2018-20

The Effect of College Education on Health and Mortality: Evidence from Canada

Guy Lacroix
Francois Lalibertée-Auger
Pierre-Carl Michaud
Daniel Parent

Décembre / December 2018

**Centre de recherche sur les risques
les enjeux économiques et les politiques publiques**

www.crrep.ca

Abstract

We investigate the returns to college attendance in Canada in terms of health and mortality reduction. To do so, we first use a dynamic health microsimulation model to document how interventions which incentivize college attendance among high school graduates may impact their health trajectory, health care consumption and life expectancy. We find large returns both in terms of longevity (4.1 years additional years at age 51), reduction in the prevalence of various health conditions (10-15 percentage points reduction in diabetes and 5 percentage points for stroke) and health care consumption (27.3% reduction in lifetime hospital stays, 19.7 for specialists). We find that education impacts mortality mostly by delaying the incidence of health conditions as well as providing a survival advantage conditional on having diseases. Second, we provide quasi-experimental evidence on the impact of college attendance on long- term health outcomes by exploiting the Canadian Veteran's Rehabilitation Act, a program targeted towards returning WW-II veterans and which incentivized college attendance. The impact on mortality are found to be larger than those estimated from the health microsimulation model (hazard ratio of 0.216 compared to 0.6 in the simulation model) which suggest substantial returns to college education in terms of healthy life extension which we estimate around one million canadian dollars.

JEL Codes: I14, I26
Keywords: mortality, education, microsimulation, quasi-experimental, instrumental variables, veterans

Guy Lacroix
Francois Lalibertée-Auger
Pierre-Carl Michaud
Daniel Parent

Corresponding author: Pierre-Carl Michaud, Department of Applied Economics, HEC Montréal, 3000 chemin Côte-Ste-Catherine, Montréal, H3T2A7. We acknowledge funding from the National Institute of Aging trough grant R01AG040176-06. We are grateful to Statistics Canada trough the Quebec Interuniversity Center for Social Statistics (QICSS) for granting access to the data used for this project. We thank Sébastien Box-Couillard and Aurélie Coté- Sergent for excellent research assistance. Part of the analysis presented in this paper was conducted at the Quebec Interuniversity Centre for Social Statistics which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the QICSS are made possible by the financial or in-kind support of the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institutes of Health Research (CIHR), the Canada Foundation for Innovation (CFI), Statistics Canada, the Fonds de recherche du Qubec - Socit et culture (FRQSC), the Fonds de recherche du Qubec - Sant (FRQS) and the Quebec universities. The views expressed in this paper are those of the author(s), and not necessarily those of the CRDCN or its partners. All errors are our own.

1 Introduction

Large differences in health outcomes, in particular mortality, are observed across education groups (and more generally socio-economic status) and there is some evidence that these differences are widening (Goldman and Smith, 2011; Tjepkema et al., 2012). An important question is whether education causally improves health instead of the association being due to other factors or, more simply, reflecting a case of reverse causality (Fuchs, 1980; Smith, 1999). This has important implications for assessing the effect of education policies on health and health care utilization, as well as determining optimal subsidies for college education (which should be increasing in the social return to education).

Using quasi-experimental variation in education, a burgeoning literature has attempted to assess whether education improves health and lowers mortality. Galama et al. (2017) provide a critical review of the literature. The nature of the quasi-experimental variation in education used in those studies is important, in particular the level of education at which variation is induced. For example, variation in high school education is likely to induce effects which are different from those arising from variation in college education. A few randomized controlled trials have been conducted in the U.S. to estimate the effect of early education on later life health. For example Conti et al. (2016) reports effects on adult health from the Perry Pre-School project and the Carolina Abecedarian Project. Many studies look at high school education. For example, changes in mandatory schooling have been exploited to estimate the effect of education on health and mortality. While some studies such as Lleras-Muney (2005) and van Kippersluis et al. (2009) find large effects for the U.S. and the Netherlands, other studies point to little or no effect of mandatory high school and primary schooling on mortality (Albouy and Lequien, 2009; Clark and Royer, 2013; Meghir et al., 2018). But much less is known about the effect of college education on health and mortality. For example, Buckles et al. (2016) use draft avoidance in the Vietnam war and find significant effects on mortality of those cohorts that were affected.

To the extent that the relationship between education is causal, a related question involves identifying through which channels education improves health and reduces mortality. A good starting point is the health capital model of Grossman (1972). In that model, initial conditions and investments in health over the life-cycle guide the dynamics of health. Investments react to changes in economic resources as well as various inputs such as education. Since education affects initial conditions, economic resources and the inputs into the production of health, it can affect health at older ages through a multitude of pathways. For example, education could reduce mortality because it induces changes at the beginning of adult life that have long-lasting effects:

for example by reducing smoking or obesity prevalence. It could also reduce the incidence of various chronic diseases through better prevention or life styles (investments). Finally, it could improve survival conditional on having various diseases, trough better self-management of chronic diseases or better use of the health care system. Hence, understanding the pathways through which education may impact mortality is paramount.

In this paper, we make two contributions that shed light on the relationship between college education and health/mortality. First, using a flexible dynamic microsimulation model for Canada, we decompose the mortality differences by education into differences that arise from differences in initial conditions, incidence of diseases, and finally survival conditional on having diseases. We can attribute differences to these three components and to particular diseases, and also investigate differences in health care use. While this descriptive evidence is interesting, it is unclear whether it is causal. We then ask the question whether these simulated differences in mortality and college education in the microsimulation model match the reduced-form causal effect of college education on old age mortality using quasi-experimental evidence from Canada. We exploit the Canadian Veteran's Rehabilitation Act – the Canadian "G.I. Bill", a program targeted towards returning WW-II veterans and which incentivized college attendance. Since both exposure to the war and educational attainment before the war create differential exposures the program, we exploit differences in education and mortality across Canadian provinces to estimate these effects. We make use of a unique match of the 1991 Canadian Census to the Mortality Registry in Canada 1991-2011 and estimate flexible Cox-proportional hazard models and competing risk logits for cause of death using a control function approach.

