IMPACTS OF PUBLIC POLICY ON CHILD DEVELOPMENT AND WELLBEING

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Lauren Eden Jones

May 2014

IMPACTS OF PUBLIC POLICY ON CHILD DEVELOPMENT AND WELLBEING

Lauren Eden Jones, Ph. D.

Cornell University 2014

This dissertation examines how public policy initiatives in the areas of education, health and consumer safety regulation can impact the development and wellbeing of children. Methodologically, I employ data-intensive approaches that correct for selection problems by exploiting variation in policy exposure, imaginative sample partitions, and demanding robustness checks. Chapter one examines an educational policy in Canada that provided missionary boarding school to indigenous children in Canada throughout the 20^{th} century. Using the gradual phase-out of the policy after 1950, I examine how exposure to the schools affected adulthood health and social outcomes. I find that the schools led to increased risky health behavior, decreased social cohesion and potentially negative mental health effects. Chapter two investigates how the expansion of prescription drug coverage in the province of Quebec affected rates of stimulant use for treatment of ADHD, and whether use of stimulants improved educational or behavioral outcomes in the medium and long terms. The results suggest that those children who began taking the medication as a result of the policy experienced increased depression and anxiety, and fared worse on educational outcomes. Chapter three examines whether child-safety seats – the use of which are mandated in many jurisdictions – are effective in preventing death in auto accidents. After replication of a previous study (Levitt, 2003), the chapter also describes how driver characteristics, the increased prevalence of SUVs, and incorrect use of safety seats impact the results. Combined, the chapters of this dissertation

reveal unintended policy effects across a broad range of areas affecting children.

BIOGRAPHICAL SKETCH

Lauren Eden Jones was born on November 10, 1982 in Mississauga, Ontario, Canada. In 2001, she enrolled at the University of Toronto as an undergraduate. Over the next four years, she completed an Honors Bachelor of Arts degree with a major in Philosophy and minors in French and Art History. She graduated in 2005 with distinction. After spending time living and working in France, she returned to academia to pursue her Masters of Science degree in Agricultural, Environmental and Development Economics at The Ohio State University. Working with Cäzilia Loibl and David Kraybill, she completed her thesis on savings behavior in low-income families and graduated in 2007. She arrived at Cornell University directly, and commenced work on her Doctoral degree in the department of Policy Analysis and Management. While at Cornell, Lauren worked closely with Sharon Tennyson and Nicolas Ziebarth; she also spent time as a visiting student at the School of Public Policy and Governance at the University of Toronto, where she worked primarily with Mark Stabile. This dissertation marks the culmination of her studies.

ACKNOWLEDGMENTS

I feel very lucky to have a long list of people to thank. Firstly, I wish to thank my advisor, Sharon Tennyson, without whose committed support and encouragement I would not have completed this work. As a teacher, mentor, and co-author, Sharon inspired me throughout my doctoral studies, and kept my head above water when it was most difficult to swim. She helped me navigate the politics of the academy (they don't teach you that in class!), and did so with grace and poise. I feel very fortunate to have her as a model as I embark on my journey as a professor.

I also thank Mark Stabile for facilitating my work at the University of Toronto. Even though I walked into Mark's office with nothing more than a letter to recommend me, he took a chance and hired me as a research assistant. Since, Mark has included me on many projects, all of which have been important learning experiences. Throughout, I have been encouraged by Mark's faith in my ability, and inspired by his passionate commitment to research. I look forward to continuing our partnership next year.

My other committee members Nicolas Ziebarth and Ronald Ehrenberg are also due great thanks. Ronald was very encouraging during the developmental period of my dissertation, and although my research moved into a different area, his expertise and encouragement are still greatly valued. Although Nicolas arrived at Cornell near the end of my doctoral studies, he has been an extremely encouraging and instructive mentor, especially in advising me throughout my job search. He read drafts of my paper carefully and commented thoughtfully an invaluable gift. I also thank him for his flexibility during the job-market period.

In a professional capacity, I would also like to thank Maria Fitzpatrick, Janet Currie, Douglas Almond, Claudio Lucarelli, Sean Nicholson, Cäzilia Loibl, the other faculty in PAM, and the staff of the University of Toronto Statistics Canada Research

Data Centre for support and guidance of various forms.

I cannot imagine having spent my time at Cornell with a more wonderful group of classmates. While I met many smart, generous and interesting people during my doctoral studies, I especially thank my class of “Pammies”: Jose Balmori de la Miyar, Anna Hill, and Kathy Michelmore (and our honorary member, Kevin Sutherland), as well as Youcef Msaid, Philip Armor, and Ross Milton. You are all brilliant and I wish you the best!

Last, but certainly not least, I would like to thank my wonderful family and friends. Nance, Den, and Jamieson, I am lucky to have you. Ash Mogg, without you, life would be so grey.

TABLE OF CONTENTS

CHAPTER 1

THE EFFECTS OF CULTURALLY INTRUSIVE EDUCATION: DID CANADA'S RESIDENTIAL SCHOOLS AFFECT HEALTH AND SOCIAL OUTCOMES?

A great general has said that the only good Indian is a dead one, and that high sanction of his destruction has been an enormous factor in promoting Indian massacres. In a sense, I agree with the sentiment, but only in this: that all the Indian there is in the race should be dead. Kill the Indian in him, and save the man.

- Captain Richard H. Pratt, founder Carlisle Indian Industrial School (as quoted in Nock, 1988, p.4).

Introduction

Formal education administered in schools is a fundamental tenet of the Western tradition. Schools provide all children with safe access to education, regardless of family or community situation, and help to level the playing field among children from varied backgrounds. However, for children whose cultural backgrounds differ from that of school administrators, classroom education may provide more than academic knowledge; it also acts as an important means of imparting cultural norms and values. Childhood immigrants, charter school students or students at religious, military or disciplinary boarding schools may be exposed to cultural norms that differ drastically from those in their home country, community, or household. For many of these children, the cultural shift associated with schooling may be large.

Since Becker (1960; 1962), economists have understood that education plays a role in determining nonmarket outcomes like health (Grossman, 1972) and family composition (Becker & Lewis, 1973; Schultz 1974). Health and social outcomes may be especially affected for children whose experience of formal education includes a component of cultural change. Indeed, recent studies have found that the process of cultural assimilation, or acculturation, affects health behavior and family composition (Hunt, Schneider & Comer, 2004; Fernández and Fogli, 2006; Christopoulou and Lillard, 2013; Bleakley & Chin, 2010; Singh & Siahpush, 2002; Abraído-Lanza et al., 2006; Kimbro, 2009; Li & Wen, 2013; Adsera and Ferrer, 2011, to name a few). By investigating foreign-born children, researchers have uncovered evidence to suggest that additional exposure to host country culture is associated with increased risky health behaviors, lower marriage and fertility rates, and mental health. If certain types of schooling impart acculturative influence, those who are exposed to the schools may be especially susceptible to the health or social effects of schooling. This is an important point since children for whom the school's culture is more remote, like immigrants or those from marginal ethnic groups, may be more vulnerable to negative outcomes.

Despite the important link between culture and health and social outcomes, formal analyses of the impact of the cultural components of education are rare. The concept of culture can be vague, intractable and difficult to measure, especially in quantitative research. Further, it is difficult to disentangle the effects of culturally intrusive educational policies from those of other cultural influences, like community. Meanwhile, instances of culturally intrusive educational experiences are increasing. The urban charter school model, for example, where students are removed from their communities to attend preparatory boarding schools with rigorous educational, behavioral and moral standards is one example (Curto and Fryer, 2011). As such allencompassing educational models gain traction in the world of education policy, analysis of their effects on nonacademic outcomes becomes increasingly important.

To provide evidence on how such educational experiences affect outcomes, this paper analyzes the community-wide, long-term health and social effects of an educational policy with an explicitly acculturative mission. I explore the case of missionary boarding schools for Aboriginal[1] children in Canada – called residential schools hereafter. These schools were an important branch of the aggressive civilization policy employed by North American governments to eradicate indigenous culture. Residential schools were operated by Christian missionaries and funded by the government, and had the explicit mandate of assimilating Aboriginal children into European culture. School administrators sought to achieve assimilation on three dimensions: religious, linguistic and vocational. The curriculum was heavily weighted toward Christian religious education, and included a component of "industrial" education, wherein students were taught tradesmanship, farming or homemaking. The schools additionally forbade the use of Native languages, requiring uptake of English or French (Miller, 1996; Milloy, 1999). Students were required to leave their families to live in residence at the segregated schools; they often attended residential schools under duress, and were frequently denied permission to visit with family (Milloy, 1999).

Such boarding schools have been employed in numerous countries, in all continents, by nearly every government with assimilative policies – from North and

[1] I use the terms *Aboriginal* and *Indigenous* interchangeably to describe people who descend from the First Peoples of Canada. I additionally use the terms *Native*, *First Nations*, and *North American Indians* to refer to the subset of indigenous Canadians who are neither Inuit nor Métis, and who are the focus of this study.

Latin America, to New Zealand and Australia, to the former USSR, China, and Colonial Africa. Estimates suggest in North America and Australia, 300,000 people attended such schools (Smith, 2010). Beyond the obvious historical importance of such schools, their legacies continue to be of importance today; both for indigenous communities, and for other culturally diverse student groups.

Obtaining accurate empirical evidence on the effects of such schools, however, is challenging for several reasons. First, the overtly assimilative objective of colonial boarding school systems suggests that it is important to measure their nonmarket effects, an endeavor that requires rich data. Second, the long history of these systems, which often date from the era of first contact between indigenous peoples and European settlers, combined with non-random school attendance patterns render plausible counterfactuals rare. Third, the geographically and societally remote nature of many Aboriginal communities makes it difficult to obtain comprehensive data on the affected population. Fourth, truancy and selected student populations produced non-random attendance patterns that will bias simple estimates of the effects of schools attendance. Finally, because the schools sought to achieve such broad reaching effects on entire communities, estimating their impact requires consideration of community-wide forces.

I am able to overcome the obstacles to measuring the causal effects of missionary boarding schools by exploiting a natural experiment facilitated by their gradual closure. Between 1880 and 1986, over 150,000 Aboriginal students attended Canada's missionary boarding schools (Milloy, 1999). As schools began to close after a shift in policy in the early 1950s, Aboriginal education became largely desegregated and secular, with indigenous students shifting into the public school system (Milloy, 1999). The gradual dismantling of the residential school system generated variation in exposure to missionary education by both geography and birth cohort, variation that I use to estimate the casual effect of exposure to the schools.

My data – which derive from the Statistics Canada 1991 Aboriginal Peoples Survey (APS) – are rich and are collected from a large subset of Aboriginal communities across the country. I use adulthood survey data collected in 1991 on respondents born between 1942 and 1972 who would have been school-aged during the phase out period of the policy between 1951 and 1986. Linking respondents to local residential schools by geography, I construct a measure of policy exposure intensity that varies across community and cohort. I use this measure to estimate the effects of community-wide exposure to the policy – in comparison to exposure to a non-sectarian public school regime – on cultural, health and social outcomes. My approach allows me to comment, first, on whether the residential school policy achieved acculturation and second, whether acculturation was accompanied by social and health effects. I ask, when compared to public school, did exposure to residential schools in childhood "westernize" indigenous people on the intended dimensions? And, were there other unintended consequences of the acculturative residential school policy on adulthood health and social outcomes?

