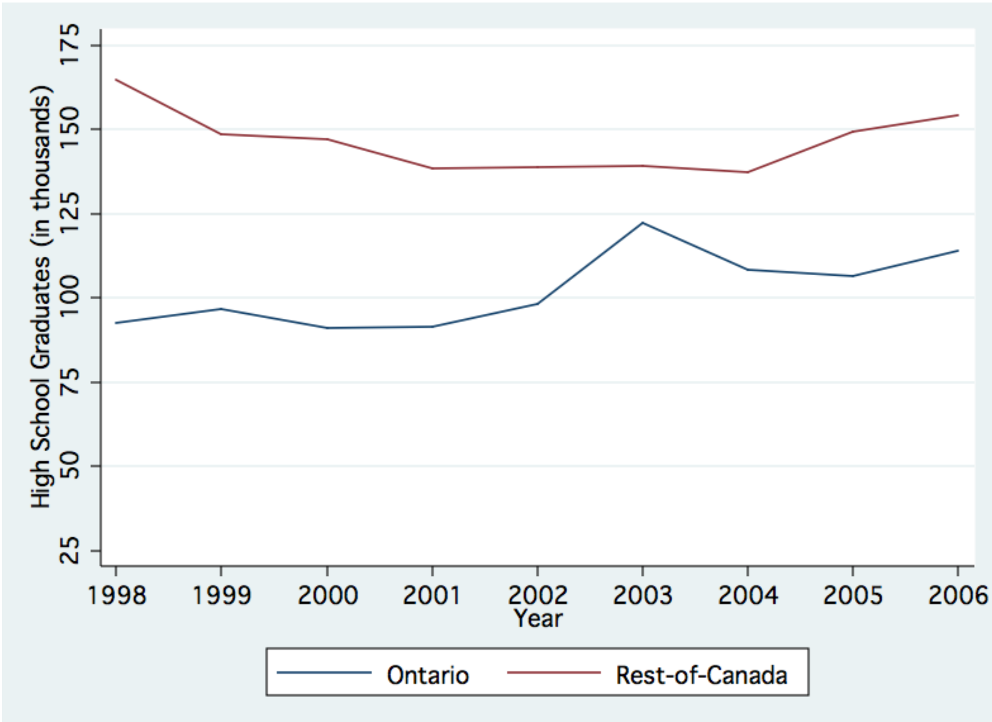

Figure 1: Number of New High School Graduates per Year

3.1 Census Master Files

The main findings of this paper are based on the Canadian Census master files. The 2001 and 2006 Census long-form questionnaires target approximately 20 percent of Canadian households. There are many advantages to using the Census master files when looking at the impact of the double cohort on youth earnings. First, the Census master files are the largest Canadian data sets available to researchers containing both detailed information on the respondents' earnings and education level. Since the main effect of the double cohort should be concentrated on a small fraction of the Canadian population, the size of the Census is crucial to compute any meaningful statistics.

Second, the long-form questionnaire is rich enough in terms of individuals' labour market activities to get a measure of one's price of labour. In particular, it contains information on the labour force status, the number of weeks worked last year, whether the individual mainly worked full- or part-time during these weeks, and their annual wages and salaries for the last year.

Third, the master files contain the year of birth of the individuals and not simply the age on the day on the survey. Since, Ontario uses December 31st as the cut-off date to determine when a child can enrol in primary school, it is straightforward to identify who is expected to have graduated from high school in 2003, and importantly, who should be a Grade 12 graduate (as opposed to Grade 13).

The Census also contains information on gender, educational attainment, visible minority status, immigrant status, marital status, the province of residence (now, one year ago, and five years ago), and workers' industrial sector. This information will be used to identify the 'treatment' and potential 'control' groups and as controls in the regression analysis. Finally, the Census also contains information on workers' occupation. This information will allow me to investigate whether the double cohort affected the types of occupation held by young Ontario workers.

The main variable of interest is the (log of) weekly wages earned in the year prior to the Census. Annual wages (i.e., gross wages and salaries before deductions) are adjusted using the provincial consumer price indices to be expressed in 2000 dollars, and divided by the number of weeks worked in the year prior to the Census to represent weekly wages.

I make a series of restrictions to help the identification of the cohort-size effect. First, I avoid having education playing any role in the wage determination by discarding Grade 12 graduates, and by focusing on individuals with a high school diploma, but no further schooling.[10] Grade 12 graduates are excluded from the analysis, in order to avoid having the effect of the cohort size being confounded with the potential (lack of) Grade 13 effect; Grade 12 graduates might have a lower level of human capital than Grade 13, thus including them in the analysis might bias the results. Indeed, when I include Grade 12 students, the estimated effects of the double cohort become more negative (by 4 percentage points on average). I further concentrate the analysis to full-time[11] workers as is done in studies where the number of hours worked is not perfectly observed (e.g., Katz and Murphy (1992), Card and Lemieux (2001), and Boudarbat, Lemieux and Riddell (2010), Green and Sand (2011)). In order to focus on high-school graduates who had fully entered the labour market, I restrict the sample to individuals who did not go to school, and worked 48 weeks in the year prior to the Census.[12] Looking ahead, in Section 6.4, I show that the employment status (i.e., working full-time, full-year) was also affected by the double cohort and compute bounds for

[10]Some changes to the educational-attainment questions in 2006 make it impossible to have a perfect match between the 2001 and the 2006 educational attainment variables. In particular, unlike the 2001 Census, the 2006 Census does not disentangle high-school graduates with further training (but no certificate) from high-school graduates without further training. I therefore labelled as high-school graduates without further schooling in 2001 high-school graduates regardless of whether they have further training, as long as they do not have a certificate above high school diploma. Excluding 2001 high-school graduates with further training increases the magnitude of the supply shock by about 1.5 percentage points. Hence, the estimates presented in this paper could be seen as being on the conservative side in that regard.

[11]The Canadian Census and the LFS define working full time as working 30 hours or more per week.

[12]The results obtained from looking at individuals who worked 26 weeks or more, or 39 weeks or more (three quarters of the year) are very similar to the ones presented in this paper.

the impact of the double cohort on wages that take this attrition/selection into account. Finally, I discard individuals with weekly wages of less than \$75 in 2000 dollars.[13] Appendix A presents more details on the Census data construction and restrictions.

3.2 Labour Force Survey Master Files

The Labour Force Surveys are conducted each month and they complement the Census data, here, as they allow me to concentrate on the very narrow group of individuals who should be most affected by the reform, Grade 13 graduates who entered the labour market a few months following the double cohort. By looking at a shorter time span (two years as opposed to five with the Census) I can further mitigate the potential impacts of other unrelated shocks on youth earnings. In particular, all individuals in the LFS estimations are observed significantly after the high-tech bust.

I rely on the January surveys for two reasons. First, since we only know the age of respondents in the LFS—as opposed to their year of birth in the Census—we can only disentangle Grade 13 from Grade 12 graduates in January. In January 2004, Grade 13 graduates should be 19 years old, while Grade 12 graduates should be 18. Second, full-time workers observed in January occupy regular jobs as opposed to a mix of regular and summer jobs for the months immediately following usual high-school graduation dates. Labour supply for summer jobs might be only driven by demographics (e.g., the number of individuals aged between 15 and 19) and not on schooling attainment.

Aside from allowing me to observe double-cohort graduates only a few months after their graduation, the LFS offers another advantage over the Census data. The LFS has information about workers' hourly wages, giving me a direct measure of the price of labour. Like the Census, the LFS contains information on gender, educational attainment, marital status, the province of residence, and workers' industry sector. Although there is no information about race or immigrant status in the LFS prior to 2006, the Census results suggest that the inclusion of these personal characteristics does not affect the estimated cohort effect.[14] I restrict the LFS sample to individuals who are not enrolled in school, that have a high school diploma (but no further schooling), and work full time (30 hours or more a week). I discard individuals with hourly wages less than \$2.5 in 2000 dollars,

[13]The same restriction is used by Boudarbat, Lemieux and Riddell (2010) and Green and Sand (2011). The cut-off roughly represents half of the minimum wage on a 30-hour week. A similar restriction is also used in Katz and Murphy (1992).