The paper is structured as follows. In section 2, we present a useful framework for thinking of life-cycle differences in health by education. In section 3, we use a microsimulation model to uncover some of the pathways trough which education translates into mortality differences. In section 4, we provide quasi-experimental evidence on the long-term effect of education on mortality. Finally, we conclude in section 5.

2 Life-Cycle Trajectory of Health and Education

Denote by E the education outcome of an individual at some initial age, normalized to be $t = 0$. The initial vector of risk factors, such as smoking habits and obesity is denoted Z_0. Finally denote by H_0 the vector containing the set of health conditions already diagnosed at $t = 0$. The joint distribution of these initial conditions, given education, is denoted $F_E(Z_0, H_0) = F(Z_0, H_0|E)$. Two sets of markovian equations may capture the evolution of Z_t and H_t over

the life-cycle. Let those equations be:

$$H_{t+1} = \psi_t(H_t, Z_t, E) \tag{1}$$

$$Z_{t+1} = \phi_t(H_t, Z_t, E), \tag{2}$$

where ψ_t and ϕ_t are age-specific (vector) functions of the state variables and some random exogeneous disturbance. Define mortality as $M_t = 1$ if dead and zero if alive. Let mortality be defined as a function of the state variables:

$$M_{t+1} = \xi_t(H_{t+1}, Z_{t+1}, E). \tag{3}$$

The vectors (H_{t+1}, Z_{t+1}) are only observed provided $M_{t+1} = 0$, and $M_{t+1} = 1$ if $M_t = 1$ (death is an absorbing state).

Denoting $\mathbb{E}$ as the expectation operator, long-term differences in mortality rates between education level E and E',

$$\Delta_{E',E} = \mathbb{E}[M_{t+s}|E'] - \mathbb{E}[M_{t+s}|E],$$

are a compendium of three forces: initial conditions differences, $F_E(H_0, Z_0)$ and $F_{E'}(H_0, Z_0)$, incidence rate differences in health and risk behaviors as a function of education, ψ_t and ϕ_t, and finally differences in conditional mortality rates ξ_t by education. In section 3, we use a microsimulation model to estimate long-term differences in mortality rates and decompose these differences in terms of the above three sources. In section 4, we take a different route to estimate the long-run relationship between mortality and education by studying the function:

$$M_{t+s} = \alpha_{t+s}(E, H_0, Z_0) \tag{4}$$

for s large. We do so by exploiting a policy change which plausibly holds initial conditions fixed while varying E. Both of these approaches have merit. While the microsimulation approach allows to identify pathways, the unbiased estimation of ψ_t, ξ_t and ϕ_t is necessary to obtain unbiased estimates of mortality differences at older ages and is tenuous given limited panel data. If unobserved frailty is correlated with education, for example trough selection into education due to frailty, then differences $\Delta_{E',E}$ do not provide an estimate of the causal effect of additional

education. On the other hand, while the quasi-experimental approach allows for the estimation of a causal effect, it does not allow to identify specific pathways. Furthermore, it does so for a specific policy, or source of variation, which may not translate into an estimate of the average effect of providing education to a randomly chosen individual in the population (Imbens and Angrist, 1994).

3 Decomposing Mortality Differences by Education over The Life-Cycle

Health trajectories of Canadian men aged 30 and 31 are simulated for the period from 2016 until death using COMPAS, a health microsimulation model which implements the dynamic structure presented in section 2. The model shares many features with the one developed for the U.S. by Goldman et al. (2004), but was adapted to the Canadian context by Laliberté-Auger et al. (2017).[1]

COMPAS has four parts. The base population is first generated using Statistics Canadas 2008-2009 and 2010 Canadian Community Health Survey (CCHS), which each contains a cross-section of over 60,000 records of individuals living in private households for a total of over 120 000 individuals. Each record is weighted to ensure that the base population is representative of the Canadian population in 2010. The CCHS contains information on individual characteristics (gender, age, educational attainment, immigrant status), health status (diagnosis of diabetes, hypertension, stroke, cancer, heart disease, lung disease or dementia), major risk factors (tobacco use and obesity), and disabilities (activity limitations and cognitive impairment). Hence, for the youngest individuals aged 30-31, it provides an initial distribution $F_E(X_0, Z_0)$.

The second part consists in modelling the transitions between various health statuses, the functions ϕ_t ,ψ_t and ξ_t. To that end, COMPAS uses data from Statistics Canada's National Population Health Survey (NPHS) for the years 1994-2011, a bi-annual panel survey. The NPHS comprises 17,276 respondents who were interviewed every second-year between 1994 and 2011. The NPHS is based upon a similar sampling frame as the CCHS. Developing a disease, being admitted into a long-term care facility, starting or quitting smoking, and dying are examples of changes in health status that are modelled jointly, *i.e.* transition probabilities differ for every combination of potential health statuses. In addition, the NPHS contains detailed information on activities of daily living limitations (ADL), instrumental activity of daily living

[1] COMPAS' code and full technical documentation are available in a public repository. See Laliberté-Auger et al. (2017).

limitations (IADL) and cognitive impairment.[2] Complementary log-log models are used to estimate the transition probabilities –or incidence– for each of the seven diseases considered in the model, as well as the mortality transition, as a function of age, major risk factors, other socio-demographic characteristics (gender, education attainment, immigrant status) and, in certain clinically plausible cases, the prior occurrence of other illnesses. Assumptions about the potential interactions between various diseases are based upon medical research and were validated by a panel of experts as part of the Future Elderly Model (FEM) project (Goldman et al., 2004).[3] A multinomial logistic regression model is use to estimate the transition probability between the different combinations of disability states.