I use several additional methods to support the claim that my results reveal causal policy effects. I conduct an event study analysis to show that the residential school attendance rate was significantly affected by school closures, and was not simply decreasing due to trends. Using 1941 and 1951 census data, I additionally show that the school closures are not related to pre-existing, community-level characteristics in a way that could have affected outcomes through some other channel. Finally, I report results from a series of placebo experiments that show that my results are very likely due to the school closure.

My findings indicate that the schools had their intended acculturative effect: those with more exposure to the schools appear more likely to identify as Catholic and less likely to be subsistence hunters. I additionally find evidence that adult health behaviors were negatively affected by exposure to a residential school during youth: exposure led to increased smoking and drinking. Societal characteristics of those who were exposed to residential schools also appear to have been affected. Exposure to an open school results in decreased likelihood of marriage and decreased fertility among women. Finally, I uncover some evidence to suggest that mental health outcomes might have been negatively impacted by exposure. The overall pattern of health and social effects that accompany the acculturation experience of residential schooling mirrors the pattern in the immigration literature remarkably well (Bleakley & Chin, 2004; 2010; Singh & Siahpush, 2002; Kimbro, 2009; Li & Wen, 2013; Patterson, Kyu & Georgiades, 2012; Breslau et. al, 2009). In essence, I argue that the schools did exactly what they intended to do: they assimilated attendees into Euro-Canadian culture, and the process of assimilation has long-lasting effects for those exposed.

The findings of this study extend the recent economic literature on culture, its determinants and the role governments and institutions can play in its formation. It is also a contribution to the acculturation literature. In a context outside immigration, it adds to the quantitative, empirical research on the health and social effects associated with acculturation. It also emphasizes the role that education may play in acculturation, and the role of culture in schools. Among research on indigenous communities, the current findings suggest that colonial policies have legacies that continue to affect Native peoples.

Previous Literature

The link between culture and health and social behaviors has been long studied in many disciplines. Previous studies in the immigration literature have found links between measures of acculturation and health and social outcomes (see Hunt, Schneider & Comer, 2004 for a brief introduction to the concept of acculturation in health research). While some studies link health and social behaviors among immigrants to behaviors in one's country of ancestry (Fernández and Fogli, 2006; Christopoulou and Lillard, 2013), many look at childhood immigrants and estimate how the degree of exposure to host country culture is related to adulthood outcomes. Often, age at arrival in the host country is used as a proxy for exposure to the new culture, assuming that the younger one arrives, the longer the duration of exposure, the easier assimilation can occur. Health and social outcomes are then modeled as a function of the age of arrival variable, and the estimated coefficients are interpreted as the effect of acculturation. This identification strategy has the benefit that, for children, the age at migration is to some degree exogenously determined by parents' behavior (Bleakley & Chin, 2004).

Among studies that investigate health outcomes, researchers have found a link between younger age at arrival and increased risky health behaviors, like smoking, drinking alcohol, and drug use (for example, Singh & Siahpush, 2002; Abraído-Lanza et al., 2006; Kimbro, 2009; Li & Wen, 2013). Studies have also identified a link between mental health status and acculturation, although the direction of the relationship varies between studies (see Koneru et al., 2007 for a review). Among social outcomes, studies have linked acculturation to marriage and fertility rates. Bleakly and Chin (2010) show that those immigrants who arrived in the US during the critical period of language acquisition (before age 9) are more likely to speak English, and are less likely to be married, less fertile and less likely to live in an ethnic enclave. For immigrants to Canada, Adsera and Ferrer (2013) show that fertility increases with age at immigration until late adolescence despite home country language, suggesting that the fertility-acculturation relationship may not operate solely through language acquisition.

The question remains of how the process of acculturation determines health and social outcomes, and whether the nature of educational experiences can temper the relationship. Acculturation may simply involve learning and adopting new social norms that dictate behavior, a process that schooling can facilitate. However, intrusive cultural experiences in childhood may produce unique negative effects. Evidence from the Moving to Opportunity program, which provided incentives for American families to move out of high-poverty neighborhoods into lower-poverty ones, shows that displacement can be traumatic. Intent-to-treat analysis of the policy reveals that boys who had the opportunity to move into lower poverty neighborhoods were at higher risk for major depressive disorder and post-traumatic stress disorder than those in control groups (Kessler et al. 2014). Similarly, some analyses of the effects of school desegregation in the United States have uncovered negative effects, suggesting that shifts in the cultural milieu can be traumatic (Rivkin and Welch, 2006).

Authors have questioned whether acculturative stress may be able to explain many of the pervasive problems in Native communities. Residential schools in particular have received attention in many qualitative studies. Berry (1999) describes how former attendees of residential school indicated that school attendance was a major event affecting self-reported cultural identity. Chandler and Lalonde (2009) suggest that radical change to one's culture, such as that produced by residential schooling, can undermine continuity and place people at risk for suicide. They show that among Native communities in British Columbia, suicide rates vary with community effort to rehabilitate a sense of cultural continuity. Correlative studies using survey data have additionally linked residential school attendance to low educational attainment among children of attendees (Bougie and Senecal, 2009). However, others have found that chronic health problems do not appear related to residential school attendance among Aboriginal people (Barton et al. 2005). Using the APS, Wilson and Rosenberg (2002) investigate whether measures of traditional lifestyle are related to self-assessed health measures and find some positive correlation between traditional lifestyle and health.

Residential schools represent a particular brand of cultural influence, one that is strongly associated with the missionaries who administered the schools. Missionary cultural influence has been studied, especially in the historical context. Nunn (2010) shows how earlier and more sustained contact with missionaries explains higher rates of Christianity among certain African ethnic groups. Nunn (forthcoming) shows how the placement of Protestant versus Catholic missions in Africa has affected modern day attitudes toward female education, with descendants of individuals more likely to have come into contact with Protestant missionaries more interested in female education. Other authors have investigated the diffusion of Protestant missionaries to quantitatively test the Protestant work ethic theory (Becker and Woessmann, 2009; Bai and Kung, 2010). Few studies, however, have sought to identify the effects of missionary policies on those who experienced contact first-hand.

In concurrent work on Canadian residential schools, Feir (2013) uses the 1991 APS to investigate how attendance at a residential school during childhood affects economic and cultural assimilation into Euro-Canadian society. Feir (2013) develops a comprehensive economic model of how the interaction between government, missionaries, and indigenous people determined residential school attendance, and the consequent effects on human and social capital. The model links school closures to the religiosity of the local non-native community; her primary empirical approach uses the Catholicism rate among the non-indigenous community and national trends in residential school enrollment as instruments for actual exposure. She limits her sample to those respondents living in the Western provinces, and focuses her efforts on identifying the average treatment effects for all children, and for attendees. Her main finding shows that the schools improved human capital outcomes for some students, with increased high school graduation rates and better work outcomes. In examining cultural outcomes, she finds that attendance at the schools in childhood resulted in increased westernization in adulthood, with decreased aboriginal language use and participation in traditional activities.

The present study compliments Feir (2013) by investigating the effects of the assimilative schooling policy on nonmarket health and social outcomes. Furthermore, the analytical approach used in the present study differs from Feir's work in at least one important way. Because I am interested in the effects of residential schools on health and social behaviors – behaviors that are largely subject to spillover effects (Becker and Murphy, 2001; Durlauf, 2004) – I use an analytical approach that estimates the effects of community-wide exposure to the policy, rather than the effects of attendance itself on treated individuals. The choice to look at community-wide policy effects is especially important given the fact that a large proportion of the sample live on-reserves, in segregated neighborhoods, with little mobility. In such contexts, spillovers may be even more likely to occur (Borjas, 1995; Maclean, Webber and Sindelar, 2013). I use the same approach to investigate the effects on cultural variables, which allows me to comment directly on the success of the residential school policy whose purpose was to westernize all indigenous people through spillovers to non-attending family and community members (Miller,1996).

The Residential School Policy

The first boarding school for Aboriginal youths in what is now Canada was established by French missionaries in 1620. By the mid-19th century, such schools had become the institutional norm for Aboriginal students (Miller, 1996). The schools, which were operated by the church and funded by the government, sought to teach students the, "modes of civilized life, of action, thought, speech and dress."[2] They

[2] E.F. Wilson, Reverend and head master of several early residential schools, 1890. (Nock, 1988, p. 4)

played a key function in the mission to “civilize” indigenous people, the idea being that children should be culturally converted in order to assimilate all Aboriginal peoples into European society. They were segregated schools; until the phase-out era of residential schooling, indigenous students were not permitted to attend public schools with other non-indigenous Canadian children, and the non-indigenous did not attend residential schools.

School curriculum was structured around the assimilative mission: indigenous children were taught a trade or industry, converted to Christianity, and discouraged from the use of Native languages. The focus on industry sprang from a belief that traditional Native cultures – wherein livelihoods are earned fishing, hunting, gathering and trapping – bred laziness, a lack of respect for routine, and rendered Aboriginal peoples unfit for participation in a western economy. According to a government official, “no system of Indian training is right that does not endeavor to develop all the abilities, remove prejudice against labor and give courage to compete with the rest of the world. The Indian problem exists owing to the fact that the Indian is untrained to take his place in the world,” (Milloy, 1999, p.34). Indeed, “taking one’s place in the world” was the gold standard for graduating pupils, and as such, much of every day was spent in learning a trade: gardening, farming, carpentry or domestic duties, for example.

Religious education occupied an additional chunk of each school day. Church members – missionaries, volunteers and clerics – were responsible for the day-to-day management and teaching at the schools. Along with bible study and prayer, the religious staff taught mandatory “ethics” courses with names like, *The practice of* *cleanliness, obedience, respect, order and neatness*, and *Patriotism, Self-maintenance, Charity and Pauperism* (Milloy, 1999). The adoption of English – or French in francophone parts of the country – was viewed as the final key to cultural conversion. Use of English or French was required, and students who continued to converse in their native language were disciplined (Truth and Reconciliation Commission of Canada, 2012).

Schools tended to be located away from Native communities, a fact that forced children to be removed from their families and live in residence in order to attend school. While parental visitation policies varied between schools, the general idea was that children should be kept apart from their families, thereby easing the process of civilization (Truth and Reconciliation Commission of Canada, 2012). Throughout the Canadian residential school era, about 150 residential schools existed in total, 80 of which were concurrently open at the height of the system in the mid-20th century. Figure 1-1 shows the location of recognized residential schools in Canada that I consider in this study.