[14]There is some information in the LFS that is not available in the Census data. For example, there is information on union membership. Although the LFS results presented in this paper do not control for union membership (in order to be consistent with the Census estimations), I have estimated regressions where I control for it. Controlling for union membership does not affect the estimate of the impact of the double cohort on wages. These results are available upon request.

which is consistent with the weekly-wage cut-off of \$75 applied to the Census data. Appendix A presents more details on the LFS data construction and restrictions.

4 Estimating the Impact of Cohort Size on Earnings

Basic economic theory predicts that a positive supply shock should negatively affect wages. We would therefore expect to observe lower wages for individuals who were part of the double cohort as compared to a more 'normal' cohort of high-school graduates, after controlling for other factors affecting individual wages. A major difficulty faced by researchers is that other types of shocks—unrelated to cohort size—can occur around the time of the cohort-size increase. This is especially true when observing individuals over long periods of time. Here, the short time span over which individuals are observed (five years in the case of the Census data, and two years in the case of the LFS data), and the magnitude of the cohort size increase should mitigate this difficulty.

Although Ontario's economy grew at a steady pace and did not experience any significant downturn in the early 2000's, there are two (potential) demand shocks that must be accounted for when trying to identify the cohort-size effect: a demand shock that affects all Ontario workers, and one that affects young high-school graduates across Canada.

The identification strategy in this study is to disentangle the two types of shocks mentioned above from the labour supply shock following the double cohort, using both workers from Ontario who were presumably not affected by the supply shock and recent high-school graduates from other provinces as controls. In particular, I use triple-difference estimation which essentially compares wage gaps between a 'control' group (e.g., experienced workers) and recent high-school graduate workers across provinces and across time. Shocks specifically affecting recent high-school graduates across Canada can be controlled for by comparing wages of Ontario recent high-school graduates to wages of similar workers in other provinces. Demand shocks affecting Ontario can be captured by comparing the wages of recent high-school graduate Ontario workers to wages of other Ontario workers who should not be affected by the increase in cohort size, at least in the short run, but who should be affected by demand shocks. A triple-difference estimation strategy allows me to control for these two types of shocks simultaneously. After controlling for personal characteristics and for the potential effect of labour market conditions unrelated to the double cohort, changes in the outcomes of young workers between before and after the double cohort should be due to the increase in the number of recent high-school graduates.

The implementation of a triple-difference estimation is straightforward. The difficulty comes from choosing a group of workers affected by demand shocks in a similar fashion to recent high-school graduates while not being affected by the supply of this type of labour. The next sub-section presents details about the estimation technique and different control groups used to estimate the effect of a supply shock on wages.

4.1 Estimation Strategy

The triple-difference estimation strategy is represented in a regression framework by the following equation:

$$ln(w_{igpt}) = \lambda_{gt} + \eta_{gp} + \phi_{pt} + \beta(DC_t \times Youth_g \times ON_p) + \boldsymbol{X_{igpt}\gamma} + \varepsilon_{igpt} \quad (1)$$

where i represents an individual, g a group of workers (e.g., young versus older workers), p a province, and t represents time. $ln(w_{igpt})$ is the log of the weekly wages. Note that all individuals in the sample have the same educational attainment: a high school diploma. $Youth_g$ is a dummy variable equal to 1 if the individual is a recent high school graduate, and 0 otherwise. DC_t is a dummy variable equal to 1 if the individual is observed after the double cohort, 0 otherwise, while ON_p is equal to 1 if the individual resides in Ontario. Therefore, the $DC_t \times Youth_g \times ON_p$ term represents the 'treatment' group: Ontario high-school graduates who entered the labour market following the double cohort. The coefficient estimate for 'DC × Youth × ON', $\hat{\beta}$, captures provincial differences (i.e. Ontario versus RoC) in the changes in relative wages of young workers (compared to older workers) between 2000 and 2005. If equation (1) is correctly specified, $\hat{\beta}$ will capture the effect of the double cohort on youth earnings. λ_{gt}, η_{gp}, and ϕ_{pt} allow for the possibility that 1) the groups of workers have been affected differently by (demand) shocks across time (e.g., between 2000 and 2005 when using the Census data); 2) the average wage might differ across worker groups and that this difference might differ across provinces; 3) there were province specific shocks across time.[15] Finally, $\boldsymbol{X_{igpt}}$ is a vector of personal characteristics (e.g., gender, race, marital status, worker industry sector) that will be used to verify the robustness of my results.

To address the possibility of having a less than perfect control group, I estimate equation (1)

[15] Allowing for the average wage to differ across worker groups, and for this difference to differ across provinces, can be particularly useful here considering that high-school graduates in Quebec have eleven years of schooling, instead of twelve. Any wage gap due to this educational system difference should be captured in η_{gp}. Also, any inflation differences across provinces will be captured by ϕ_{pt}, since I look at the log of wages. The estimates for β are therefore identical whether I use nominal or real wages.

using different control groups to see whether the estimates vary significantly from one specification to another.[16] I consider workers with a high school diploma—the same level of education as the treatment group—but from different age groups and provinces as potential control groups. The idea is that more experienced workers are less likely to be close substitutes to recent high-school graduates, but would still be affected by labour demand shocks. When analysing the Census data, I divide the workers into six age groups: 21 years old (youth), 26 to 30, 31 to 35, 36 to 40, 41 to 45, and 46 to 50 years old. I do not use individuals aged 22 to 25 since many of them (especially those aged 22) might have entered the labour market at the same time as the double-cohort graduates.

5 Results

Before presenting the results from estimating equation (1), it is worthwhile to present summary statistics on the evolution of the average weekly wages between 2000 and 2005. Table 1 presents average weekly wages (in 2000 dollars) by age group and region (Ontario versus the RoC) for full-time, full-year workers. The number of observations for each group is presented in square brackets. One can notice an important strength of the Census data: its large sample size. For both the 2001 and 2006 Censuses, I observe more than 2,000 full-time, full-year Ontario workers that are 21 years of age and have a high-school diploma. The second striking finding from Table 1 is that the average weekly wages of young Ontario workers actually decreased by 8.3 percent between 2000 and 2005. This sharp decrease in wages is by far the most significant among all worker groups considered in Table 1. In the absence of any other shock to Ontario's economy, this drop in wages would be indicative of a significant labour-supply effect.

There is a significant difference between the growth rates experienced in Ontario and the RoC, which could indicate that other shocks (e.g., the high-tech bust) might have affected both regions differently. Importantly, the difference in growth rates does not seem to be driven by one specific group of workers in the RoC. In particular, young workers in the RoC saw their wages increase by 3.8 percent (an average annual growth rate of 0.7 percent), which is comparable to the growth rates of most of the worker age groups in the RoC. This systematic difference in growth rates justifies the use of a triple-difference estimation strategy, as opposed to a standard difference-in-differences estimation—recall that the triple-difference estimation presented in equation (1) is designed to

[16]If one believes that any type (e.g., experienced versus inexperienced, or skilled versus unskilled) of labour can be considered (to some extent) as a substitute to another labour type, then there is no perfect control group. Recall that the perfect control group would be affected by demand shocks in a similar way as recent high-school graduates, while not being affected by the increase supply of high-school graduates.

control for region-specific shocks. Interestingly, Ontario workers aged 26–30 and 31–35 also saw their wages decrease between 2000 and 2005 while older workers in Ontario experienced some weekly wage improvement over the same period. More generally, we can see the wage growth rates improve with age in Ontario. This finding could suggest that workers aged 26–35 might not have been totally isolated from the supply shock—this is in line with a decreasing level of substitutability across workers with larger age (or experience) differences. Note that this conjecture is further supported by the fact that we do not observe this trend in the RoC. Overall, the information found in Table 1 points toward a large impact of the double cohort on wages.