In the third part, COMPAS uses data from CCHS to project health care use. The number of night of hospitalization, the number of general practitioner and specialist visits are modelled using a negative binomial model in order to account for the strong proportion of zero in utilization. Models are function of individual characteristics and health status predicted in the second phase but it is not a function of past utilization, so in this sense it can be seen as representing cross-sectional utilization.

Typically, in the final part, new cohorts enter the simulation in each period at starting ages between 30 and 31. Stages 2,3 and 4 are then repeated every 2-year, and the last cohort is added in year 2050. The model is almost entirely based on population dynamics: it tracks individuals as they move from one two-year period to the next. In each simulation cycle, the population varies randomly due to an outgoing mortality flow (or because some may have reached the maximum allowed age of 120) and an incoming flow. The model also incorporates recent trends in tobacco use, obesity and educational attainment and an entire set of scenarios to make population projections. For this paper, we focus on a single cohort of men aged of 30-31 years old in 2016 until death. This cohort approach is well suited to investigate life course changes in health and pathways through which education ultimately impacts mortality. Results using any other entering cohort from COMPAS would be very similar.

Because of the randomness associated with these simulations, 50 replications of the projection process are performed for each of the scenarios with 100 sets of possible values for the parameters of the underlying econometric models (100 bootstrap samples are used to estimate transition model parameters). Hence, a total of 5,000 replications are performed, and the estimates obtained are averaged over those replications. We also report the 5th and 95th percentiles

[2] In what follows, we refer to ADL, IADL and cognitive impairment as "disability".

[3] Information about these assumptions, as well as the data and all econometric models used in COMPAS are described in details in Laliberté-Auger et al. (2017), as are details about the modelling of mortality as well as back-casting and validation of the transition models.

of those 100 sets of simulations.

3.1 Simulation results

The parameter estimates of COMPAS are obtained from representative samples of the Canadian population, as mentioned above. Table 1 reports the mean values of some health characteristics of the initial population in COMPAS by education levels (the distribution $F_E(X_0, Z_0)$). The statistics refer to males aged 30 and 31 in 2016. As expected, proportionately more high school graduates smoke and have weight problems. In some cases the differences are sizable and in particular with respect to smoking habit. Thus, once COMPAS is rolled out, twice as many high school as university graduates are active smokers and a third more exhibit weight problems initially. COMPAS attempts to replicate the expected health trajectories of this group of men relatively precisely.

Table 1: Health Characteristics of Initial Population

Variables	High School	University
Current smoker	0.37	0.19
Former smoker	0.37	0.39
Never smoked	0.26	0.42
Non obese (BMI <30)	0.76	0.84
Obese (30 $\leq$ BMI <35)	0.17	0.12
Very obsese (BMI $\geq$ 35)	0.07	0.04

We compare lifetime outcomes of college educated and those with high school education. This amounts to allowing education to impact future transitions both trough a direct effect and an effect stemming from different initial conditions (rates of smoking and obesity are higher in the high school group). Trough transitions which generate the incidence of new health conditions, education changes risk factors for other transitions. This chain reaction culminates with differences in mortality at older ages as outlined in section 2. We first focus on life expectancy and mortality as reported in Table 2. The first column reports these vital statistics for base case high-school graduates. The second column reports the statistics of the counterfactual health trajectories when high school graduates attend college. The last column computes the difference between the two, either in terms of ratio (line 1) or differences (lines 2-4). The simulation results indicate that earning a university degree decreases the average annual mortality between the ages of 51 and 96 by two percentage points. The difference is statistically significant and corresponds to an odds-ratio of death of 0.61 in favour of the college graduates. Tjepkema et al. (2012) find using Canadian census data merged with death records (the same data we

use latter but over a shorter period), a age-standardized hazard ratio of 0.767 for ages 25 and older. The lower mortality rates we find among college graduates translate into an additional life expectancy of 4.1 years at age 51. COMPAS further allows computing the expected number of disability-free and disability-ridden years. College education is found to increase the former by 5.5 years and decrease the latter by 1.3 years. Both differences are statistically significant. For the U.S., Meara et al. (2008) reports a difference of 5.4 years in 1990 between those with high school or less and any college at age 25. For European countries, Kunst and Mackenbach (1994) report among those 55-64 in the 1970s and 1980s a hazard ratio of 0.73 for Sweden, 0.55 for Finland and 0.438 for France. Hence, our simulated hazard ratios are within the range of what has been reported elsewhere for different countries.

Table 2: Mortality Rates and Life Expectancy

Variables	High school	University	Variation
Mortality rate, 51–96	0.054 [0.049 ; 0.058]	0.031 [0.026 ; 0.036]	0.606 [0.487 ; 0.710]
Life expectancy at 51	35.2 [34.6 ; 35.8]	39.3 [38.3 ; 40.6]	4.1 [2.8 ; 5.6]
Life expectancy with disability at 51	6.7 [6.2 ; 7.4]	5.4 [4.5 ; 6.3]	-1.3 [-2.3 ; -0.4]
Life expectancy without disability at 51	28.5 [27.8 ; 29.2]	33.9 [32.7 ; 35.4]	5.5 [4.2 ; 7.0]

Disability: at least one of the following: ADL, IADL and cognitive impairment

COMPAS allows to compute health care use and extent of care conditional on health status. Hence, the counterfactual simulations yield different utilization profile since diseases and other conditions change as a result of the change in education. Table 3 reports the expected number of hospital stays and number of visits to general practionners (GP) and specialists by education. Not surprisingly given the previous results, high school graduates are found to have more numerous hospital stays (+ 27.3%) and consultations with specialists (+19.7%), but are no more likely to consult with GP's. These differences are large and potentially lead to large gaps in lifetime use of the health care system.