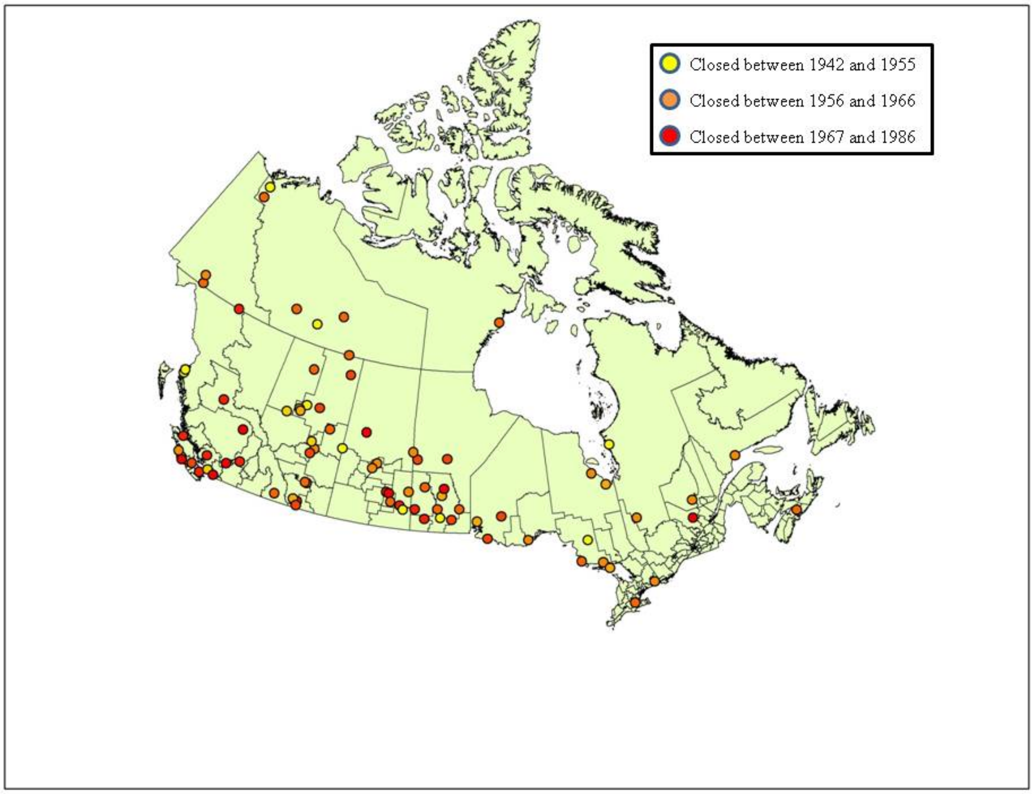

Figure 1-1. Locations of residential schools considered in the present study

Underfunding, sickness, poor quality teachers and abuse were common in schools, and anecdotal evidence abounds (Milloy, 1999; Miller, 1996; Canada, 1996; CBC, 2008). In light of these facts, it is not surprising that even after the passage of laws in 1920 that made school attendance compulsory for all indigenous children between ages 7 and 15, families were reluctant to send their children to residential schools. Despite the law, a large network of truancy officers to enforce it, and tax incentives for parents who sent children, the schools were plagued by imperfect attendance. In 1927 for example, only about one-third of Canada's 20,000 school-aged Native children were enrolled in residential school, with another 8,000 attending segregated day-schools; the remainder did not attend any formal school (Miller, 1996).

Enrollment grew to a peak by the mid-20th century, with about 10,000 indigenous students living at residential school (Truth and Reconciliation Commission of Canada, 2012). The resulting attendance pattern at the schools was certainly biased, with accounts that orphaned, sick or troubled children were recruited first in order to keep schools operational (Milloy, 1999).

School Closures

By 1948, officials had begun to understand that the boarding school system was failing. With the passage of the Indian Act in 1951, the official government policy on Native education shifted, and integration rather than segregation began to blossom as the new approach to Native education (Canada, 1996). The sectarian residential schools began to be eliminated – either transformed into government-run hostels where rural Native students could live while attending nearby public schools or shut down altogether – and students were transferred into the public school system. By 1969, about 60 percent of Native students were attending public schools (Miller, 1996). Figure 1-2 shows the number of operational residential schools by year.

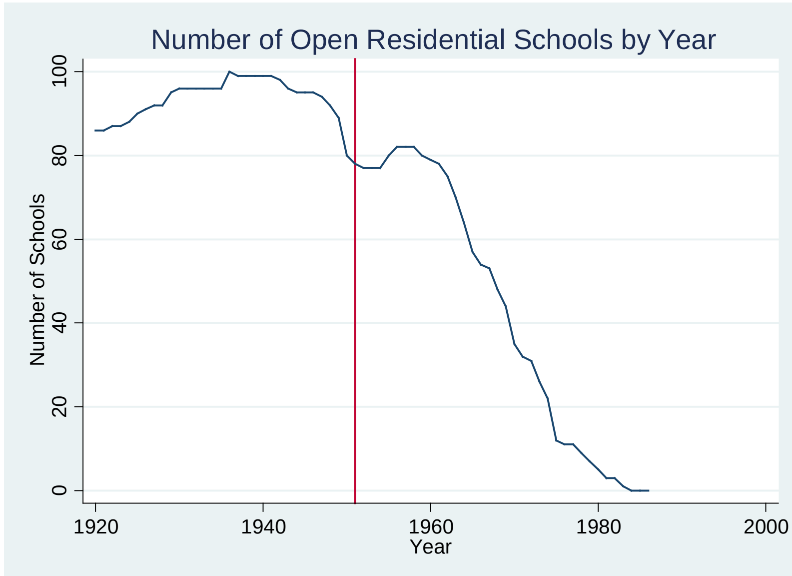

Figure 1-2. Number of open residential schools by year

The closure of the schools was characterized by an ad-hoc process: as Milloy, one of the primary researchers and authors on the 1996 Royal Commission on Aboriginal People, states, "progress in closing the schools was difficult and slow…the Department [of Indian Affairs] had not only to fashion a program that linked Aboriginal communities with local non-Aboriginal school boards and provincial ministries of education…[but also] had to contend with its old allies the churches, who continued to insist upon the importance of denominational education," (Milloy, 1999, p.190). As such, the closure process took over four decades to complete, with any particular school shutting its doors only once a willing public school board and provincial ministry of education were identified to accept the Native pupils. Resistance from the churches, especially the Catholic Church, additionally slowed the process. Finally, in some cases, Native groups themselves were reluctant to lose a local school since without the residential school system, many remote communities would have been without an easily accessible formal education option. In seven cases, Native resistance to school closure was so strong that control of the local residential school was transferred to the Band and education continued (Miller, 1996). By 1986, though, students were no longer exposed to the assimilative, sectarian residential schools; rather, they were attending schools that provided typical public school curricula.

Research Design

Data

The primary data employed in this study are drawn from the restricted access files of the 1991 Aboriginal People's Survey, a cross-sectional survey of self-identified indigenous people conducted by Statistics Canada. For the present study, I restrict the sample of respondents to those who identify as North American Indian (NAI), since the residential school experience differed between the different sub-groups of Canada's indigenous population.[3] Due to a change in the residential school question asked to those over 49 years of age versus those under 50, I further restrict the sample to those born in 1942 or later. These restrictions yield a base sample of 34,260 respondents. I additionally partition the sample to investigate the effects of residential school on those who live on-reserve in 1991 (N=23,810).[4]

---

[3] North American Indian, Inuit and Métis are the three main groups of indigenous origin in Canada. Inuit people are those who live in the northern regions of the country, primarily in the three territories. The Métis descend from North American Indian ancestors who mixed with early French colonizers.

[4] Due to Statistics Canada privacy rules, all reported sample sizes have been rounded to the nearest 10.

The APS survey data contain responses to census questions. Additionally, respondents are asked detailed questions on a range of topics, including education, employment, housing, language, mobility, health and lifestyle. While the APS does not ask directly about values or beliefs, it does provide information on a host of behaviors that are good indicators of the degree to which respondents have shifted away from traditional activities. Specifically, I construct indicator variables that capture whether the respondent speaks an Aboriginal language, whether she identifies as Catholic and whether she obtains at least half her meat for consumption from hunting. I additionally construct an acculturation measure variable which is the sum of the religion, language and vocation indicator variables, and ranges from 0 to 3 with higher scores indicating greater degrees of acculturation.

The APS also provides information on social and health behaviors. Specifically, I look at the likelihood that a respondent is married, the number of live-born children among women, the likelihood that the respondent drinks alcohol weekly or smokes cigarettes daily, body mass index (BMI) and the likelihood that the respondent has a chronic illness.[5] I also investigate the one quasi-mental health measure available in the APS: whether the respondent believes that suicide is a problem in the community. Controlling for community, it stands to reason that individuals who believe suicide is a problem in the community have had increased contact with mental health problems, either personally or among their social circle. To the extent that mental health problems are a hallmark of modern, Western society, it is informative to gauge whether residential school attendees have had increased contact

[5] Chronic illnesses include diabetes, high blood pressure, heart disease, arthritis, asthma, tuberculosis, emphysema and epilepsy.

with suicide. Finally, I use the questions that asks respondents if they lived at a residential school during elementary or high school to construct my main residential school attendance indicator, which equals 1 for respondents who ever attended a residential school. Details on variable construction and survey questions are included in Appendix 1. Summary statistics for various outcome and control variables for the entire sample, and by residential school attendance status, are presented in Table 1-1.

Table 1-1. Summary Statistics

North American Indians			
	Full sample mean/sd	Attended res scll mean/sd	Did not attend res scl mean/sd
Attended residential school	0.207	1	0
	(0.405)		
Graduated high school	0.371	0.343	0.385
	(0.483)	(0.475)	(0.487)
Working for pay in 1990/1991	0.680	0.668	0.694
	(0.466)	(0.471)	(0.461)
Receiving welfare	0.361	0.396	0.350
	(0.480)	(0.489)	(0.477)
Speaks Aboriginal language	0.561	0.783	0.497
	(0.496)	(0.412)	(0.500)
Obtains half meat from hunting	0.218	0.307	0.192
	(0.413)	(0.461)	(0.394)
Catholic	0.464	0.519	0.453
	(0.499)	(0.500)	(0.498)
Lives on reserve	0.345	0.459	0.309
	(0.475)	(0.498)	(0.462)
Acculturation scale	1.70	1.40	1.80
	(0.84)	(0.83)	(0.826)
Married	0.404	0.424	0.397
	(0.491)	(0.494)	(0.489)
Number of live-born babies	2.30	3.00	2.10
	(1.92)	(2.08)	(1.78)
Worried about suicide	0.410	0.468	0.400
	(0.492)	(0.499)	(0.490)
Drinks weekly	0.268	0.207	0.285
	(0.443)	(0.405)	(0.451)
Smokes Daily	0.467	0.470	0.470
	(0.499)	(0.499)	(0.499)
Body Mass Index	25.90	26.60	25.80
	(4.87)	(4.87)	(4.84)
Has a chronic illness	0.367	0.406	0.356
	(0.482)	(0.491)	(0.479)
N	31,630	7,590	24,040

The summary statistics suggest that those who attended residential schools are less likely to have graduated high school, and faring worse in the labor market. They appear more culturally indigenous than non-attendees, a finding that is contrary to the predicted effects of residential school attendance. They are more likely to be married and have more babies. Attendees appear less likely to drink alcohol than non-attendees; the two groups are fairly similar on other health dimensions.