Table 1: Average Weekly Wages of Full-Time, Full-Year Workers (Census Data)

	Ontario			Rest of Canada		
Weekly Wages	2000	2005	Difference	2000	2005	Difference
Youth	438.46	402.18	-8.3%***	398.25	413.20	3.8%***
	(206.44)	(182.82)		(202.70)	(212.56)	
	[2,215]	[2,110]		[3,590]	[4,095]	
Aged 26–30	664.47	631.09	-5.0%***	604.11	600.48	-0.6%
	(373.34)	(318.64)		(333.48)	(332.59)	
	[15,965]	[15,445]		[19,440]	[21,520]	
Aged 31–35	765.47	747.75	-2.3%**	678.25	709.61	4.6%***
	(793.31)	(579.72)		(614.26)	(833.56)	
	[19,265]	[16,685]		[24,490]	[21,730]	
Aged 36–40	821.15	811.05	-1.2%	720.55	744.63	3.3%***
	(607.97)	(775.26)		(456.75)	(660.62)	
	[24,050]	[21,150]		[33,495]	[26,345]	
Aged 41–45	867.01	861.18	-0.7%	763.25	773.47	1.3%*
	(885.60)	(906.62)		(844.92)	(617.92)	
	[24,465]	[27,525]		[37,215]	[36,930]	
Aged 46–50	885.23	902.78	2.0%**	781.69	816.96	4.5%***
	(710.25)	(976.62)		(566.68)	(956.13)	
	[20,110]	[25,480]		[30,805]	[37,535]	

Notes: The average wages are expressed in 2000 dollars using provincial consumer price indexes. Standard deviations are in parentheses. The observations are weighted using the Census weights. The numbers of observations, rounded to a base of 5, are in square brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

5.1 Census Regression Results

Table 2 presents the regression results from estimating equation (1) using the Census data and workers aged 26 to 30 as the control group (a group of workers relatively close in age to the double-cohort graduates). Specification (1) only includes a set of fixed effects and interaction terms for time, province, and age group (η_{gp}, λ_{gt}, and ϕ_{pt} in equation (1)), along with the $DC_t \times Youth_g \times ON_p$

dummy variable.[17] Recall that the parameter estimate of $DC_t \times Youth_g \times ON_p$ is meant to capture the effect of the double cohort on the wages of young Ontario workers. Specification (2) adds personal characteristics (i.e., gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status) to the regression equation. Specification (3) includes industry fixed effects (based on 20 sectors of activity), while specification (4) allows for the personal characteristics to have differential effects, and the industry fixed effects to vary across worker groups *and* across time. Such interaction terms could capture potential industry-specific (e.g., the Mining and oil and gas extraction or the Professional, scientific and technical services sectors) shocks occurring between 2000 and 2005. Under this specification, any impact of a demand shock that was industry specific, province specific, or youth specific should not bias the estimator for the impact of the double cohort. Such specification should minimize the potential impact of shocks such as the high-tech bust on the estimated impact of the double cohort. All Census regression results presented in this paper are done using weights.[18]

The results coming out of Table 2 suggest that the double cohort had a statistically and economically significant impact on wages. All else equal, workers from the double cohort are earning on average about 6 percent less than similar workers who were part of a normal cohort. Adding control variables does not materially affect any of the estimates. In particular, the estimate of the double cohort effect ranges from -7.1 percent when only including basic controls (specification (1)) to -5.6 percent when industry fixed effects are included (specification (3)). Finally, the point estimate under specification (4) (6.0 percent) is very close the ones presented under specifications (1) through (3), making it unlikely (given the set of interactions terms included in the regression) that I am capturing the impact of the high-tech bust on high school graduates.[19]

A few other findings are worth mentioning. Workers aged 21 earned on average 34 percent less than workers aged 26 to 30 (from specification (3)) in 2000. This wage gap is fairly constant across provinces as only New Brunswick and Alberta have wage gaps that are significantly different from Quebec.[20] The wage gap decreased by about 3 percent between 2000 and 2005. Finally, the average

[17]Quebec is the base province in equation (1).

[18]In the case of the Canadian Census, only weighted estimation results can be released to the public. Unweighted regression results are almost identical. I use robust standard errors, instead of clustered at the province-year, province-age-group, or province-year-age-group level, as the robust standard errors are larger (especially when using the LFS data).

[19]Note that under specification (4), the interpretation of the coefficient estimates for 'Youth' and 'DC' are not the same as in specifications (1) to (3): under specification (4) they represent wage gaps (e.g., between younger and older workers for 'Young') in a specific industry (i.e., Agriculture, forestry, fishing and hunting). This is why the point estimates for these coefficients are so different between specifications (3) and (4).

[20]The differences in wage gap are 6.6 and 2.8 percent for New Brunswick and Alberta, respectively.

Table 2: Census Results Using Workers Aged 26-30 as the Control Group (Weekly Wages for Full-Time, Full-Year Workers)

	(1)	(2)	(3)	(4)
DC × Youth × ON	-0.071***	-0.062***	-0.056***	-0.060***
	(0.022)	(0.021)	(0.020)	(0.020)
Youth	-0.390***	-0.383***	-0.336***	-0.438***
	(0.013)	(0.013)	(0.012)	(0.153)
DC	-0.010	-0.008	-0.002	0.071**
	(0.009)	(0.009)	(0.008)	(0.031)
DC × ON	-0.050***	-0.052***	-0.047***	-0.045***
	(0.011)	(0.010)	(0.010)	(0.010)
DC × Youth	0.039***	0.031**	0.026**	0.034***
	(0.013)	(0.013)	(0.012)	(0.012)
Province Fixed Effects	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	No	Yes	Yes	Yes
Industry Fixed Effects	No	No	Yes	Yes
Personal Characteristics × Youth Fixed Effects	No	No	No	Yes
Industry × Youth Fixed Effects	No	No	No	Yes
Industry × DC Fixed Effects	No	No	No	Yes
R^2	0.10	0.17	0.25	0.25
N	84,375	84,375	84,375	84,375

Notes: The dependent variable is the ln of real weekly wages. The sample consists of individuals who worked 48 weeks or more during the year prior to the Census and worked full time during these weeks. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age on January 1st of the Census year. The omitted provincial dummy variable is Quebec. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The industry fixed effects reflect 20 sectors of activity (based on NAICS). The estimation was done using Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

real weekly wage of workers aged 26-30 did not change in Quebec, but did decrease significantly in Ontario by about 5 percent.

Table 3 compares the estimates of the effect of the increased cohort size for different aged-based control groups. The first column reports the results from specification (4) in Table 2. The next columns present the results from estimating the same specification, but for workers aged 31–35, 36–40, 41–45, and 46–50 respectively. One can clearly see that, as we move from younger to older control groups, the estimated effect of the double cohort increases significantly. When using workers aged 46 to 50, the estimated effect is -9.2 percent. At the same time, one can see that the difference in wage growth rates between Quebec and Ontario shrinks as we use older workers as control groups. For both workers aged 41–45 and 46–50, the difference is very close to zero and is no longer statistically significant. The results support the idea that similarly educated workers are seen as substitutes with the level of substitutability decreasing as age separating the workers increases.

Table 3: Double Cohort Effect and Aged-Based Control Groups (Weekly Wages for Full-Time, Full-Year Workers)

Age Group	26–30	31–35	36–40	41–45	46–50
DC × Youth × ON	-0.060***	-0.049**	-0.074***	-0.084***	-0.092***
	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)
DC × ON	-0.045***	-0.032***	-0.020**	-0.007	-0.001
	(0.010)	(0.010)	(0.009)	(0.008)	(0.008)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Personal Characteristics × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.25	0.29	0.30	0.30	0.32
N	84,375	94,180	117,050	138,145	125,935

Notes: The dependent variable is the ln of real weekly wages. The sample consists of individuals who worked 48 weeks or more during the year prior to the Census and worked full time during these weeks. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age on January 1st of the Census year. The omitted provincial dummy variable is Quebec. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The industry fixed effects reflect 20 sectors of activity (based on NAICS). The estimation was done using Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

5.2 LFS Results

In this subsection, I present the results obtained using the LFS to complement the Census results. The LFS allows me to estimate the immediate impact of cohort size on wages using two surveys that are only two years apart (the 2002 and 2004 January LFS). Looking at such a small time span should minimize the likelihood of having other unrelated trends (or demand shocks) affecting the estimation of the double cohort impact on wages. The estimation strategy is exactly the same as the one used with the Census data. The main difference is that the 'Youth' worker group is composed of 2001 and 2003 high-school graduates born in 1982 and 1984, as opposed to 1979 and 1984, when using the Census data. By using the January 2004 LFS, I can observe the wages of young workers only a few months after their graduation.