Table 3: Resource Utilization

Variables†	High school	University	Variation
# hospital stays	100 [91 ; 108]	73 [61 ; 83]	-27.3 [-40.2 ; -15.6]
# visits to GPs	100 [97 ; 103]	102 [96 ; 107]	1.8 [-4.7 ; 7.7]
# visits to Specialists	100 [96 ; 105]	80 [75 ; 84]	-19.7 [-26.2 ; -13.8]

† High-school normalized to 100

Greater life expectancy is associated with more disability-free years as shown above. ADL

limitations may be intimately related to the occurrence of various diseases. Health transitions in COMPAS consider up to seven different diseases. Below, we report the incidence at ages 65 and 80 of six of them. As shown, college graduates are less prone to cancer, heart disease, hypertension, diabetes, stroke and dementia than high school graduates. The differences are slightly more pronounced at age 80. Most of these are statistically significant, although the confidence intervals between high school and college graduates overlap slightly in a few cases.

Table 4: Disease Prevalence

Variables	High school	University	Variation
Cancer at 65	0.04 [0.04 ; 0.06]	0.03 [0.02 ; 0.04]	-0.01 [-0.02 ; -0.00]
Cancer at 80	0.06 [0.05 ; 0.08]	0.05 [0.03 ; 0.06]	-0.02 [-0.03 ; -0.00]
Heart disease at 65	0.16 [0.14 ; 0.17]	0.14 [0.11 ; 0.17]	-0.02 [-0.05 ; 0.01]
Heart disease at 80	0.31 [0.27 ; 0.34]	0.28 [0.23 ; 0.33]	-0.03 [-0.07 ; 0.03]
Hypertension at 65	0.63 [0.60 ; 0.65]	0.56 [0.52 ; 0.60]	-0.06 [-0.11 ; -0.03]
Hypertension at 80	0.83 [0.80 ; 0.85]	0.78 [0.73 ; 0.82]	-0.05 [-0.10 ; -0.02]
Diabetes at 65	0.26 [0.23 ; 0.29]	0.16 [0.13 ; 0.20]	-0.10 [-0.14 ; -0.06]
Diabetes at 80	0.40 [0.37 ; 0.44]	0.26 [0.21 ; 0.32]	-0.14 [-0.20 ; -0.09]
Stroke at 65	0.03 [0.03 ; 0.04]	0.01 [0.01 ; 0.02]	-0.02 [-0.03 ; -0.01]
Stroke at 80	0.10 [0.08 ; 0.12]	0.05 [0.03 ; 0.07]	-0.05 [-0.08 ; -0.02]
Dementia at 65	0.01 [0.01 ; 0.02]	0.01 [0.01 ; 0.01]	-0.01 [-0.01 ; -0.00]
Dementia at 80	0.12 [0.10 ; 0.15]	0.08 [0.06 ; 0.11]	-0.04 [-0.07 ; -0.01]

The previous simulation results are consistent with the empirical regularity according to which education and health are positively correlated. In particular, the results of Table 2 show that the differences between high school and college graduates in terms of mortality rates and life expectancy can be quite large. Yet, these simulations do not allow to distinguish between the impact of having a college degree *per se* and the contributions of the differing initial conditions between the two educational groups $F_E(X_0, Z_0)$, nor do they measure the contributions of the education-specific incidence rates between the various diseases and the occurrence of ADLs (ψ_t and ϕ_t) or different mortality rates conditional on disease, (ξ_t). In order to identify the contributions of each of these specific components, we run a series of additional simulations in which we block pathways sequentially, initial conditions, incidence and survival conditional on prevalence of diseases. For example, an intitial condition simulation changes $F_E(X_0, Z_0)$ for $F_{E'}(X_0, Z_0)$ but keep other components to the model with education at E. We report the results of this decomposition in Table 5. The first column replicates the last column of Table 2, *i.e.* the total impact of earning a college degree. The second column reports the results of simulating the health trajectories of high school graduates while assuming they share the same initial conditions as college graduates (smoking incidence, obesity, *etc.*). The third column assumes that high school graduates have the same risk of contracting a given disease as college graduates, effectively changing education level in the set of functions ϕ_t and ψ_t. Finally, column 4 reports

the results of assuming that high school graduates share the same mortality and disability risks as college graduates, and that of requiring long-term care, conditional on having a particular set of diseases and health conditions. Hence, it only varies ξ_t.

Table 5: Decomposition of the Impact of a University Degree

Variables	Total[1]	Initial[2] Conditions	Diseases[3] Incidence	Mortality[1] & Disability
Hazard ratio 51-96	0.61 [0.49;0.71]	0.97 [0.95;0.99]	0.82 [0.78;0.87]	0.73 [0.59 ;0.86]
Life expectancy at 51	4.1 [2.8;5.6]	0.2 [0.1;0.4]	1.5 [1.1;1.9]	2.5 [1.3;3.8]
Life expectancy with disability at 51	-1.3 [-2.3;-0.4]	-0.0 [-0.1;0.1]	-0.3 [-0.6;-0.1]	-0.8 [-1.7;0.2]
Life expectancy without disability at 51	5.5 [4.2;7.0]	0.3 [0.1;0.4]	1.8 [1.4;2.4]	3.2 [2.3;4.5]

The table shows unequivocally that the health differentials between high school and college graduates have little to do with the initial conditions. Indeed, at baseline, high school graduates are more likely to smoke, to suffer from obesity, and to have various diseases. Yet, these differences have next to no impact on health trajectories in the long run. Thus, the relative risk of death between the ages of 51 and 96 is only 3 percentage points lower when high school graduates are assigned their initial conditions rather than those of college graduates as shown in column 2. Furthermore, life expectancy with or without ADL limitations is insensitive to the assumed initial conditions. This suggests that poorer health conditions at the age of 30-31 have little impact on long run life outcomes. On the other hand, the third column reveals that the differential incidence of various diseases plays a major role in explaining life expectancy. Thus when high school graduates are assumed to face the same transition probabilities as college graduates across diseases, their life expectancy increases by 1.5 years and that without ADL limitations by over 1.8 years. Likewise, column 4 shows that the differential probability of death and disability between high school and college graduates is the main driver of the gap in life expectancy. Hence, when assigning a college degree to high school graduates, their relative death ratio between the ages of 51 and 96 is only 73% that of its original value, and their life expectancy increases by 2.5 years (3.2 years without ADL limitations).