The confidential files of the APS that I use in this study provide geo-codes at the enumeration area level, which I refer to as *sub-counties* throughout. The sub-county divisions I employ herein identify geographic areas of between 125 and 650 dwellings. While the sub-county divisions available in the APS are useful in matching respondents to residential schools at a very fine level, they do not delimit areas with any political significance. Furthermore, there are very few respondents living in many sub-counties. As such, I also employ Federal Electoral District partitions – the next largest unit of geography – which are composed of enumeration areas. Federal Electoral Districts (referred to as *counties* henceforth) are each represented by a member of the House of Commons and in this regard, respondents within any given county experience some degree of policy homogeneity. There are about 300 counties in Canada.

I use the 1941 and 1951 waves of data from the Canadian census to help illustrate that residential school closure dates are not determined by pre-existing county level characteristics. I calculate average characteristics at the county level for each wave of survey data, using the person weights included with each wave of census data.

Linking the APS and the School Closure Data

To identify school closure dates, I use the list of recognized residential schools compiled for the Indian Residential Schools Settlement Agreement, the class-action settlement between former attendees of residential schools and the Canadian government. Closure dates were acquired from the Aboriginal Healing Foundation (2007) for the approximately 150 recognized residential schools. However, since the present study only addresses the effects of residential schools on attendees of schools born after 1941, I eliminate schools that closed prior to 1942. I further eliminate schools for which closure dates are missing (N=10). These restrictions leave me with a list of 95 potential residential schools that I link to the APS data.

I link the APS data with the school closure data geographically by sub-county. Unfortunately, the APS data only provides information on respondents' province of birth and sub-county of residence in 1991; it does not include respondents' or sub-county of birth. Using information on mobility, band membership, and province of birth, I impute sub-county of birth for the respondents in my sample.[6] A detailed description of the imputation process is presented in Appendix 1.

Once I determine sub-county of birth for all respondents, I am able to link each to their local residential schools. For each residential school, I identify latitude and longitude coordinates using Google Maps. I use ArcMAP10 software to plot the location of each residential school on a 1991 sub-county-level base map of Canada

[6] All analyses were repeated using respondents' current county of residence as birth county. Results were not sensitive to use of the imputed birth county.

obtained from Statistics Canada. For each school, I define a “local catchment area” of radius 100 kilometers. If a sub-county is intersected by a given school's catchment area, then the school is considered local to that sub-county. I merge the local school data into the APS survey data by sub-county. For each respondent, I identify all schools that are local – in many cases, there is only one local school, but about 30 percent of the sample has multiple local schools – and I determine the date of first school closure. This first local school closure date is the relevant policy changes date for each respondent in my sample.[7]

Analytical Approach

Estimating the causal effects of residential schools using an ordinary least squares (OLS) method will not uncover the true causal effect of residential school attendance. While enrollment in residential schools became compulsory in 1895, enforcement of the law was less than perfect and that schools tended to enroll orphaned, sick or neglected children (Milloy, 1999). Indeed, the simple comparisons of mean demographics reported in Table 1-1 indicate that attendees are worse off on many dimensions than non-attendees.

To overcome bias, I use the closure of schools identified at the sub-county-level as a proxy for actual attendance at a residential school. I argue that residential school closure was a pseudo-random process and, as such, determines whether a student attended residential school, but does not affect adult outcomes directly. Further, by including county-specific and cohort controls to capture any pre-existing

[7] All analyses were repeated using the last local school closure date as the relevant date. Estimates decreased in magnitude, but remained significant and similarly signed.

regional variation and trends in outcomes that could be correlated with the local school closing date, the effect of exposure to an open, local school can be identified. Finally, by using the sub-county, cohort-specific exposure to the residential schooling policy, I am to account for the intended cultural spill-over effects of the policy on non-attendees. This approach allows me to estimate the overall, community-wide effect of exposure to the residential school policy in contrast to a traditional, public school regime.

I construct the residential school exposure variable as follows: For respondents born in year *y*, in sub-county *sk*, with first school closure year $Close_{sk}$, I define age at first local closure, or exposure years,

$$A_{ysk} = Close_{sk} - y \tag{1}$$

I set the exposure years variable, A_{ysk} to 0 for all respondents not were yet born at the time of school closure (N=5,540), and who were born in a sub-county with no local school (N=16,190). The exposure years variable A_{ysk} is a measure of the within-community, cohort-specific exposure to residential schooling.

Using School Closure as an Exogenous Source of Variation

In order for school closure to be an appropriate indentifying variable, there must be an indication that the closure process was unrelated to outcomes except insofar as it determined exposure to the schools. Figure 1-1, which suggests that school closures were randomly distributed over geography, provides the first sign that the exposure years variable may be a reliable proxy. A second concern is that the school closure process – which was determined in part by the willingness of public schools to take on displaced indigenous children after a residential school closed – could be related to variation in characteristics of the non-indigenous population. In this case, the closure process cannot be used as an exogenous determinant of residential school attendance; cultural, health, and social outcomes of the indigenous population may be related to the closure process through a non-causal pathway.

I allay this concern using the 1941 and 1951 censes of population county-level characteristics. Using a survival analysis approach, I show that the process of school closures cannot be explained by pre-existing conditions among the non-native population. Suppose that the variable t represents the number of years between the first year of my relevant data time period (1942) and the first school closure year in county k. I estimate the following model:

$$h_k(t) = \alpha + \boldsymbol{X_k} \cdot \boldsymbol{\beta} + \varepsilon_k \quad , \tag{2}$$

where $h_k(t)$ is the likelihood that the school in sub-county k closes at time t, given it has remained open until at least time t, and $\mathbf{X_k}$ a vector of county level covariates. If school closure dates are unrelated to pre-existing county characteristics, the vector of coefficients $\boldsymbol{\beta}$ should have little predictive power, and the coefficients should be jointly insignificant.

A final concern is that residential school attendance rates could have been decreasing, regardless of the school closures. In this case, the exposure years variable would appear to be related to the likelihood of residential school attendance by chance. I use an event study approach, which estimates the effect of the school closure on attendance net of any pre-existing cohort attendance tends, to show that the school closures did indeed induce a shift in attendance rates. I estimate the following model:

$$R_i = \sum_{\boldsymbol{a=0}}^{\boldsymbol{22}} \boldsymbol{D_{sk}^{a}} \cdot \boldsymbol{\Gamma} + \mathbf{X_i} \cdot \boldsymbol{\Delta} + \theta_y + \lambda_k + \varepsilon_i \tag{3}$$

Equation (3) explains an indicator of residential school attendance R_i for student *i*, born in year *y* in sub-county s*k*. The right hand side variables include individual-level covariates $\boldsymbol{X_i}$, cohort and county fixed effects, θ_y and $\lambda_{k,,}$ and a set of 23 dummy variables, each representing an event-specific cohort. Each dummy variable d_{sk}^{a} equals 1 for individuals born in sub-county s*k* who were exposed to an open local school at age *a*, and 0 otherwise; the excluded base category is exposure age over twenty-two, since most individuals are finished schooling by this age. By plotting the coefficients $\boldsymbol{\Gamma}$ by exposure years, I am able to illustrate the effect of exposure years on attendance, net of cohort or county-level trends.

Reduced-Form Models of the Effect of Policy Exposure

I use the exposure years variable A_{ysk} as a measure of community-wide residential school exposure, and I estimate a reduced form model of the form:

$$R_i = \alpha + \boldsymbol{X_i} \cdot \boldsymbol{\Delta} + \gamma \cdot A_{ysk} + \theta_y + \lambda_k + \varepsilon_i \quad (4)$$

In equation (4), R_i is an indicator that equals 1 for those who attended residential school, $\boldsymbol{X_i}$ is a vector of individual level covariates that determine the likelihood of having attended a residential school, and θ_y and λ_k are cohort and county fixed effects. The coefficient of interest is therefore γ, which estimates the change in probability of residential school attendance induced by one additional year of exposure to an open residential school, net of cohort and community level differences in residential school attendance.

Turning to the outcomes, I estimate equation (4) on the cultural, social and health variables described above. In these models, the coefficient γ estimates the effect of an additional year of community-wide residential school exposure. In all cases, I estimate linear models and cluster standard errors at the county-cohort level.

Results

School Closure: Exogeneity and Residential School Attendance

Figure 1-3 plots average residential school attendance rates by exposure years A_{ysk}, separately for males and females. An initial law enacted in 1895 compelled all indigenous children under age sixteen to attend residential school, and while the law was changed in the 1920s to require children between ages seven and fifteen to attend, there is evidence that children both younger than seven and older than fifteen attended residential schools (Miller, 1996). Figure 1-3 reveals a typical school attendance pattern, where children begin kindergarten around age five and complete high school around age eighteen. Children who were under age five at the time of the local school closure appear less likely to have attended residential school. Similarly, children over age eighteen at the time of school closure appear more likely to have attended.

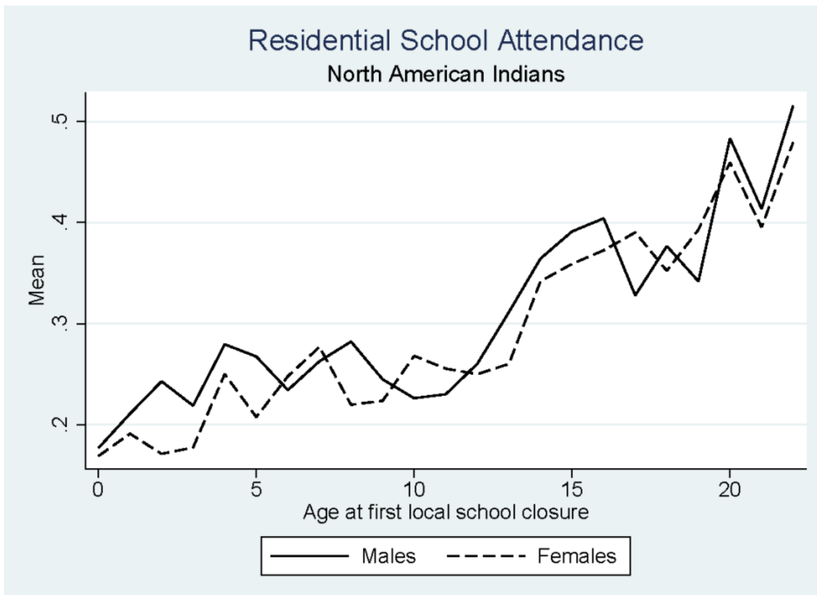

Figure 1-3. Average residential school attendance rates by exposure years

The most distinct feature of the figure, however, is the spike in residential school attendance rates that coincides with exposure years equal to twelve. If students had been attending the schools in compliance with the truancy laws, the spike in the figure should appear at exposure years equal to seven. Instead, the pattern tells us that children under age eleven at the time of first school closure were less likely to have attended. While there is little official documentation available concerning the age patterns of residential school attendees, the figure suggests that the onset of puberty changed residential school attendance habits. The relevance of age eleven is likely that it coincides with the onset of puberty, an age at which parental protective instincts tend to lessen, and when it may have become more acceptable to send a child to boarding school (Laursen, Coy, and Collins, 1998; Steinberg, 1988).