Table 4 presents the regression results from estimating equation (1) using the LFS data and workers aged 24 to 28 as control group.[21] In order to be consistent with the Census results, I present the LFS results from using the log of weekly wages as the dependent variable. Using hourly wages instead of weekly wages gives very similar results. The four specifications in Table 4 are the same as in Table 2 with the exception that the LFS data do not contain information about race or immigrant status.

The results from Table 4 corroborate the Census results, suggesting that the supply shock had a significant effect on youth wages. The estimated double-cohort effect is larger in magnitude than when using the Census data, but it is also less precisely estimated. This is not surprising given the smaller sample size in the LFS. Workers from the double cohort earn on average about 20 percent less than similar workers who were part of a normal cohort, suggesting that the labour market reacted strongly to the supply shock, at least in the very short run.

Table 5 compares the estimates of the effect of the increased cohort size for different aged-based control groups. Unlike the results coming out of the Census data, the estimates do not show a clear pattern across age groups, and not surprisingly, the estimates also fluctuate more in Table 5 than in Table 3. The estimates fluctuate between -11 and -25 percent. Nevertheless, all estimates are statistically significant and far from zero, indicating that the choice of the control group is not critical. Overall, despite being less precise than the Census results, the LFS results suggest that the double cohort had a significant impact of the youth labour market.

It is important to note that the difference between the LFS and the Census results should not necessarily be interpreted as a fadeout of the supply shock effect. This difference could be due to a couple of other factors. First, as mentioned above, the small number of observations in the LFS makes the estimates less precise.[22] A 95 percent confidence interval for the impact of the double cohort on wages based on the LFS regression results would include the Census estimate in most cases. Second, one could see the results of the LFS as an attempt to estimate the treatment-on-the-treated effect while the Census results inform us on the intent-to-treat effect. Furthermore, the results from the LFS are more subject to self-selection than the Census results. In particular, Figure 1 shows that the number of high-school graduates increased significantly in 2002 and stayed

[21] Workers aged 24 to 28 in January 2004 would be aged 26 to 30 in January 2006.

[22] I have estimated equation (1) using the both the December and January LFS surveys in order to increase the number of observations, without adding too much measurement error on individuals' birth year. The results are very similar to the ones using the January surveys only. These results are available upon request. It should be noted that, since most individuals are interviewed for six consecutive months in the LFS, only about a sixth of the observations differ between December and January.

Table 4: LFS Results Using Workers Aged 24-28 as the Control Group (Weekly Wages for Full-Time Workers)

	(1)	(2)	(3)	(4)
DC × Youth × ON	-0.238**	-0.215**	-0.184**	-0.196***
	(0.095)	(0.084)	(0.074)	(0.071)
Youth	-0.492***	-0.462***	-0.432***	-0.575***
	(0.056)	(0.056)	(0.053)	(0.208)
DC	0.040	0.043	0.052	0.253
	(0.063)	(0.060)	(0.059)	(0.176)
DC × ON	-0.020	-0.027	-0.033	-0.017
	(0.075)	(0.072)	(0.070)	(0.070)
DC × Youth	0.115**	0.114**	0.101**	0.100**
	(0.048)	(0.046)	(0.043)	(0.043)
Province Fixed Effects	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	No	Yes	Yes	Yes
Industry Fixed Effects	No	No	Yes	Yes
Personal Characteristics × Youth Fixed Effects	No	No	No	Yes
Industry × Youth Fixed Effects	No	No	No	Yes
Industry × DC Fixed Effects	No	No	No	Yes
R^2	0.23	0.32	0.39	0.39
N	1,900	1,900	1,900	1,900

Notes: The dependent variable is the ln of real weekly wages. The sample is composed of full-time workers. 'Youth' is an indicator variable is equal to 1 if the individual is 19 during the LFS reference week. All individuals in the sample have a high school diploma, but no further schooling. The personal characteristics include: gender, a rural area indicator, and marital status. The industry fixed effects are constructed using 9 sectors of activity (based on NAICS). The estimation was done using the LFS weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

relatively high in 2004 suggesting that some high school students 'fast-tracked' high-school while others slowed down in order to avoid to the double cohort. This selection could affect the LFS results since these individuals are not accounted for when estimating the impact of the double cohort. This is not the case with the Census results as I observe almost all double-cohort graduates in 2006.

6 Identification Issues

Overall, both the Census and LFS results presented so far suggest, unlike the findings from previous studies looking at the impact of immigration shocks, that a supply shock of 'native' workers can significantly affect wages, and that this effect can spread to similarly educated workers. In

Table 5: Double Cohort Effect and Aged-Based Control Groups (Weekly Wages for Full-Time Workers)

Age Group	24–28	29–33	34–38	39–43	44–48
DC × Youth × ON	-0.196***	-0.250***	-0.188***	-0.110*	-0.180***
	(0.071)	(0.081)	(0.071)	(0.065)	(0.067)
DC × ON	-0.017	0.054	-0.032	-0.070	0.015
	(0.070)	(0.079)	(0.060)	(0.053)	(0.052)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Personal Characteristics × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.39	0.45	0.44	0.44	0.43
N	1,900	1,815	2,405	2,890	2,925

Notes: The dependent variable is the ln of real weekly wages. The sample is composed of full-time workers. 'Youth' is an indicator variable is equal to 1 if the individual is 19 during the LFS reference week. All individuals in the sample have a high school diploma, but no further schooling. The personal characteristics include: gender, a rural area indicator, and marital status. The industry fixed effects are constructed using 9 sectors of activity (based on NAICS). The estimation was done using the LFS weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

this section, I present a series of robustness checks investigating the main potential identification issues. Before looking at whether the supply shock also affected other important outcomes (like the employment status), I present 1) results from using a control group chosen specifically to tackle the issue regarding the potential impact of the high-tech bust on youth earnings, 2) results from a falsification test, and 3) I investigate whether immigration between 2000 and 2005 could be a threat to the identification of the impact of the double cohort on wages.

6.1 An Alternative Control Group

One potential concern with the results presented so far is that, if a shock unrelated with the Ontario double cohort affected young Ontario workers with a high school degree more than the other Ontario workers, then some of this shock could be captured by the DC × Youth × ON coefficient estimates. One such potential shock is the 2000-2001 high-tech bust that may have affected Ontario more than the other provinces. In this section, I present results from which the main 'control' group was known to be negatively affected by the tech bust: recent university graduates. Bowlby and Langlois (2002) suggest that university graduates suffered the most (in terms of job losses) from the high-tech bust—this is not surprising given the nature of the jobs in this sector. By using such a control group, I should introduce a bias going against finding a double-cohort effect on high-school graduates.

Table B.1 presents the results from using recent university graduates (i.e., individuals 21 to 25 years of age on January 1st of the Census year and holding a bachelor's degree) as control group as opposed to older individuals with the same educational attainment (as done in Tables 2 and 3). The coefficient estimates are in line with the results presented so far: the estimate being statistically significant and around -5.5 percent. Although, the point estimates are slightly lower than the ones using older workers (aged between 36 and 50), Table B.1 findings support the the idea that I am not simply capturing the impact of the high-tech bust on young Ontario workers with a high-school degree in Tables 2 and 3. This is especially true if we think that a severe high-tech bust effect on recent Ontario university graduates should bias downward the estimator for the double-cohort effect on young Ontario workers with a high-school degree.

6.2 Falsification Test

If one wanted to play devil's advocate, one could argue that the exact same estimate could be obtained if 2000 was particularly *good* for Ontario young workers, and that things went back to normal in 2005. In order to investigate this possibility, I use the 1996 and 2001 Censuses to conduct a falsification test. The 1995-2000 period was one of solid economic expansion in Canada (and Ontario), but without any supply shock comparable to the double cohort. I generated a false double cohort (for 1996) and estimated equation (1) where 'DC $\times$ Youth $\times$ ON' is replaced by 'False DC $\times$ Youth $\times$ ON'. Table 6 presents the results of the falsification test. As in equation (1), the coefficient estimates for 'False DC $\times$ Youth $\times$ ON' captures provincial differences (i.e., Ontario versus ROC) in the changes in relative wages of young workers (compared to older workers), but between 1995 and 2000. All of the 'False DC $\times$ Youth $\times$ ON' coefficient estimates are small and statistically insignificant, supporting the idea that I am capturing the impact of the double cohort and not some other unrelated shock.