The COMPAS microsimulation model is based upon various high-quality panel survey and administrative data sets. It captures most features of the dynamic transitions between healthy and disease-ridden spells of a representative ageing Canadian sample. The relation between education and life expectancy it unearths is driven by these random but conditional (on observable characteristics) transitions. An alternative approach to investigate the link between the two is

to focus on the long term consequences of a large-scale policy innovation that may have induced a given population to increase its educational level. In Canada, such a policy was implemented after World War II and was designed to ease enrollment of returning veterans into vocational and university programs. Fortunately, recently released files linking census and mortality data allow us to track cohorts of potential returning veterans into old age and in many cases until death. In what follows, we briefly overview the Veterans Rehabilitation Act (VRA) and provide *prima facie* evidence on its likely impact on college attendance. We next attempt to link the shift in the latter with mortality data. By doing so, we provide additional empirical support about the correlation between education and mortality.

4 Quasi-Experimental Evidence

Identifying the causal impact of education on life expectancy is a demanding task. Indeed, exogenous policy-induced changes in schooling attainment must be investigated long after their implementation and thus require quality data on health outcomes or death. Yet, this is precisely the empirical strategy used by most papers in the literature. Thus Buckles et al. (2016) exploit exogenous variations in years of completed college induced by draft-avoidance behavior during the Vietnam War to examine the impact of college on adult mortality. They find that increasing college attainment from the 25^{th} to the 75^{th} percentile of the education distribution would decrease the cumulative mortality by 8 to 10 percent relative to the mean. Lleras-Muney (2005), using compulsory education laws from 1915 to 1939 as instruments for education find that education has a relatively large causal impact on mortality in the U.S. Likewise, Oreopoulos (2006) finds that an additional year of compulsory schooling lowers the likelihood of reporting a physical or mental health disability by 1.7 percentage points and that of reporting a disability that limits one's daily activity by 2.5 percentage points. Similar results are found to hold for the U.K. (e.g. Cutler and Lleras-Muney (2010)). However, Clark and Royer (2013) exploit two changes to British compulsory schooling laws that induced sharp cross-cohort differences in educational attainment. Unlike the previous papers, they find no evidence of a causal link between the reforms and health.

Unfortunately, there exists very little evidence on the causal link between life expectancy and education in Canada. Yet, Tjepkema et al. (2012) provide convincing evidence that age standardized mortality rates are highest among people with less than secondary graduation and lowest for university degree-holders. Thus, while there exists empirical evidence to the effect that education and life expectancy are correlated in Canada, there is yet no evidence of a causal

link in the literature. In what follows, we exploit a large policy change that was implemented after World War II to incentivize returning veterans to further their schooling (see Lemieux and Card, 2001).

4.1 The G.I. Bill

As described in Lemieux and Card (2001), following the adoption of Veterans Rehabilitation Act (VRA) by the Canadian government, returning veterans who wished to enroll in either a vocational training program or a university program were provided financial assistance. Thus, honorably discharged veterans who chose to enter university had their tuition fully paid for and received a monthly living allowance of approximately $CAN 650 in today's dollars, or $60 in 1946, if they had no dependents (more otherwise). The benefits were available for up to four years. Although only veterans who volunteered into the armed forces were eligible for the financial assistance, it made little difference as the vast majority of the enlisted men in Canada were volunteers.

Interestingly, French-speaking men from the Province of Quebec were much less likely to participate in the war effort. Indeed, fewer than 20% of eligible men in their late teens or early twenties served during the war whereas approximately 50% of similarly aged men did so in Ontario. Lemieux and Card (2001) state a number of reasons that may explain why there were proportionately fewer French-speaking servicemen. First, with the exception of a few French-speaking infantry regiments, the Canadian armed forces were first and foremost unilingual English. Since most of the French-speaking population of Quebec did not speak English, the infantry was their only option. Second, most French Canadians had very weak ties to France and so were reluctant to volunteer abroad. As reported in Lemieux and Card (2001), the proportion of U.K. immigrants living in Ontario and the province of British Columbia was relatively high. French immigration to Quebec, on the other hand, had stopped in the mid-eighteenth century so that ties to France were very limited.[4]

In addition to having proportionately fewer servicemen, the educational attainment among French speakers in Quebec at that time was considerably lower than that of English speaking Canadians. Thus, for the cohort of men who would have been eligible to participate in WWII, the modal number of years of schooling among French speaking Quebec men was five

[4]Indeed, ever since the French colonies in New France were ceded to Great Britain following the Seven Years' War, the population in French Canada had no special attachment to Britain, as opposed to the English-speaking population with British ancestry in other provinces. In addition, by the time of the Second World War, French Canadians had more or less been cut off from France for close to two hundred years, thereby severely limiting the kind of attachment felt towards the old country that was felt elsewhere in English Canada.

Figure 1: University attendance (%) by year of birth

or six, compared to eleven or twelve for their English speaking counterpart. Additionally, while under-prepared returning veterans from Ontario benefited from remedial programs, no such accommodation existed in Quebec. Consequently, even for those who volunteered in Quebec, the fact that their schooling level was well below high school made the VRA financial assistance program essentially irrelevant. The combination of lower educational attainment and lower participation rates in the war effort makes French Canadian men a potentially adequate control group for returning veterans from English-speaking Ontario.