Another potential explanation for the changing attendance rates may be due to the nature of curriculum at residential schools. The vocational component of school curricula had the additional benefit of providing free labor to run the schools. Many residential school administrators employed their students as unpaid workers, exploiting them to complete school chores such that the school could be financially self-sufficient (Milloy, 1999; Canada, 1996). If older students were able to contribute more consistent, valuable labor, it stands to reason that school administrators would have had an increased incentive to attract and keep adolescent students.

Table 1-2 provides evidence that school closures are not related to pre-existing local characteristics. Column (1) of the table shows results from the estimation of the survival model using the 1941 county characteristics, and column (2) shows the results using 1951 characteristics. In general, there does not appear to be a significant relationship between county characteristics and the duration of school operation. The only significant coefficient estimate is found on the proportion of male residents variable in the 1941 regressions, a correlation that can be explained by the large volume of men who left Canada to fight in World War II. However, since it does not persist into 1951, it does not cause concern. Furthermore, Wald tests of the joint significance of the coefficients shows that, in combination, the community-level pre-characteristics do not help predict the closure date of the local school. It is reassuring that the school closure process is not correlated with characteristics of the local, non-native population.

Table 1-2. Results of survival analysis predicting school closure using 1941 and 1951 census data

Years until first school closure	1941 Census data	1951 Census data
Proportion Married	0.7032	1.6576
	(1.1692)	(1.6533)
Years of Schooling	0.0043	0.0928
	(0.0825)	(0.1237)
Proportion Catholic	-0.5866	1.0582
	(0.7947)	(0.6996)
Average age	0.0714	0.0112
	(0.0455)	(0.0391)
Proportion male	-4.6344*	0.9723
	(2.6303)	(2.0837)
Proportion Canadian citizen	2.6059	0.6063
	(1.8764)	(2.4471)
Proportion who speak French	0.5740	0.1940
	(0.7637)	(0.7019)
Proportion under 18	2.5999	0.8797
	(2.5312)	(1.7030)
Population (thousands)	-0.0132	-0.0095
	(0.0435)	(0.0285)
Number of Children per household	0.0042	N/A
	(0.1363)	
Proportion of households owned by family	0.5276	N/A
	(0.6961)	
Number of Counties	63	62
Wald test for joint significance (X^2)	10.1	5.85

County-level characteristics computed from the 1941 and 1951 20 percent sample census response data, using appropriate weights. Estimated models are proportional hazard models fit via maximum likelihood. Standard errors reported in brackets.*p <0.05; **p<0.01

Figure 1-4 provides evidence that residential school closures did affect the attendance rates beyond the overall decreasing trends in attendance. The event study graph plots the coefficients γ_a and their 95 percent confidence intervals by the exposure years variable. Several things are worth noting in this graph. First, after controlling for cohort and county, there does not appear to be any trend in the likelihood of attending a residential school by exposure years. The abrupt shift in attendance rates at exposure years twelve indicates that closures did shock attendance rates in nearby communities. It appears that students over age twelve at the time of school closure are just as likely to have attended residential school as those who were older than 22 in the time of first closure. The figure suggests that using school closure as an exogenous source of variation is valid.

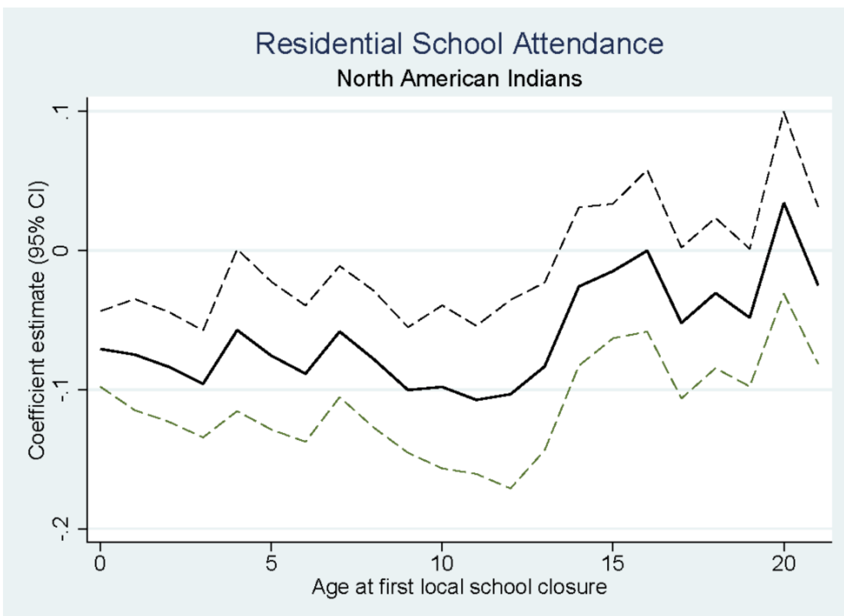

Figure 1-4. Event study of residential school attendance

Table 1-3 formally illustrates the relationship between exposure years and residential school attendance with the results of the estimation of model (4). Moving across the table from column (1) to (3), the time and geographic controls become more demanding. Column (1) results derive from a model specification that includes year of birth controls; column (2) adds county fixed effects; and column (3) includes a county-specific linear cohort trend. The results confirm that there is a positive, significant relationship between residential school attendance and the exposure years variable. While the effect estimate decreases in magnitude with the addition of county controls (columns (2) and (3)), it remains significant even with the more demanding specification. What's more, the addition of a county-specific cohort trend does not change the coefficient estimate, implying that once cohort is controlled for, the relationship between age at school closure and attendance is not due to a trend of decreasing attendance over cohorts within counties, but rather the closure itself. In sum, full 18-year exposure to an open residential school increases the likelihood of having attended residential school by between 4.1 and 8.8 percentage points, an effect size of between 16 and 35 percent.

Table 1-3. Results of reduced form models explaining residential school attendance with the exposure years variable

	North American Indians Residential school attendance		
	(1)	(2)	(3)
Exposure Years	0.0045**	0.0024**	0.0023**
	(0.0007)	(0.0007)	(0.0007)
Demographic controls	X	X	X
Year of Birth FE	X	X	
County FE		X	
County-level linear cohort trend			X
N	31,630	31,630	31,630
adj. R-sq	0.1128	0.1622	0.1618
F	46.7259	36.1622	22.42
Variable mean	0.207	0.207	0.207
Full exposure effect (coefficient x 18)	39%	21%	21%

Data are the 1991 Aboriginal Peoples Survey. Models estimated using responses from those who identify as Native, who indicate North American Indian status, and who were born between 1942 and 1971. Estimated models are linear probability. Demographic controls include gender, an indica- tor for Official Indian Status, an indicator for multiple ethnic origin, and indicator for non-Canadian birth, and 3 indicators for geographic region: North, mid-North or South. Standard errors clustered at the county-cohort level reported in brackets. *p <0.05; **p <0.01

Long-run Effects of Exposure to an Open Residential School on Acculturation

Table 1-4 shows the results of model estimations explaining the cultural variables using the exposure years variable as a measure of policy treatment. The results explain the likelihood of being Catholic, the likelihood of obtaining meat from hunting, the likelihood of speaking an aboriginal language, as well as the acculturation scale, which ranges from 0 to 3 and where higher scores indicate a greater degree of acculturation. I additionally explain the likelihood of living on-reserve. In all cases, I report results from models that include year of birth and county controls, along with demographic controls.

Table 1-4. Results of models explaining the cultural variables using the exposure years variable

	North American Indians				
	Catholic	Hunts	Speaks Ab	Culture	On-reserve
Exposure Years	0.0035**	-0.0015**	0.0001	0.0052**	-0.0004
	(0.0010)	(0.0007)	(0.0006)	(0.0013)	(0.0005)
N	33,830	31,020	33,830	31,020	33,830
adj. R-sq	0.2785	0.1556	0.3526	0.2862	0.5236
F	51.0843	9.7061	25.1016	35.4017	25.3940
Variable mean	0.464	0.218	0.561	1.70	0.345
Full exposure effect (coefficient x 18)	13.6%	-12.4%	-	5.5%	-

Data are the 1991 Aboriginal Peoples Survey. Models estimated using responses from those who identify as Native, who indicate North American Indian status, and who were born between 1942 and 1971. Dependent variables are an indicator for Catholic religion, for obtaining at least half of meat from hunting, for speaking an Aboriginal language, and for living on-reserve; Culture is a combined measure of Catholicism, Hunting and Speaking Ab that ranges from 0 to 3, where someone who is Catholic, not a hunter and does not speak and Aboriginal language receives a score of 3. Estimated models are linear probability. All models include year of birth and county fixed effects. Demographic controls include gender, an indicator for Official Indian Status, an indicator for multiple ethnic origin, and indicator for non-Canadian birth, and 3 indicators for geographic region: North, mid-North or South. Standard errors clustered at the county-cohort level reported in brackets. *p <0.05; **p <0.01

The results show that additional exposure to residential schools results in higher levels of acculturation. Specifically, those for whom a residential school was open for all 18 school-aged years are 15 percent more likely to identify as Catholic, and 11 percent less likely to hunt than those who had no exposure. Among the full sample, there does not appear to be a significant difference in the likelihood of speaking an Aboriginal language for those exposed. The results for the acculturation scale summarize the full acculturative effect of exposure to residential school: 18 years of exposure to residential school results in a 0.09 point increase in the acculturation scale on a base of 1.70, a 6 percent increase. The likelihood of living on-reserve does not appear to be related to residential school exposure.

The results reveal a connection between residential school exposure and the cultural variables that is not apparent in simple means comparisons (or in OLS regressions of the residential school attendance indicator on outcome variables, results of which are not reported here). While the summary statistics in Table 1-1 suggest that attendees are less acculturated than non-attendees, the analysis using exposure years identification reveals the opposite pattern.

Long-run Effects of Exposure to an Open Residential School on Health Behaviors

Panel A of Table 1-5 shows the results for the health behaviors and outcomes. Those exposed to an open school appear more likely to smoke and drink than less exposed people (18-year exposure results in 5 and 13 percent increases, respectively, in the probability of smoking daily and drinking weekly). However, BMI does not appear affected, nor is the measure of likelihood of having a chronic illness related to exposure.

Table 1-5. Results of models explaining the health and social variables using the exposure years variable

	Panel A: North American Indians, Health Variables			
	Drinks	Smokes	BMI	Chronic
Exposure Years	0.0011**	0.0019**	0.0001	0.0000
	(0.0005)	(0.0006)	(0.0052)	(0.0005)
N	33,480	33,480	28,270	33,480
adj. R-sq	0.0850	0.0400	0.1224	0.0860
F	30.8081	3.9191	58.6200	71.4818
Variable mean	0.268	0.467	25.90	0.367
Full exposure effect (coefficient x 18)	7.5%	7.3%	0%	0%

	Panel B: North American Indians, Social Variables		
	Married	Babies	Suicide
Exposure Years	-0.0014**	-0.0066**	0.0034**
	(0.0007)	(0.0034)	(0.0009)
N	33,830	17,340	25,100
adj. R-sq	0.1655	0.3203	0.0791
F	146.7052	103.7448	9.4417
Variable mean	0.404	2.30	0.410
Full exposure effect (coefficient x 18)	-6.2%	-5.2%	14.9%

Data are the 1991 Aboriginal Peoples Survey. Models estimated using responses from those who identify as Native, who indicate North American Indian status, and who were born between 1942 and 1971. Dependent variables in Panel A are an indicator for drinking alcohol at least weekly, smoking cigarettes daily, and having a chronic illness. The Body Mass Indicator was calculated from reported height, weight and age. Dependent variables in Panel B are an indicator for being married and an indicator for whether the respondent believes suicide is a problem in the their community; the Babies variable is the number of liveborn babies, and is explained for the sample of women. Estimated models are linear probability. All models include year of birth and county fixed effects. Demographic controls include gender, an indicator for Official Indian Status, an indicator for multiple ethnic origin, and indicator for non-Canadian birth, and 3 indicators for geographic region: North, mid-North or South. Standard errors clustered at the county-cohort level reported in brackets. *p <0.05; **p <0.01

The positive relationship between smoking and drinking, and exposure to the acculturative schooling policy confirms expectation, and aligns with results from other studies of acculturation that suggest that risky health behaviors may worsen with acculturation. That BMI and the presence of a chronic illness remain unaffected is interesting. The effects on these variables, which may take time to manifest, could present later for exposed individuals, especially since they appear to be engaging in risky health behaviors.