6.3 Immigration

Immigration flows were relatively high in Canada between 2000 and 2005, and Ontario has historically been the major recipient of immigrants (Chui, Tran and Maheux 2007). If the increase in the share of immigrant population between 2000 and 2005 was more pronounced in Ontario, and that immigrants mainly compete with inexperienced (or young) workers, then some of the esti-

Table 6: Falsification Test Based on the 1996 and 2001 Censuses

Age Group	26–30	31–35	36–40	41–45	46–50
False DC × Youth × ON	-0.007	0.004	0.002	-0.014	-0.021
	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Personal Characteristics × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.25	0.29	0.30	0.30	0.33
N	86,085	105,740	123,585	125,205	101,020

Notes: The dependent variable is the ln of real weekly wages. The sample consists of individuals who worked 48 weeks or more during the year prior to the Census and worked full time during these weeks. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age on January 1st of the Census year. 'False DC' is an indicator variable that equals 1 if the individual is from the 1996 Census. The omitted provincial dummy variable is Quebec. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The industry fixed effects reflect 20 sectors of activity (based on NAICS). The estimation was done using Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

mated impact of the double cohort on youth earnings could be due the disproportionate increase in immigrants in Ontario. I directly address this issue by investigating whether the share of the immigrant population increased more in Ontario than in other provinces. More specifically, I estimate regressions where the dependant variable is a dummy variable equal to 1 if the individual is an immigrant, 0 otherwise. As explained below, there is no clear indication that immigration should be a cause for concern for the identification of the double cohort effect on wages.

Whether I look at all individuals aged 21–50, individuals not enrolled in school aged 21–50, high-school graduates, not enrolled in school and aged 21–50, or FTFY workers aged 21–50, simple difference-in-difference estimations do not suggest that the share of immigrants increased more in Ontario than in the RoC between 2001 and 2006. If anything, the share of immigrants increased slightly less in Ontario than in the RoC. Note that if recent immigrants, regardless of their age, compete with young Ontario workers for jobs, the difference-in-difference approach (Canada versus RoC) is probably more suitable than estimating an equation like (1). Nevertheless, I get a similar conclusion when I estimate a regression using specification (1) of equation (1) on individuals not enrolled in school aged 21–50, high-school graduates, not enrolled in school and aged 21–50, or FTFY workers aged 21–50: if anything the proportion of immigrants among young Ontarians increased less than among the other groups of workers (older workers in Ontario, and young and older workers in the RoC). The only case where I find an increase, albeit small, in the proportion of immigrants among young Ontarians relative to the other groups of individuals is when I include individuals enrolled in school. Note that when I am looking at these individuals, I include individuals enrolled in university and college, whom are not likely to compete for jobs against the group of workers studied here.

6.4 Increased Cohort Size, Employment, Immigration, Out-Migration, and School Attendance

When interpreting the results coming out of the Census data, one should keep in mind that I restricted the sample to full-time, full-year (FTFY) workers. It is quite possible that the FTFY status itself and other important outcomes have been affected by the supply shock. I now estimate the impact of the double cohort on FTFY status, out-migration, schooling attainment and school enrolment.

6.4.1 Full-Time, Full Year Status

Simple descriptive statistics suggest that the fraction of FTFY workers among Ontario youth decreased by 1.6 percentage points between 2000 and 2005.[23] This difference is both statistically and economically significant since the fraction of FTFY workers was 19.5 percent in 2000. I further investigate the potential impact of the double cohort on the likelihood to be a FTFY worker by estimating equation (1), using a FTFY dummy as dependent variable (instead of the log of wages). The results are presented in Table 7.

There does not appear to be any change in the likelihood of being FTFY when using individuals aged 26–30 as the control group, but a statistically significant difference appears as we move to older control groups. The estimates obtained when using individuals aged 36–40, 41–45, or 46–50 are in the vicinity of the 1.6 percentage point difference when simply looking at the change in proportions. This finding is interesting as it suggests the same age-based pattern for the impact of the supply shock as the one found when looking at wages. It is quite possible that the FTFY

[23] This fraction is obtained by dividing the number of 21 year-olds with a high school diploma that work full-time, full-year by the total number of 21 year-olds with a high school diploma.

Table 7: Double Cohort and Full-Time, Full-Year Status

Age Group	26–30	31–35	36–40	41–45	46–50
DC × Youth × ON	0.005	-0.009	-0.015*	-0.018**	-0.012
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
R^2	0.13	0.15	0.15	0.15	0.15
N	247,930	255,135	291,505	321,385	295,360

Notes: The dependent variable is a dummy variable equal to 1 if the individual worked full-time, full-year in the year prior to the Census. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The estimation was done using the Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

status of workers aged 26–30 was affected by the supply shock, supporting the idea that there is some level of substitutability between similarly educated workers given that they are close in age.

There is no obvious valid instrumental variable to deal with the endogeneity of the FTFY status, but I can still compute 'worst-case' scenario bounds on the impact of the double cohort on wages based on the work of Lee (2009).[24] The estimated upper and lower bounds for the average treatment effect on wages are -3 and -17 percent, respectively. The fact that the upper bound for the effect is negative is somewhat surprising given that I compute it assuming that the individuals for which the FTFY status was affected by the supply shock would have had the largest wages in the absence of the shock. In practice, this would be very surprising given the positive correlation between weekly wages and the number of weeks worked found in the Census data. In the end, the estimation results based on 'worst-case' scenarios emphasize the important impact of the supply shock on wages.

[24] When applying his methodology to the analysis of the Job Corps program, Lee (2009) trims the treatment group data, as the program is assumed to *positively* affect both wages and the probability to be employed. Since the supply shock is expected to have a *negative* impact on both the FTFY status and wages, I trim the 'control' group data instead of trimming the treatment group data. In particular, I trim the top (bottom) 8.2 percent—the percent change in the proportion of full-time, full-year workers among Ontario youth between 2000 and 2005—of the of Ontario 2001 Youth group data and then re-estimate equation (1) to get an estimate of the lower (upper) bound. More detailed results are available upon request.

6.4.2 Out-Migration

In an attempt to explain the small impact of immigration on the wages of local worker, some studies explored the possibility that native workers react to an increase in immigration by moving to another labour market (Borjas, Freeman and Katz 1997, Borjas 2003, Kugler and Yuksel 2008).[25]

The Canadian Census allows me to investigate the possibility that, in order to avoid facing the increased labour supply, some Ontario workers moved to a different labour market. In particular, the Census possesses information on the province of residence five years prior to the Census. I can therefore test whether a significant portion Ontarians moved to a different province between 2001 and 2006. Table 8 presents the results from estimating equation (1) with a dummy equal to 1 if the individual moved to a different province between one and five years prior to the Census.

Table 8: Double Cohort and Out-Migration

Age Group	26–30	31–35	36–40	41–45	46–50
DC × Youth × ON	-0.006	0.000	-0.003	-0.003	-0.001
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
R^2	0.13	0.15	0.15	0.15	0.15
N	247,930	255,135	291,505	321,385	295,360

Notes: The dependent variable is a dummy variable equal to 1 if the individual the province of residence 5 five years ago is different from the province of residence one year ago. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The estimation was done using the Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8 does not suggest that Ontario students tried to escape the double cohort by moving to another province. The coefficient estimates are small and statistically insignificant for all age-based control groups. As point of comparison, 2.4 percent of young Ontarians in 1996 had moved to a different province in 2000. These results are in line with studies by Card and DiNardo (2000), Card (2001), and Glitz (2012), suggesting little migration response from natives to labour supply shocks.

[25] Aydemir and Borjas (2011) suggest an alternative explanation based on measurement error.

6.4.3 School Enrolment and Schooling Attainment

An expected increase in cohort size could affect school enrolment and educational attainment which could, consequently, affect the estimation of the supply shock effect on earnings (Connelly 1986, Stapleton and Young 1988). Therefore, I now investigate the possibility that the double cohort affected school enrolment and/or schooling attainment of high-school graduates.

Competition for post-secondary institutions increased significantly following the double cohort (Morin 2011). It is therefore quite possible that a number of college-bound students were unable to join a post-secondary institution in 2003 (due to the increased admission standards) and joined the labour force instead. If so, the proportion of 21 year olds enrolled in post-secondary schooling would decrease in Ontario relative to the other provinces. Recall that the double-cohort college-bound students were expected to enter college in 2003 and many of them were expected to graduate from university in 2007. Therefore, they were expected to be enrolled in school at the time of the 2006 Census.