Figure 1 plots the proportion of men with a least one year of university education by year of birth and province. Men born between 1921 and 1927 were potential candidates for overseas service. A first noteworthy feature of the figure concerns the (much) lower educational attainment of French-speaking men, irrespective of their birth year. A second one concerns the dramatic increase in university enrollment among Ontario men born between 1918 and 1928, and especially among those born between 1921 and 1927. These men were between 18 and 24 during the war and they all qualified for the VRA program upon returning home. No such increase in similarly aged French-Canadians is observed in the figure. Two conclusions can be drawn from the figure. First, one can legitimately claim that the program indeed incentivized many veterans to attend university.[5] Second, it can also be legitimately claimed than French-Canadians constitute an

[5]This is especially true given the fact that there is little indication that the bulge in university full-time enrollment

appropriate control group for English-Canadians.

4.2 Empirical Strategy

Following the above discussion, we proceed to estimate the impact of the exogenous policy change on future health and mortality in two steps. First, we investigate the extent to which the VRA has indeed induced a discernable change in university attendance beyond what would have been observed in its absence. *Prima facie* evidence in Figure 1 suggests it did. We next model (monthly) life duration taking into account the potential endogeneity of university attendance.

The estimation strategy is cast within an instrumental variable triple difference estimator. Thus, in the first step we compare university attendance of men born prior to 1921 to that of men born after 1921. The former are assumed "too old" to engage in university training after War II. This second difference relates to the province of residence. Finally, the third difference focuses on men born after 1927 as they were too young for service. To fix ideas, we specify the following linear probability model:

$$UNIV_i = \alpha ONT_i + \beta A1835_i + \gamma A1835ONT_i + \delta A1824_i + \eta A1824ONT_i + \epsilon_i, \tag{5}$$

where $UNIV_i$ equals 1 if individual i ever attended university for at least one year (E is the dynamic model of section 2), ONT_i equals 1 if born in Ontario, and the remaining terms are age dummy variables as of June 1945.[6] The treatment effect is captured by the estimate of the parameter η, *i.e.* men between 18 and 24 of age born in Ontario, and thus eligible for the VRA.

The second step involves estimating the following Cox proportional hazard model:

$$\begin{aligned}\lambda(t|X_i) =& \lambda_0(t) exp(\rho\hat{\nu}_i + \alpha UNI_i + \beta ON_i + \gamma A1835_i + \delta A1835ONT_i + \\ & \eta A1824_i + \upsilon_1 BY + \upsilon_2 BY^2 + \upsilon_3 BY^3 + \epsilon_i),\end{aligned} \tag{6}$$

where t is age (in months), $\hat{\nu}_i$ is a residual computed from the first stage regression, BY, BY^2, BY^3 is a third-order polynomial in birth year, and the remaining variables are similarly defined as

reflected some pent up demand resulting from veterans having had to interrupt their education during the war. Indeed, although financial assistance for individuals wishing to pursue university education was available for any eligible returning veteran, it was naturally particularly attractive for younger men relative to slightly older men who may have had a job prior to their military service and who would presumably have preferred to re-enter the labor force rather than attending college, perhaps benefiting from the vocational training component of the VRA.

[6] Note that only English-speaking men from Ontario and French-speaking men from Quebec are included in the regressions.

above. The treatment effect is captured by the estimate of the parameter α and potential endogeneity is assessed through a simple statistical test on the estimate of the parameter ρ.

4.3 First-Stage Estimation Results

The estimation of the model relies on the recently released Canadian Census Health and Environment Cohort (CanCHEC). CanCHEC was derived on a linkage-based dataset compiling socio-economic, demographic, cancer, mortality, and place of residence data for 2.5 million Canadians over a 20-year follow-up period. It thus provides detailed information of the individual characteristics on 15% of the entire population aged 20 and over in 1991 as well as their date of death between 1991 and 2011.

Figure 2 below uses CanCHEC data to plot the five-year death rates for males aged between 55 and 80 by province in 1991. The top lines relate to high-school graduates and the bottom lines to university graduates. Notice first that the profiles are relatively similar in both provinces. Furthermore, university graduates are 3-4 percentages points less likely to die over the next five years at every age. This translates into an life expectancy differentials of 3.09 and 3.46 years in Quebec and Ontario, respectively.

Figure 2: Five-Year Death Rates by Age

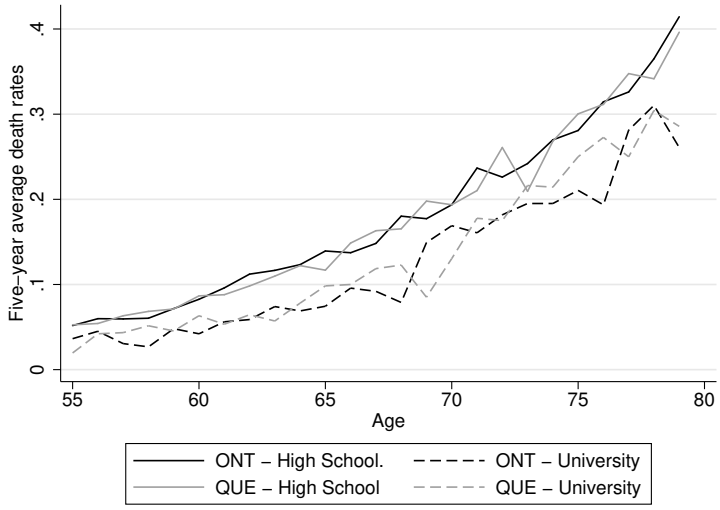

Table 6 reports the parameter estimates of the probability of attending university for at least one year among men born between 1910 and 1940. Consistent with Figure 1, men born in English-speaking Ontario are found to be 3 to 4 percentage points more likely to have attended university than those born in French-speaking Quebec. The most noteworthy feature of the table

concerns the treatment effect associated with the variable $1824ONT$. According to its parameter estimates, men aged between 18 and 24 and born in Ontario were roughly 3-4 percentage points more likely to attend university presumably due to the VRA than otherwise. All parameter estimates are fairly robust to the inclusion of various age polynomials. The parameter estimates of the regressions are next used to compute individual residuals which are then incorporated in the second-stage Cox model to control for the endogeneity of university participation.