Long-run Effects of Exposure to an Open Residential School on Social Outcomes

Panel B of Table 1-5 shows the results for the social outcomes: marriage, fertility and concern that suicide is a problem in the community. More exposure to an open residential school is associated with a lower probability of marriage (8 percent decrease with 18 years of exposure), and women with higher exposure appear to have fewer live-born babies (full exposure results in a 5 percent decrease in fertility). Exposed individuals also appear more likely to worry about suicide in the community. Among the NAI population, those with 18 years of exposure are 13 percent more likely than those with no exposure to worry that suicide is a problem.

The results of the analyses explaining the social variables align with results from the acculturation literature. Marriage and fertility rates appear to decrease with higher rates of acculturation. Similar patterns are also found in the economics of education literature, where fertility and marriage rates appear to decrease with additional years of schooling. The suicide effect is especially interesting. After controlling for county, higher levels of concern about suicide in the community imply that exposed individuals have more personal experience with suicide or are more inclined to suicidal ideation. However, the result could also imply an increased awareness of, or concern for, community issues.

Robustness Checks

The above discussion suggests that there is a relationship between exposure to the residential school policy and the cultural, health and social outcomes. However, the analysis relies on the within-community changes across cohorts in the degree of exposure to the policy. If other unobserved factors also vary across cohorts, the above results could be spurious. In order to buttress my argument against this objection, I conduct several robustness checks. For one, I re-estimate all my models with county-specific linear cohort trends (results in the Appendix 1). If my results are due to trending changes across successive cohorts in the outcomes in question, then the inclusion of county-specific trends should eliminate or significantly change the coefficient estimates on the exposure years variable. My results, however, persist with the addition of the county-specific linear and quadratic cohort trends. In all cases, estimated coefficients remain similar in magnitude and significance level. This indicates that the effects on outcomes are not due to changing trends over cohorts, but rather to the change in community-wide residential school exposure induced by the closures.

Secondly, I run a placebo experiment with falsified school closure dates. I randomly assign exposure years values by sub-county. Using observations from respondents who had no exposure to open residential schools, I run the models on all outcomes using the placebo dates and record the coefficient estimate on the placebo exposure years variable. I repeat the entire process 1000 times for each outcome, and plot the estimated placebo coefficients. In the case that my actual estimated results reveal an effect truly attributable to the closure of residential schools rather than to chance, they should fall well outside the range of the distribution of the placebo coefficients.

Figure 1-5 shows the results of the placebo tests for residential school attendance, acculturation, smoking and drinking behavior, marriage and fertility, and worry about suicide. In all cases, the bars in the graph represent the distribution of the placebo coefficient estimates, which I have normalized to mean 0, standard deviation 1. The vertical line shows the actual coefficient estimate from the regression estimated with the true exposure years variable. The figure confirms that, by and large, the actual coefficient estimates fall within the tails of the distributions of placebo coefficient estimates. The figure suggests that it is very unlikely that I would find effects as extreme as the ones I have if the closures had been ineffective (as the placebo closures are by design).

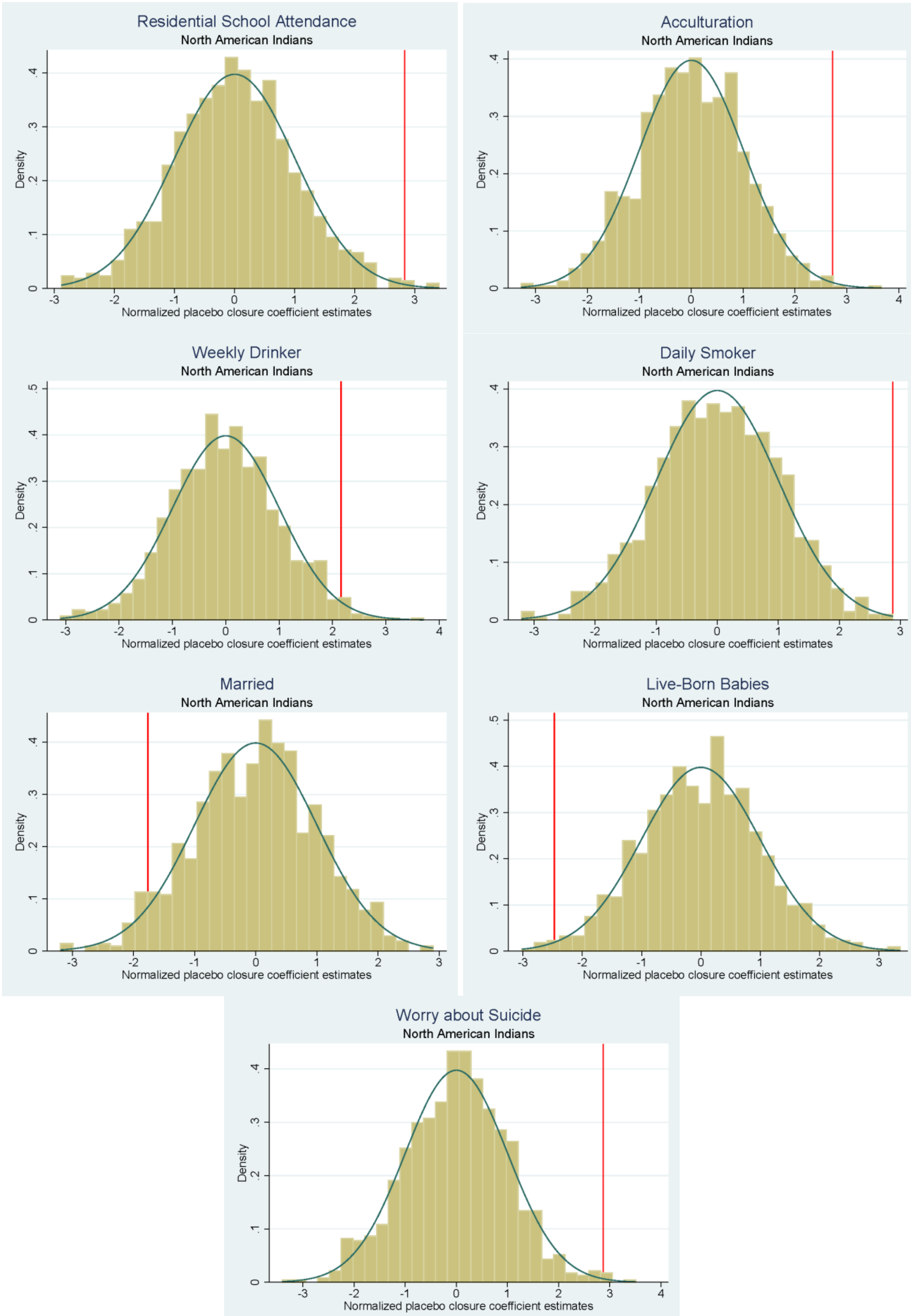

Figure 1-5. Results of placebo experiment

Do Effects Differ by Educational Attainment?

The above discussion provides evidence of how residential schooling affected outcomes, but the question of mechanisms remains. Are the effects due to acculturation, or are they attributable to some other aspect of the experience? One possible alternative is that the residential schooling policy facilitated educational access, and that the observed effects are due to additional education, rather than to the acculturative nature of the curriculum.

To address this question, I partition the sample by years of completed education and estimate models for each group. Using respondents who live on-reserve,[8] I split observations into a low education group, which includes people with less than a tenth grade education, and a high education group, including people with grade 10 or more. I then estimate models for each education subgroup. If the effects I uncover are due to additional educational attainment rather than to acculturation, I should find limited effects among the low education group.

[8] I do this because educational attainment is highly correlated with on-reserve status. When I consider the entire population, the high education group is significantly different than the low education group on many dimensions. Using the on-reserve population helps eliminate some of these confounding differences.

Table 1-6. Results by educational attainment

	On-Reserve					
	Panel A - Grade 10 or less					
	Res Scl	Culture	Drinks	Smokes	Married	Suicide
Exposure Years	0.0020*	0.0059**	0.0017**	0.0026**	-0.0023**	0.0041**
	(0.0011)	(0.0017)	(0.0008)	(0.0008)	(0.0009)	(0.0012)
N	16,620	16,620	17,630	17,630	17,800	13,740
adj. R-sq	0.1532	0.2675	0.0810	0.0504	0.1709	0.0957
F	24.3628	17.0117	16.4826	5.0715	85.7761	6.5910
	Panel B - Grade 11 or more					
	Res Scl	Culture	Drinks	Smokes	Married	Suicide
Exposure Years	0.0052**	0.0019	0.0017*	0.0029**	-0.0021**	0.0050**
	(0.0012)	(0.0026)	(0.0009)	(0.0011)	(0.0012)	(0.0016)
N	5,380	5,630	5,770	5,770	5,810	4,598
adj. R-sq	0.2015	0.2602	0.0820	0.0518	0.1838	0.1044
F	12.9278	8.6547	9.3504	2.1043	40.3739	3.6833

See Table 1-4. *p <0.05; **p <0.01

Results of these analyses are reported in Table 1-6. The coefficient from the attendance regression is significant for both subgroups, even though the likelihood of having attended residential school is smaller for the low education group. Turning to the nonmarket effects, the estimated acculturation effect appears larger for the low education group. Considering that the high education group is on average, much more acculturated than the low education group, this result makes sense. Among the low education group, exposure to residential school is related to increased smoking and drinking, decreased likelihood of marriage, and increased concern about suicide in the community, and effect sizes are similar across educational attainment. That the social and health effects associated with residential schooling are present among respondents with fewer completed years of education implies that they do not derive exclusively from the increased educational attainment the experience provided.

Conclusion

The present study seeks to estimate the long-term impacts of attendance at one of Canada's Indian residential schools. Because of the stated goal of the schools – to assimilate attendees into the dominant Euro-Canadian culture – I search for the impact of the schools on culture. I additionally investigate a broad set of outcomes, including health behaviors and outcomes, and social variables. Using the gradual closure of residential schools across Canada as a source of variation, I am able to comment on the causal effect of increased community-wide exposure to the schooling policy. This is especially important given the documented non-random nature of attendance patterns.