There are some data quality issues with the school enrolment question in the 2006 Census, according to Statistics Canada, which make it impossible to know whether enrolment to post-secondary institutions changed over time.[26] We can nevertheless know whether school enrolment (at any type of institution) changed following the double cohort. The double cohort does not seem to have affected the proportion of 21 year olds enrolled in school—the enrolment rate increased by (a non-statistically significant) 0.7 percentage points relative to the other provinces. This result is consistent with students not admitted to a post-secondary institution in 2003 being admitted (and enrolled) in 2004, for example.

The double cohort did affect the proportion of 21 year olds with a post-secondary diploma, but mainly among individuals still enrolled in school. When compared to other Canadians of the same age, the proportion 21-year-old Ontarians with a post-secondary diploma increased by 2.8 percentage points. This difference is statistically significant. When we concentrate on individuals that are not enrolled in school the increase drops to 1.5 and is no longer statistically significant. This result seems to suggest that the double cohort did not affect the decision to attend post-secondary schooling, but did affect the type of post-secondary institution attended.

Overall, the proportion of 21 year old high-school graduates not enrolled in school and without a post-secondary diploma—the group of individuals on which my analysis is based—does not seem

[26] See Statistics Canada (2008) for more details.

to have changed significantly between 2001 and 2006, relative to the RoC. This proportion only decreased by a non-statistically significant 0.9 percentage point relative to the other provinces.[27]

7 Distributional and Occupational Impacts

While the Census and LFS results on the mean impacts of the double cohort are informative, it is quite likely that the double cohort did not simply shift the wage distribution downward. I therefore investigate 1) whether the double cohort affected the wage distribution in a non-trivial way, and 2) whether the observed wage differences could be explained by shifts in the distribution of occupations of young workers.

7.1 Wage Distribution

In order to investigate the potential distributional effects of the double cohort, I estimated quantile regressions based on equation (1) for the bottom quartile, median and top quartile of the weekly wage distribution. Table 9 presents the estimates for DC $\times$ Youth $\times$ ON for each quartile and each age group.[28]

The results from Table 9 suggest that the impact of the double cohort is largest at the bottom quartile. The estimated impact at the top quartile is less than 60 percent the estimated impact at the bottom quartile. This holds true for all age groups. In fact, the results in Table 9 suggest that the impact of the double cohort decreases monotonically with the distribution quantile.[29] This finding suggests that relatively highly paid jobs (among high-school graduates) were less affected by the double cohort.[30] I examine the potential sources of the heterogeneity in the impact of the double cohort in the next subsection.

7.2 Occupations

In this subsection, I investigate the potential impact of the double cohort on the types of occupations held by young Ontario workers. This will shed light on the potential mechanisms by which the double cohort shock reduced real wages among young workers. Specifically, I look at whether the

[27] All regression results from which the numbers presented in this sub-section were taken from are available upon request.

[28] Note that each estimate presented in Table 9 are obtained from separate regressions.

[29] I have estimated quantile regressions for 5th to 95th quantiles. These results confirm that the impact of the double cohort decreases (in magnitude) as we move up the wage distribution. These results are available upon request.

[30] Note that this is contrary to what we would expect if my results were driven by the high-tech bust (if we expect the high-tech bust to mainly affect relatively 'good' jobs for high-school graduates).

Table 9: Quantiles Regression Estimates for the Impact of the Double Cohort on Weekly Wages

Age Group	26–30	31–35	36–40	41–45	46–50
DC × Youth × ON on Bottom Quartile	-0.106***	-0.098***	-0.116***	-0.133***	-0.128***
	(0.021)	(0.023)	(0.022)	(0.021)	(0.023)
DC × Youth × ON on Median	-0.080***	-0.052***	-0.082***	-0.086***	-0.101***
	(0.018)	(0.019)	(0.019)	(0.020)	(0.018)
DC × Youth × ON on Top Quartile	-0.037*	-0.017	-0.044**	-0.049**	-0.074***
	(0.021)	(0.022)	(0.019)	(0.023)	(0.022)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Personal Characteristics × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
N	84,375	94,180	117,050	138,145	125,935

Notes: The dependent variable is the ln of real weekly wages. Note that each estimate are obtained from separate regressions. The sample consists of individuals who worked 48 weeks or more during the year prior to the Census and worked full time during these weeks. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age on January 1st of the Census year. The omitted provincial dummy variable is Quebec. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The industry fixed effects reflect 20 sectors of activity (based on NAICS). The estimation was done using Census weights. The numbers of observations are rounded to a base of 5. Standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

impact of the double cohort that we observe in Table (3) was due to 1) a generalized wage decrease across all occupation types, 2) a compositional change in the type of occupations held by young Ontario workers, or 3) a combination of 1) and 2).

The Canadian Census has information on workers' occupations, which allows me to investigate whether young Ontario workers were more likely to end up working in 'bad' jobs following the double cohort. To do so, I first ranked the occupations based on a wage measure, as is done in many studies on job polarization (e.g., Goos and Manning (2007), Dustmann, Ludsteck and Schönberg (2009), and Green and Sand (2011)). I use the 2000 weekly wage distribution of full-time full-year workers (with a high-school diploma) to rank more than 500 occupation profiles.[31] I then constructed a dummy variable equal to one if the occupation profile was part of the bottom quartile of the wage distribution in 2000, and zero otherwise.

Table 10 presents the results from estimating equation (1) with a bad-job dummy variable as dependent variable. All but one age-based control groups suggest that the double cohort increased

[31] See Green and Sand (2011) for more details on the occupational classification in the Canadian Census data.

Table 10: Double Cohort and Bad Occupations

Age Group	26–30	31–35	36–40	41–45	46–50
DC × Youth × ON	0.031*	0.019	0.037**	0.036**	0.042**
	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Personal Characteristics × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.32	0.31	0.30	0.28	0.27
N	84,375	94,180	117,050	138,145	125,935

Notes: The dependent variable is a dummy variable equal to 1 if the individual worked in a low-paying occupation according to the 2001 Census. An occupation is considered as low paying if its average weekly wage was in the bottom quartile of the occupation wage distribution. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The estimation was done using the Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

young Ontario workers' probability of having a bad job by about 4 percentage points, all else equal. This increase is not small if we consider that 34 percent of the 2001 Ontario 'Youth' group had a low-paying occupation. This suggests that part of the impact of the double cohort on wages (found in Table 3) could be due to a change in the distribution of occupation types.

At the other end of the occupation spectrum, Table B.2 shows no clear indication that young Ontario workers were significantly less likely to have a 'good' job (i.e., an occupation profile in the top quartile of the wage distribution in 2000) following the double cohort. This last finding suggests that the impact of the double cohort was not uniform across occupation profiles. Such a finding is further evidence that I am not capturing (in Table 3) the impact of the high-tech bust. If the high-tech sector offered 'good' jobs for high-school and university graduates, then we could expect its bust to push high-school and university graduates out of these occupations. This push out of good jobs could be worse for high-school graduates if we expect some sort of trickle-down effect following the demand shock (e.g., see Beaudry, Green and Sand (2013))–university graduates could now occupy 'good' jobs that used to be occupied by high-school graduates. None of the coefficients estimates presented in Table B.2 suggest that such push out of 'good' jobs occurred following the double cohort, at least not more in Ontario than in the RoC.

Finally, if we concentrate the weekly wages of workers in bad jobs (according to the 2001 Census occupation ranking), we see that the wage decrease following the double cohort is, for these workers, about twice as large as the one estimated in Table 3. This is in line with the results in subsection 7.1 suggesting that the impact of the double cohort is concentrated in the lower tail of the wage distribution. The results not presented here are available upon request.

8 Conclusion

For years, economists have been interested in estimating the impact of cohort size on labour market outcomes. Given the small year-to-year variations in cohort size, researchers have typically focused on long-term fluctuations. Doing so introduces an important identification issue as it becomes difficult to separate the effect of cohort size from other unrelated trends—an issue which becomes more serious as the period under study lengthens.

This paper studies the effect of the 2003 Ontario double cohort on youth earnings. The double cohort generated a large and sudden influx of workers, making it possible to clearly identify the impact of cohort size on wages. In particular, the short time span over which the supply shock occurred helps resolve the identification problem faced by previous studies looking at cohort size effects.