Table 6: Probability of Attending University at Least One Year

Variables	(1)	(2)	(3)
ONT	0.037^{***}	0.037^{***}	0.037^{***}
	0.002	0.002	0.002
1835	0.024^{***}	0.025^{***}	0.024^{***}
	0.008	0.008	0.006
1824	-0.011^{**}	-0.012^{*}	-0.007^{*}
	0.005	0.006	0.004
1835ONT	-0.025^{***}	-0.025^{***}	-0.025^{***}
	0.004	0.004	0.004
1824ONT	0.038^{***}	0.038^{***}	0.038^{***}
	0.005	0.004	0.005
Third-order polynomial in birth year	✓		
Fourth-order polynomial in birth year		✓	
Linear spline in birth year			✓
N	165K	165K	165K
F-test Instrument	62.51	76.18	67.65

Note: Bootstrap standard error between parentheses. Sample: men born between 1910 and 1940 in the provinces of Quebec and Ontario

4.4 Second-Stage Estimation Results

As stressed earlier, the impact of education on life expectancy is investigated through a Cox proportional hazard model in which university attendance is instrumented by the VRA.[7] The estimation results are reported in Table 7. Each specification includes the same age polynomials as in their respective first-stage estimation. According to the parameter estimates, men who were induced to attend university as a result of the VRA have a much longer life expectancy. The hazard ratios are all statistically significant and are roughly equal to 0.2. The table also reports the confidence intervals for each specification. Note how the latter overlap with the one derived from COMPAS (Table 2). Indeed, the confidence intervals from the Cox model range between 0.076 and 0.606, whereas the one computed from COMPAS ranges between 0.487 and 0.710. Both approaches thus conclude that education has a sizeable impact on life expectancy, though COMPAS yields a slightly more conservative estimate. The first-stage residuals are

[7]Time is measured monthly. Yearly data yield essentially the same results.

incorporated into the model to account for the potential endogeneity of university education. The associated parameter estimates are all statistically significant, thus underlining its likely endogeneity.

Table 7: Cox Proportional Hazard Model

	(1)	(2)	(3)
University attendance[1]	0.216***	0.216***	0.216***
Confidence Interval	[0.076, 0.606]	[0.077, 0.610]	[0.076, 0.609]
First-stage Residual	3.216**	3.208**	3.205**
Standard error	(1.697)	(1.692)	(1.693)
Third degree polynomial birth year trend	✓		
Fourth degree polynomial birth year trend		✓	
Linear spline birth year trend			✓
First-stage specification (see Table 6)	(1)	(2)	(3)
N	165K	165K	165K

[1]Instrument - 18-24 years old in June 1945 and born in Ontario. Exponentiated coefficients.
Note: Bootstrap standard error between parentheses. Sample: men born between 1910 and 1940 in the provinces of Quebec and Ontario

4.5 Education and Health Outcomes

Finally, we use the same estimation strategy as above to investigate the potential causal links between education and various health outcomes. Table 8 reports the marginal effects of a university degree induced by the VRA. Not surprisingly given the sample size, the parameter estimates are quite stable across specifications and most are highly statistically significant. According to the table, earning a university degree leads to a significant reduction in the probability of experimenting heart diseases, non-respiratory cancers, and other causes of death. This is particularly the case for the latter two. The parameter estimates are consistent with the differential incidence of similar diseases generated by COMPAS as reported in Table 4.

Interestingly, while we find statistically and economically significant effects of university education on many diseases, the estimates in the first row show no effects on contracting respiratory diseases, including lung cancer. This may seem counterintuitive given that one may expect that getting an arguably exogenous dose of university education would have particularly large effect on smoking behaviour (see *e.g.* Grimard and Parent, 2007). Recall, however, that the VRA intervention affected cohorts of men who lived in an era where smoking was quite common.[8] In addition, those men served in the armed forces where smoking was quite prevalent and cigarettes were part of their daily rations. So it is not completely surprising that of all health behaviours

[8]The "official" warning from the U.S Surgeon General stating that smoking causes cancer came much later in 1964.

that may have been changed due to attending university, smoking might have been the least affected.[9]

Table 8: Multinomial logistic model - Marginal Effects of University Education on Health Outcomes

Marginal Effect	(1)	(2)	(3)	(4)
Respiratory diseases(including lung cancer)	−0.020	−0.021	−0.021	−0.019
	(0.089)	(0.087)	(0.087)	(0.090)
Heart diseases	−0.059**	−0.059**	−0.060***	−0.058***
	(0.028)	(0.028)	(0.029)	(0.028)
Cancer (excluding lung)	−0.128***	−0.128***	−0.128**	−0.128**
	(0.041)	(0.041)	(0.041)	(0.041)
Other causes of death	−0.171***	−0.172***	−0.172*	−0.171*
	(0.053)	(0.053)	(0.053)	(0.054)
Third degree polynomial in birth year	✓			✓
Fourth degree polynomial in birth year		✓		
Linear spline in birth year			✓	
Age fixed effects	✓	✓	✓	
Age spline				✓
First stage speification (see Table 6)	(1)	(2)	(3)	(1)
N	2.5M	2.5M	2.5M	2.5M

[1]Instrument - 18-24 years old in June 1945 and born in Ontario.
Note: Bootstrap standard error between parentheses. Sample: men born between 1910 and 1940 in the provinces of Quebec and Ontario

5 Conclusion

Determining whether education has a causal impact on health is important from a policy perspective. In this paper we investigate the link between the two using both a dynamic microsimulation model and a major policy intervention that was implemented in Canada to foster college education among returning WW-II veterans. COMPAS, the microsimulation model, shares many features with a similar model developed for the U.S. but focuses on the health trajectories of a representative cohort of Canadian men aged 30 and 31 in 2016.