My results reveal an interesting pattern in adulthood outcomes for those exposed to an open residential school for longer. It is clear that the most important purpose of the schools, acculturation, was achieved. Exposed cohorts score significantly higher on the acculturation scale than the unexposed. Health behaviors like smoking and drinking are worse for the exposed, and marriage and fertility rates are lower.

Additionally, I find that exposure leads to increased concern that suicide is a problem in the community. It is impossible to say for certain whether this effect is due to increased community awareness or to increased personal experience with suicide. The suicide effect of the schools, however, coincides with the social and community distancing that acculturation implies. That the effect persists among the low-education group of respondents who may be less sophisticated additionally suggests that it may be due to increased personal experience with suicide, and potentially worse mental health. If increased negative health behaviors, mental health issues and weakened family bonds imperil community coherence, the legacy of residential schools continues to threaten indigenous culture today. Indeed, to the extent that habitual smoking and drinking, lower marriage rates and fertility, and increased exposure to suicide are hallmarks of western culture, residential school appears to have westernized attendees.

The results herein are of further interest when considered in conjunction with Feir (2013). Her work reveals similar cultural effects to those in the present study. She also discovers that for many attendees, residential schooling improved human capital outcomes like high school graduation and employment.[9] Thus, the overall pattern of policy effects includes increased economic assimilation, increased social distancing, increased risky health behaviors, and potentially increased mental health problems – it mirrors the pattern of the acculturation literature on immigration nearly perfectly.

To account for the spillover, I report results from an intent-to-treat (ITT) reduced form analysis, rather than from an instrumental variables (IV) analysis that would estimate the causal effects of the school on those who attended. Because the focus of this study is cultural, and because culture is a fluid trait with a highly social

[9] Although I do not discuss it here, when I apply my analytical approach to the same human capital outcomes of Feir (2013), I find similar results with exposure increasing high school graduation rates and improving work-related outcomes.

nature, the effects of having a local residential school open are likely to spillover to non-attendees in the community. Indeed, the reader can infer from the coefficient estimates from the residential school attendance model and the outcome models that IV estimates of the average treatment effects are large. As such, I choose to present the ITT results; to the extent that residential schools have had lasting effects at the community level, a well-established fact in the literature, this approach is valid (Smith, Varcoe, and Edwards, 2005; Canada, 1996).

Despite the large effect estimates of an IV approach, the ITT approach produces estimated effects of an appropriate magnitude. The identification strategy employed is also robust to the progressive addition of cohort and region controls: county and year of birth fixed-effects and county-specific cohort trends do not alter coefficient estimates significantly, implying that results do capture causal effects of the schools. The placebo tests provide additional support for the causal argument.

One important caveat of the work is that the results only apply to the generation of North American Indian people contemporary to the phase-out period of the residential school policy. I am unable to comment on the overall effects of residential schools on all former attendees. The fact that I am able to detect cultural effects in 1991 – by which time the majority of cultural assimilation of Native peoples into Euro-Canadian culture had already occurred – is remarkable. It suggests that my estimates are a lower bound on the assimilative power of the policy.

While the policy of missionary boarding schools may appear to be of primarily historical interest, current educational initiatives exist that mirror the residential school context in important ways. In many developing nations, indigenous education interventions often include a cultural component, with missionary boarding schools persisting in some countries. Boarding schools in the western world may also seek to transform nonacademic norms and values, and charter schools that seek to reform students from marginalized cultures may also be designed in such a way as to produce unintended nonacademic effects. In all cases, the residential school program provides important lessons. First, that broad-reaching, culturally intrusive educational initiatives may have unintended effects in health and social domains; and second, that an educational initiative cannot be judged on human capital outcomes alone.

REFERENCES

Aboriginal Healing Foundation. (2007). *Directory of residential schools in Canada*. Ottawa, ON: Author. Retrieved from http://www.ahf.ca/publications/residential-school-resources.

Abraído-Lanza, A.F., Armbrister, A.N., Flórez, K.R., & Aguirre, A.N. (2006). Toward a theory-driven model of acculturation in public health research. *American Journal of Public Health*, 96, 1342-1346.

Adsera, A., & Ferrer, A. (2014). Factors influencing the fertility choices of child immigrants in Canada. *Population Studies*, 68, 65-79.

Bai, Y., & Kung, K. (2012). Diffusing knowledge while spreading God's message: Protestantism and economic prosperity in China, 1849-1920. Unpublished.

Barton, S.S., Thommasen, H.V., Talli, B., Zhang, W., & Michalos, A.C. (2005). Health and quality of life of aboriginal residential school survivors, Bella Coola valley, 2001. *Social Indicators Research*, 73, 295–312.

Becker, G. (1960). *An Economic Analysis of Fertility*. Princeton, NJ: Princeton University Press.

Becker, G. (1962). Investment in human capital: A theoretical analysis. *The Journal of Political Economy*, 70, 9–49.

Becker, G., & Lewis, G. (1973). On the interaction between the quantity and quality of children. *The Journal of Political Economy*, 81, S279–S288.

Becker, G., & Murphy, K. (2001). *Social Economics: Market Behavior in a Social Environment.* Cambridge, MA: Harvard University Press.

Becker, S.O., & Woessmann, L. (2009). Was Weber wrong? A human capital theory of protestant economic history. *The Quarterly Journal of Economics*, 124, 531–596.

Berry, J.W. (1999). Aboriginal cultural identity. *Canadian Journal of Native Studies*, 19, 1–36.

Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *The Review of Economics and Statistics*, 86, 481–496.

Bleakley, H., & Chin, A. (2010). Age at arrival, English proficiency, and social assimilation among U.S. immigrants. *American Economic Journal: Applied Economics*, 2, 165–192.

Borjas,G. (1995). Ethnicity, neighborhoods, and human-capital externalities. *The American Economic Review*, 85, 365-390.

Bougie, E., & Senecal, S. (2009). Registered Indian children's school success and inter- generational effects of residential schooling in Canada. *International Indigenous Policy Journal*, 1. Retrieved from http://ir.lib.uwo.ca/iipj/vol1/iss1/5.

Breslau, J., Borges, G., Hagar, Y., Tancredi, D., & Gilman, S. (2009). Immigration to the USA and risk for mood and anxiety disorders: Variation by origin and age at immigration. *Psychological Medicine*, 39, 1117–1127.

Canada. (1996). *Report of the Royal Commission on Aboriginal Peoples*. Retrieved from http://www. aadnc-aandc.gc.ca/eng/1307458586498/1307458751962.

CBC News. (2008). *Residential schools: A history of residential schools in Canada*. Retrieved from http: //www.cbc.ca/news/canada/a-history-of-residential-schools-in-canada-1.702280.

Chandler, M.J., & Lalonde, C.E. (2009). Cultural continuity as a moderator of suicide risk among Canada's First Nations. In L.J. Kirmayer & G.G. Valaskakis (Eds.), *Healing traditions: The mental health of Aboriginal Peoples in Canada* (221-248). Vancouver, BC: University of British Columbia Press.

Christopoulou, R., & Lillard, D.R. (2013). Is smoking behavior culturally determined? Evidence from British immigrants. *NBER Working paper No. 19036*. Cambridge, MA: National Bureau of Economic Research.

Curto, V.E., & Fryer, R.G. Jr. (2011). Estimating the Returns to Urban Boarding Schools: Evidence from SEED. *National Bureau of Economic Research Working Paper # 16746.*

Durlauf, S.E. (2004), "Neighborhood Effects," in Handbook of Regional and Urban Economics Vol. 4, J.V. Henderson and J-F. Thisse (Eds.), Amsterdam: Elsevier Science, pp. 2173-2242.

Feir, D. (2013). The long term effects of forcible assimilation policy: The case of Indian boarding schools. Department discussion paper no. 1301. Victoria, BC: University of Victoria, Department of Economics.

Fernández, R., & Fogli, A. (2006). Fertility: The role of culture and family experience. *Journal of the European Economic Association*, 4, 552–561.

Hunt, L.M., Schneider, S., & Comer, B. (2004). Should "acculturation" be a variable in health research? A critical review of research on US Hispanics. *Social Science & Medecine*, 59, 973-986.

Kessler, R.C., Duncan, G. J., Gennetian, L.A., Katz, L.F., Kling, J. R., Sampson, N. A., Sanbonmatsu, L., Zaslavsky, A.M., & Ludwig, J. (2014). Associations of housing mobility interventions for children in high-poverty neighborhoods with subsequent mental disorders during adolescence. *Journal of the American Medical Association,* 311, 937-947.

Kimbro, R.T. (2009). Acculturation in context: Gender, age at migration, neighborhood ethnicity, and health behaviors. *Social Science Quarterly*, 90, 1145–1166.

Koneru, V.K., Weisman de Mamani, A.G., Flynn, P.M., & Betancourt, H. (2007). Acculturation and mental health: Current findings and recommendations for future research. *Applied and Preventative Psychology*, 12, 76-96.

Truth and Reconciliation Commission of Canada. (2012). *They came for the children*. Winnipeg, MB: Author.

Laursen, B., Coy, K., & Collins, W. A. (1998). Reconsidering changes in parent-child conflict across adolescence: A meta-analysis. *Child Development, 69*, 817-832.

Li, K., & Wen, M. (2013). Substance use, age at migration, and length of residence among adult immigrants in the United States. *Journal of Immigrant and Minority Health*, August.

Maclean, J.C., Webber, D., & Sindelar, J. L. (2013). The roles of assimilation and ethnic enclave residence in immigrant smoking. *NBER Working paper No. 19753*. Cambridge, MA: National Bureau of Economic Research.

Miller, J.R. (1996). *Shingwauk's Vision: A History of Native Residential Schools*. Toronto, Ontario: University of Toronto Press.

Milloy, J. (1999). *A National Crime: The Canadian Government and the Residential School System*. Winnipeg, MB: University of Manitoba Press.

Nock, D. (1988). *A Victorian Missionary and Canadian Indian Policy.* Waterloo, ON: Canadian Corporation for Studies in Religion.

Nunn, N. (2010). Religious conversion in colonial Africa. *American Economic Review: Papers and Proceedings*, 100, 147–152.

Nunn, N. (Forthcoming). Gender and missionary influence in colonial Africa. In E. Akyeampong, R. Bates, N. Nunn, & J.A. Robinson (Eds.), *Development in Historical Perspective*. Cambridge: Cambridge University Press.

Patterson, B., Kyu, H.H., & Georgiades, K. (2012). Age at immigration to Canada and the occurrence of mood, anxiety, and substance use disorders. *Canadian Journal of Psychiatry*, 58, 210–217.

Rivkin, S., & Wlech, F. (2006). Has school desegregation improved academic and economic outcomes for blacks. In E. Hanushek & F. Welch (Eds.), *Handbook of Economics of Education 2.* Amsterdam: North-Holland.

Schultz, T. W. (Ed.). (1974). *Economics of the Family: Marriage, Children and Human Capital.* NBER. Retrieved from http://www.nber.org/chapters/c2960.pdf

Singh, G. K., & Siahpush, M. (2002). Ethnic-immigrant differentials in health behaviors, morbidity, and cause-specific mortality in the United States: An analysis of two national data bases. *Human Biology*¸74, 83-109.