My results suggest that the double cohort significantly depressed the wages of young workers. The Census results suggest that the wages of full-time, full-year workers decreased by 5 to 9 percent due to the supply shock—this effect being estimated two years after the double cohort. Interestingly, the estimated impact of the supply shock becomes more negative as the control group is further away in age to the treatment group, suggesting that workers close in age to the double-cohort graduates were also affected by the supply shock. The Census findings are corroborated by the LFS results, suggesting that the immediate impact of the double cohort was to depress wages by 11 to 25 percent. Not only did the supply shock affect the wages of full-time, full-year workers, but it also affected the likelihood to be working full time and for a full year by about 1.5 percentage points. Accounting for this effect on labour market participation, I estimate the impact of the supply shock on wages to be between -3 and -17 percent. Interestingly, the impact of the double cohort seems to have both increased the proportions of young workers taking bad jobs and depressed the wages in these kinds of job. The end result is that the impact of the double cohort is concentrated at the bottom of the wage distribution. In contrast to the findings from the literature on the effects of sudden immigration inflows, the findings from this paper suggest that a sudden inflow of 'native' workers can have significant negative effects on other native workers. One interesting question that this paper cannot address is whether the contrast in conclusion with the immigration literature is due to the lack of demand shock associated with the double cohort or to potential differences between immigrants and native workers.

References

Aydemir, Abdurrahman and George J. Borjas, "Attenuation Bias in Measuring the Wage Impact of Immigration," *Journal of Labor Economics*, 2011, *29* (1), 69–112.

Beaudry, Paul, David A. Green, and Benjamin M. Sand, "The Great Reversal in the Demand for Skill and Cognitive Tasks," NBER Working Papers 18901 2013.

Berger, Mark C., "The Effect of Cohort Size on Earnings Growth: A Reexamination of the Evidence," *Journal of Political Economy*, 1985, *93* (3), 561–573.

——, "Demographic Cycles, Cohort Size, and Earnings," *Demography*, 1989, *26* (2), 311–321.

Borjas, George J., "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 2003, *118* (4), 1335–1374.

——, "Native Internal Migration and the Labor Market Impact of Immigration," *Journal of Human Resources*, 2006, *41* (2), 221–258.

——**, Richard B. Freeman, and Lawrence F. Katz**, "Searching for the Effect of Immigration on the Labor Market," *American Economic Review*, 1996, *86* (2), 246–51.

——**,** ——**, and** ——, "How Much Do Immigration and Trade Affect Labor Market Outcomes?," *Brookings Papers on Economic Activity*, 1997, *28* (1), 1–90.

Boudarbat, Brahim, Thomas Lemieux, and W. Craig Riddell, "The Evolution of the Returns to Human Capital in Canada, 1980-2005," *Canadian Public Policy*, 2010, *36* (1), 63–89.

Boustan, Leah Platt, Price V. Fishback, and Shawn Kantor, "The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression," *Journal of Labor Economics*, 2010, *28* (4), 719–746.

Bowlby, Geoff and Stéphanie Langlois, "High-Tech Boom and Bust," *Perspectives on Labour and Income*, 2002, *3* (4), 12–18.

Brunello, Giorgio, "The Effects of Cohort Size on European Earnings," *Journal of Population Economics*, 2010, *23* (1), 273–290.

Card, David, "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 1990, *43* (2), 245–257.

——, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

——, "Immigration and Inequality," *American Economic Review*, 2009, *99* (2), 1–21.

—— **and John DiNardo**, "Do Immigrant Inflows Lead to Native Outflows?," *American Economic Review*, 2000, *90* (2), 360–367.

—— **and Thomas Lemieux**, "Can Falling Supply Explain the Rising Return to College For Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

Carrington, William J. and Pedro J. F. de Lima, "The Impact of 1970s Repatriates from Africa on the Portuguese Labor Market," *Industrial and Labor Relations Review*, 1996, *49* (2), 330–347.

Chui, Tina, Kelly Tran, and Hélène Maheux, "Immigration in Canada: A Portrait of the Foreign-born Population, 2006 Census," Catalogue no. 97-557-XIE, Statistics Canada 2007.

Connelly, Rachel, "A Framework for Analyzing the Impact of Cohort Size on Education and Labor Earning," *Journal of Human Resources*, 1986, *21* (4), pp. 543–562.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg, "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, 2009, *124* (2), 843–881.

Friedberg, Rachel M., "The Impact of Mass Migration on the Israeli Labor Market," *Quarterly Journal of Economics*, 2001, *116* (4), 1373–1408.

—— **and Jennifer Hunt**, "The Impact of Immigrants on Host Country Wages, Employment and Growth," *Journal of Economic Perspectives*, 1995, *9* (2), 23–44.

Glitz, Albrecht, "The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany," *Journal of Labor Economics*, 2012, *30* (1), 175–213.

Goos, Maarten and Alan Manning, "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *Review of Economics and Statistics*, 2007, *89* (1), 118–133.

Green, David A. and Benjamin Sand, "Has the Canadian Labour Market Polarized?," Technical Report HRSDC 9548-10-0001, Human Resources and Skills Development Canada 2011.

Hunt, Jennifer, "The Impact of the 1962 Repatriates from Algeria on the French Labor Market," *Industrial and Labor Relations Review*, 1992, *45* (3), 556–572.

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

King, Alan J.C., Wendy Warren, Will Boyce, Peter Chin, Matthew King, and Jean-Claude Boyer, "Double Cohort Study: Phase 2 Report," Technical Report, Social Program Evaluation Group, Queen's University 2002.

—, —, —, —, —, —, and Barry O'Connor, "Double Cohort Study: Phase 3 Report," Technical Report, Social Program Evaluation Group, Queen's University 2004.

—, —, —, —, —, —, and —, "Double Cohort Study: Phase 4 Report," Technical Report, Social Program Evaluation Group, Queen's University 2005.

Korenman, Sanders and David Neumark, "Cohort Crowding and Youth Labor Markets: A Cross-National Analysis," in David G. Blanchflower and Richard B. Freeman, eds., *Youth Employment and Joblessness in Advanced Countries*, University of Chicago Press, 2000, pp. 57–105.

Kugler, Adriana and Mutlu Yuksel, "Effects of Low-Skilled Immigration on U.S. Natives: Evidence from Hurricane Mitch," Working Paper 14293, National Bureau of Economic Research 2008.

Lee, David S., "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 2009, *76* (3), pp. 1071–1102.

Lemieux, Thomas and Kevin Milligan, "Incentive Effects of Social Assistance: A Regression Discontinuity Approach," *Journal of Econometrics*, 2008, *142* (2), 807–828.

Macunovich, Diane J., "The Fortunes of One's Birth: Relative Cohort Size and the Youth Labor Market in the United States," *Journal of Population Economics*, 1999, *12* (2), 215–272.

Morin, Louis-Philippe, "Gender and Competition: Evidence from a Major Education Reform," 2011. Unpublished Manuscript.

——, “Estimating the Benefit of High School for College-Bound Students: Evidence of Subject-Specific Human Capital Accumulation,” *Canadian Journal of Economics*, 2013, *46* (2), 441–468.

Ottaviano, Gianmarco and Giovanni Peri, “Immigration and National Wages: Clarifying the Theory and the Empirics,” Working Paper 14188, National Bureau of Economic Research 2008.

Peri, Giovanni, “Rethinking the Area Approach: Immigrants and the Labor Market in California,” *Journal of International Economics*, 2011, *84* (1), 1–14.

Stapleton, David C. and Douglas J. Young, “Educational Attainment and Cohort Size,” *Journal of Labor Economics*, 1988, *6* (3), pp. 330–361.

Statistics Canada, “2006 Census Dictionary,” Statistics Canada Catalogue no. 92-566-XWE 2007.

——, “2006 Census Code Book,” Census Operations Division, Statistics Canada 2008.

Welch, Finis, “Effects of Cohort Size on Earnings: The Baby Boom Babies’ Financial Bust,” *Journal of Political Economy*, 1979, *87* (5), S65–S97.

Appendix A

Data Construction and Identification of the Treatment Group

I first describe how the data was constructed from the Canadian Census and the Labour Force Survey and then discuss how the 'treatment' group was identified using the available information.