Interestingly, both approaches find that the better educated benefit from greater health and longer life expectancy. According to COMPAS, the latter will require fewer hospital stays and fewer visits to specialists. Their additional years of life are expected to be free of limitations in activities of daily living. This is consistent with the quasi-experimental approach according to which the college educated are less likely to suffer from heart diseases, cancer and other major causes of death.

[9]See Bedard and Deschenes (2006) for evidence on the effect of the U.S. armed forces policy of distributing cigarettes during W.W.II and the Korean War.

The microsimulation and the quasi-experimental approaches focus on two different populations. Yet, they both conclude similarly. If anything, COMPAS perhaps slightly underestimates the impact of a college education on health relative to the quasi-experimental estimate. Interestingly, the simulations reveal that college education has most of its protective effect in terms of delaying the incidence of chronic diseases. This is important as it highlights one of the mechanisms by which the health-SES gradient expands over the working years. Furthermore, it highlights that the more education have a survival advantage, conditional on the diseases they have, which in the context of Canada cannot be explained by differential access to health care.

It is possible to monetize these health returns to attending college. First, there could be cost savings. If costs are paid by the individual, lower costs would be part of the private return to attending college. However, in a public health insurance system, the effect of education on health care costs merely raises transfers across different groups in the population unless subsidized care provides a disincentive to invest in education. We are not aware of studies which have shown such a negative effect of public health insurance on college attendance. Second, both in term of private and social returns, longer and better lives is certainly part of the welfare calculus. Our estimates from the microsimulation approach, which probably underestimate the true effect, suggest a causal effect of at least 5 years on healthy life expectancy. This translates, using 200 thousand dollars as the value of a statistical life year, into a welfare benefit of nearly one million dollars Viscusi and Aldy (2003). On the labor market side, Lemieux and Card (2001) estimate private returns on earnings of the order of 15% per year. Frenette (2014) estimates that there is roughly a 500 thousand dollar difference at the median in cumulative earnings (over 20 years) of college graduates compared to high school graduates (Table 8). Hence, the health benefits could be twice as large as the labor market returns (in terms of earnings).

References

Albouy, V. and Lequien, L. (2009). Does compulsory education lower mortality? *Journal of Health Economics*, 28(1):155–168.

Bedard, K. and Deschenes, O. (2006). The long-term impact of military service on health: Evidence from world war ii and korean war veterans. *American Economic Review*, 96(1):176–194.

Buckles, K., Hagemann, A., Malamud, O., Morrill, M., and Wozniak, A. (2016). The effect of college education on mortality. *Journal of Health Economics*, 50:99 – 114.

Clark, D. and Royer, H. (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review*, 103(6):2087–2120.

Conti, G., Heckman, J. J., and Pinto, R. (2016). The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour. *The Economic Journal*, 126(596):F28–F65.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, 29:1–28.

Frenette, M. (2014). An investment of a lifetime? the long-term labour market premiums associated with a postsecondary education. Technical Report Cat. 11F0019M, no. 359, Statistics Canada.

Fuchs, V. (1980). Time preference and health: An exploratory study. NBER Working Papers 0539, National Bureau of Economic Research, Inc.

Galama, T. J., Lleras-Muney, A., and van Kippersluis, H. (2017). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. SSRN Scholarly Paper ID 3081283, Social Science Research Network, Rochester, NY.

Goldman, D., Shekelle, P., and Bhattacharya, J. (2004). Health status and medical treatment of the future elderly. Technical Report TR-169-CMS, Rand Corporation.

Goldman, D. and Smith, J. P. (2011). The increasing value of education to health. *Social Science & Medicine (1982)*, 72(10):1728–1737.

Grimard, F. and Parent, D. (2007). Education and smoking: Were vietnam war draft avoiders also more likely to avoid smoking? *Journal of Health Economics*, 26:896–926.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2):pp. 223–255.

Imbens, G. W. and Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475.

Kunst, A. E. and Mackenbach, J. P. (1994). The size of mortality differences associated with educational level in nine industrialized countries. *American journal of public health*, 84(6):932–937.

Laliberté-Auger, F., Marchand, S., and Michaud, P.-C. (2017). COMPAS: GitHub Archive, https://github.com/CEDIA-models/compascvd2017.

Lemieux, T. and Card, D. (2001). Education, earnings, and the 'Canadian G.I. Bill'. *Canadian Journal of Economics/Revue canadienne d'économique*, 34(2):313–344.

Lleras-Muney, A. (2005). The Relationship between Education and Adult Mortality in the United States. *The Review of Economic Studies*, 72(1):189–221.

Meara, E. R., Richards, S., and Cutler, D. M. (2008). The gap gets bigger: Changes in mortality and life expectancy, by education, 19812000. *Health Affairs*, 27(2):350–360. PMID: 18332489.

Meghir, C., Palme, M., and Simeonova, E. (2018). Education and Mortality: Evidence from a Social Experiment. *American Economic Journal: Applied Economics*, 10(2):234–256.

Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter. *American Economic Review*, 96(1):152–175.

Smith, J. P. (1999). Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status. *The journal of economic perspectives : a journal of the American Economic Association*, 13(2):144–166.

Tjepkema, M., Wilkins, R., and Long, A. (2012). Cause-specific mortality by education in Canada: a 16-year follow-up study. *Health Reports*, 23(3):23–31.

van Kippersluis, H., O'Donnell, O., and van Doorslaer, E. (2009). Long Run Returns to Education: Does Schooling Lead to an Extended Old Age? *The Journal of human resources*, 4:1–33.

Viscusi, W. K. and Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1):5–76.