Census Data

The main outcome variable is the average real weekly wages (in 2000 dollars). This variable is constructed using the WAGES variable from the 2001 and 2006 Censuses and Statistics Canada's provincial consumer price index (CPI). The WAGE variable consists of gross wages and salaries before deductions (e.g., income tax), and it includes commissions and cash bonuses. The WEEKS variable is used to convert the annual wages into weekly wages, and I classify a worker as full- or part-time using the FPTIM variable. An individual is considered to be working full time if she "worked mainly full-time weeks" (i.e., 30 hours or more) in the year prior to the Census (Statistics Canada, 2007). Finally, an individual is considered to be working full time, full year if she worked 48 weeks or more in the year prior to the Census. I restrict the sample to full-time, full-year workers.

Since the outcome of interest is weekly wages in the year prior to the Census, I assign respondents to their province of residence as of June of the previous year (PR1 variable). Individuals that lived out of the country in that year (about 0.8 percent of the sample of individuals aged between 20 and 50) are discarded from the analysis.

Individuals that attended school in the year prior to the Census are excluded. They are identified using the 2001 Census ATTENDR and the 2006 Census ATTSCHSUM variables. The school attendance indicator variable is equal to one if the individual attended school between September and May prior to the Census, regardless of whether the individual attended to school part-time or full-time. It is not possible to differentiate part-time and full-time attendance in the 2006 Census.

I construct the age of the respondent on January 1$^{\text{st}}$ of the Census year using their birth date. Since the last cohort of Ontario's Grade 13 program are expected to be 21 on January 1$^{\text{st}}$ 2006, I restrict the sample to individuals aged 21, and individuals aged between 26 and 50.

The educational-attainment variable is constructed using the 2001 Census SECGRADR and the 2006 Census SSGRAD variables. It corresponds to the highest educational degree obtained by the individual. A high school graduate in this paper is defined as an individual "with high school certificate or equivalency certificate without further schooling" (Statistics Canada, 2007).

Finally, I use the Class of Worker variable (COWD) to identify self-employed workers. Self-employed are excluded from the analysis since their wage-setting process is different from paid workers. The number of self-employed is very small, especially among workers aged 21. A detail about COWD that could introduce some measurement error is the fact that the question relates to labour market activity on the month of, instead of on the year prior to the Census. Including these workers does not affect the results.

Labour Force Survey Data

The main outcome variable is the real weekly wages (in 2000 dollars). The conversion from current to real wages is done using Statistics Canada's provincial CPI. One significant difference between the Census and Labour Force Survey (LFS) wages is that the LFS (hourly and weekly) wages are directly observed, in the sense that they are not estimated using the number of weeks worked in a year. The wage variables are observed only for employees.

As in the Census data, an individual is considered to be working full time if she usually works 30 hours or more per week at her main job. I restrict the sample to full-time workers.

I determine whether someone is attending school in the survey month using STUDENT. Full- and part-time students are excluded from the sample.

The LFS does not release the respondent's date of birth. I therefore use age of the respondent on the week of the survey to define my age groups and to identify the treatment group. As will be explained below, the January LFS is the only one allowing me to identify the treatment group. Since the last cohort of Ontario's Grade 13 program is expected to be aged 19 on January 1^{st} 2004, I restrict the sample to individuals aged 19, and individuals aged between 24 and 48 to be consistent with the Census.

The LFS educational-attainment variable is constructed using two variables, EDUCLEV and HSGRAD. A high school graduate in the LFS data is defined as an individual who completed 11 to 13 years of schooling (based on EDUCLEV) and who graduated from high school (HSGRAD). This measure is somewhat cleaner than the Census measure, especially given the fact that it did not change between 2002 and 2004.

Finally, I exclude self-employed workers from the LFS data using COWMAIN. This variable identifies the class of worker at the respondent's main job.

Identification of the Treatment Group

The identification of the treatment group in the Census data is easy since it contains the exact date of birth of the individuals. Since the cut-off birth date for beginning primary school is December 31st in Ontario, one only needs to know the year of birth of an individual to know if she was supposed to be part of the double cohort or not. Graduates from the last Grade-13 cohort are expected to be born in 1984, while graduates from the first cohort of the Grade-12 program should be born in 1985. In order to avoid having the results contaminated by the potential value-added of Grade 13, I exclude Grade 12 graduates.

In the Census data, the treatment group is hence defined as 2006 Census respondents who were: 1) born in 1984, 2) high-school graduates, and 3) Ontario residents in 2005. The main analysis is done on full-time and full-year workers (and not enrolled in school).

Since the LFS does not release the date of birth of their respondents, one has to rely on age only. I use the January LFS since it is the only one that allows me to get a good measure of one's date of birth. In January 2004, almost all LFS respondents aged 19 should be born in 1984, corresponding to the birth year of the last cohort of Ontario Grade 13 graduates.

In the LFS data, the treatment group is composed of January 2004 respondents who were both: 1) 19 year old in the survey week, and 2) Ontario high-school graduates. The main analysis is done on individuals who worked full time, and did not attend school in January 2004.

Cohort Size

I use the August LFS to estimate the annual cohort size numbers presented in Figure 1.[32] Between the months of May and August, LFS respondents aged 15 to 24 are asked if they were in school in the previous March. Since students can graduate from secondary school during the summer, the August survey has the advantage of including many recent graduates, giving a better picture of the expected increase in labour supply to come. In this paper, a cohort of graduates is composed of 17, 18 and 19 year-old individuals who had graduated from high school when first interviewed by the LFS, and who were full-time students in March of the same year in a secondary school institution. I estimated cohort sizes using recent high-school graduates aged 17 to 19 to include both Grade 12 and Grade 13 graduates.

[32]The Ministry of Education and Training grants diplomas at any time during the year to students who have successfully completed the necessary secondary school requirements. Hence, there is no specific month where all eligible students graduate from high school. Nevertheless, most students complete Ontario Secondary School Diploma (OSSD) requirements by the end of the spring.

Appendix B – Additional Tables

Table B.1: Census Results Using Workers with a Bachelor's Degree as the Control Group (Weekly Wages for Full-Time, Full-Year Workers)

	(1)	(2)	(3)	(4)
DC × Youth × ON	-0.056**	-0.055**	-0.045**	-0.055**
	(0.024)	(0.024)	(0.023)	(0.022)
Province Fixed Effects	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	No	Yes	Yes	Yes
Industry Fixed Effects	No	No	Yes	Yes
Personal Characteristics × Youth Fixed Effects	No	No	No	Yes
Industry × Youth Fixed Effects	No	No	No	Yes
Industry × DC Fixed Effects	No	No	No	Yes
R^2	0.29	0.31	0.37	0.38
N	35,265	35,265	35,265	35,265

Notes: The dependent variable is the ln of real weekly wages. The sample consists of individuals who worked 48 weeks or more during the year prior to the Census and worked full time during these weeks. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age on January 1st of the Census year and had a high school diploma, but no further schooling. The omitted provincial dummy variable is Quebec. The control group consists of individuals that were between 21 and 25 years of age on January 1st of the Census year, and had a bachelor's degree, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The industry fixed effects reflect 20 sectors of activity (based on NAICS). The estimation was done using Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table B.2: Double Cohort and Good Occupations

Age Group	26–30	31–35	36–40	41–45	46–50
DC × Youth × ON	0.003	-0.001	0.009	-0.000	-0.011
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Province Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
Province × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Controls for Personal Characteristics	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Personal Characteristics × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × Youth Fixed Effects	Yes	Yes	Yes	Yes	Yes
Industry × DC Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.18	0.18	0.18	0.18	0.19
N	84,375	94,180	117,050	138,145	125,935

Notes: The dependent variable is a dummy variable equal to 1 if the individual worked in a high-paying occupation according to the 2001 Census. An occupation is considered as high paying if its average weekly wage was in the top quartile of the occupation wage distribution. 'Youth' is an indicator variable that equals 1 if the individual was 21 years of age. All sampled individuals have a high school diploma, but no further schooling. The personal characteristics include: gender, immigrant status, a visible minority indicator, a rural area indicator, and multiple indicators for marital status. The estimation was done using the Census weights. The numbers of observations are rounded to a base of 5. Robust standard errors are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.