Smith, A. (2010). *Indigenous peoples and boarding schools: A comparative study.* Secretariat of the United Nations Permanent Forum on Indigenous Issues. Retrieved from http://www.un. org/esa/socdev/unpfii/documents/E_C_19_2009_crp1.pdf.

Smith, D., Varcoe, C., & Edwards, N. (2005). Turning around the intergenerational impact of residential schools on aboriginal people: Implications for health policy and practice. *Canadian Journal of Nursing Research,* 37, 38–60.

Steinberg, L. (1988). Reciprocal relation between parent-child distance and pubertal maturation. *Developmental Psychology*, 24, 122–128.

Wilson, K., & Rosenberg, M.W. (2002). Exploring the determinants of health for First Nations peoples in Canada: Can existing frameworks accommodate traditional activities? *Social Science and Medicine*, 55, 2017–2031.

.

CHAPTER 2

DO STIMULANT MEDICATIONS IMPROVE EDUCATIONAL AND BEHAVIORAL OUTCOMES FOR CHILDREN WITH ADHD?

WITH JANET CURRIE AND MARK STABILE

Introduction

Over the past twenty years, mental disabilities have overtaken physical disabilities as the leading cause of activity limitations in children. Today, ADHD is three times more likely than asthma to be contributing to childhood disability in the United States (Currie and Kahn, 2011). Recent research indicates that children with ADHD have lower standardized test scores than others (including their own siblings) and are more likely to be placed in special education, to repeat grades, and to be delinquent (Miech et al., 1999; Nagin and Tremblay, 1999; Currie and Stabile, 2006, 2007; Fletcher and Wolfe, 2008, 2009). Moreover, untreated children with ADHD impose significant costs on their classmates by disrupting learning and/or diverting teacher resources (Aizer, 2009).

According to the most recent data from the Centers for Disease Control and Prevention, approximately eleven percent of U.S. children aged 4 to 17 have ever been diagnosed with ADHD and more than half of them are taking stimulant medications such as Ritalin for their condition (Schwarz and Cohen, 2013; Centers for

Disease Control and Prevention, 2005). [1] Both diagnosis and treatment rates are lower outside the U.S., but have been rapidly increasing (Polanczyk et al, 2007).

Despite, or perhaps because of the millions of children taking stimulants, drug treatment for ADHD remains controversial. The National Institute of Mental Health recommends treatment with stimulants and says that they are safe if used under medical supervision (U.S. NIMH, 2012). However, concerns continue to surface about both short term side effects, and possible side effects due to long-term use. For example, the U.S. Food and Drug Administration voted in 2006 to recommend a warning label describing the cardiovascular risks of stimulant drugs for ADHD (Nissen, 2006). Other side effects can include decreased appetite, insomnia, headache, stomach ache, dizziness and mood changes including anxiety and depression (Schachter et al., 2001, NIMH, 2012). Some studies have also found growth deficits in treated children (Joshi and Adam, 2002). Aside from the possibility of physical side effects, inappropriate use of stimulant medication could also harm children by stigmatizing them or by crowding out other interventions that might be more helpful.

Lack of evidence regarding long-term benefits of stimulant medications is a key element of this controversy. Drugs are often prescribed with the goal of helping children to be successful in school. If the drugs do not actually lead to scholastic benefits in the medium and long run, then the case for subjecting children to even a small risk of side effects is weakened. The main problems involved in assessing the

[1] Schwarz and Cohen tabulate data from the 2011-2012 wave of the National Survey of Children's Health. Methylphenidate (sold under the trade names Ritalin, Biphentin, and Concerta) is the most commonly used central nervous system stimulant for ADHD. Others include: dextroamphetamine (Dexedrine); and mixed amphetamine salts (Adderall) (Therapeutics Initiative, 2010).

long-run efficacy of stimulant medication are first, that most drug trials follow children only for a short time – between one and two months after treatment (Griffin et al., 2008) – and second, that families (and children) choose whether or not to seek treatment for ADHD, and whether to take medication if it is prescribed.

Our paper assesses the medium and long run benefits of treatment for ADHD with stimulant medication using longitudinal data from the National Longitudinal Survey of Canadian Youth (NLSCY), and a unique policy experiment which expanded insurance coverage for drugs in Quebec in 1997. Our study improves on the previous literature in many respects. First, we have a large sample of children who have been followed from 1994 to 2008. We are able to observe medium term outcomes such as grade repetition and math scores, as well as long term outcomes like graduation from high school and whether children ever attended college. Moreover, we know whether children were taking stimulant medication as of each wave. An important feature of the NLSCY is that all children were assessed for ADHD symptoms, so we do not have to deal with selection into diagnosis. A third innovation is that we are able to exploit exogenous variation in the availability of drugs due to the policy experiment. Fourth, in our analysis of medium term outcomes we are able to use individual fixed effects to control for unobservable differences between children that might influence both treatment and outcomes.

We find that the introduction of the prescription drug insurance program increased the use of stimulants in Quebec relative to the rest of Canada. However, we find no evidence that the performance of children with ADHD improved. In fact, the increase in medication use among children with ADHD is associated with increases in the probability of grade repetition, lower math scores, and a deterioration in relationships with parents. When we turn to an examination of long-term outcomes, we find that increases in medication use are associated with increases in the probability that a child has ever suffered from depression and decreases in the probability of post secondary education among girls.

Background

In view of the importance of ADHD and the fact that stimulant medications have been used for many years, it is perhaps surprising that most of the evidence regarding their efficacy relates to short time horizons. Controlled studies suggest that medication improves attention, short-term memory, performance on quizzes, homework completion, and note-taking (Douglas, 1999; Bedard et al, 2007; Pelham et al. 1993; Evans et al, 2001). It is often assumed that these improvements will translate into future academic gains, but few studies actually track children longer than a few months. Moreover Schachter et al. (2001) argue that the positive short-run effects on attention and behavior may be over-estimated given publication bias towards positive findings. An additional concern is that the doses that yield the most desirable behavior may not be calibrated to achieve the greatest possible improvement in cognitive functioning (Wigal et al., 1999).

One of the most widely known longer term studies of the effects of medication for ADHD is based on the U.S. National Institute of Mental Health 14 month Multimodal Treatment study (MTA). It is important to note that this study did not compare medication to non-treatment, instead, the MTA compared different types of treatment. Specifically, the MTA randomized 579 children with ADHD into four arms: Stimulants alone; behavioral therapy alone; stimulants plus behavioral therapy; or usual community care, which involved treatment with stimulants but with possibly less than optimal dosages. Blinded classroom observations did not find any significant differences in behavior between the four groups. At the end of 14 months, 49.8% of children reported mild side effects, 11.4% reported moderate side effects, and 2.9% reported severe side effects (The MTA Cooperation Group, 1999).

Molina et al. (2009) discuss a long-term follow up of children from the MTA study which included 436 of the original study children and 261 "controls" who were randomly selected from the same schools and grades 24 months after the original study began and matched with treatment children by age and gender. They find that 6 to 8 years following the initial intervention, there were still no differences between the children in the four treatment groups. They also find that the treatment children were worse off than the "controls" on virtually every measure but it is important to note that these controls were not part of the original randomized design so this comparison does not constitute an experimental evaluation of the long term benefits of drug treatment compared to non-treatment. Of those originally assigned to take medications, 62% had stopped taking them by the time of the follow up which is remarkable in itself since it suggests dissatisfaction with the drug regimen. However, adjusting for this attrition did not affect the differences between treated children and control children.

Barbaresi et al. (2007) follow 370 children with ADHD from a 1976-1982 birth cohort study. They obtained the complete school record, as well as medical records with information about stimulant use for each child. They found that in this sample of children with ADHD diagnoses, longer duration of stimulant use was associated with reductions in absences and retention in grade but had no effect on school dropout. However, endogeneity of stimulant use makes these results difficult to interpret. If the children with the worst attention difficulties were most likely to take medication, then any positive effects of medication would be biased towards zero. Alternatively, if children from the best backgrounds were most likely to take stimulants properly, then this might bias the analysis towards finding a positive effect.

Zoega et al. (2009) use registry data from Iceland, which has a measured prevalence of ADHD and a usage of stimulant medication that is similar to the U.S. They linked information from medical records to a data base of national scholastic examinations for children born between 1994 and 1996 who took standardized tests at fourth and seventh grade. In order to deal with the endogeneity of treatment, they include only children who were "ever treated" between the ages of 9 to 12, and focus on whether they were treated sooner or later. They find that children with ADHD suffered declines in test taking relative to other children, but that ADHD children who started medication earlier experienced slower declines than those who started medication later. Again, this design suffers from endogeneity, this time in terms of the choice of when treatment was started. It is possible, for instance, that children start medication in response to some crisis, and then experience reversion to their mean performance.[2]

[2] Another issue is that the authors define the start of therapy to be the first prescription after a period of at least 11 months without previous prescriptions for ADHD. This suggests that some of the "later starters" may in fact have started ADHD drugs earlier and then stopped them again.

Scheffler et al. (2009) use data from the Early Childhood Longitudinal Study—Kindergarten Class of 1998-1999 to examine the effect of medication use on standardized math and reading test scores for 594 children with ADHD. They estimate first differenced models in order to control for constant aspects of the child's background. A limitation of their data is that questions about medications were asked only in fifth grade, so it was assumed that children who were not taking medication at fifth grade had never taken it. They find that children with ADHD who took medication had higher mathematics and reading scores than other children with ADHD, though they still lagged behind their non-ADHD peers.   However, if children with ADHD are on different trajectories then their non-ADHD peers, then it is not clear that estimating the model in first differences will adequately control for the endogeneity of medication use.

Dalsgaard et al. (2013) use Danish registry data and variations in the prescription patterns of physicians to identify the effect of ADHD medication on hospital contacts, criminal activity and a limited set of school performance measures. They find that physician treatment patterns vary significantly, and that among children who receive treatment, hospital contacts decrease as do the number of interactions with police. While they find little difference in test scores for treated versus non-treated children, they note that treated children are less likely to take the exam. One limitation of their study is that higher income children were significantly more likely to go to doctors who prescribed medication more frequently which suggests that the probability of receiving a prescription was correlated with economic status.

Our study provides new evidence regarding the medium- and long-term effects of stimulants use for ADHD in a nationally representative sample of Canadian children by taking advantage of a policy experiment that expanded access to these drugs.[3] In 1997, the Canadian province of Quebec adopted a mandatory prescription drug insurance law.[4] Before 1997, many residents of Quebec received private prescription drug insurance from their employers while others went without drug insurance. The new law stipulated that all Quebecers had to be insured. Those who did not have insurance through their employer were required to participate in a new provincial public plan (Morgan, 1998). Premiums and deductibles were scaled according to income and some segments of the population received coverage for free including children whose parents were covered. Premiums were collected along with the filing of the Quebec tax return to ensure compliance with the law (Pomey et al 2007). Details on the premiums, deductibles and co-insurance rates over time are presented in the data appendix.

As a result of the insurance mandate and public plan, drug insurance rates rose quickly in Quebec. Using data from the National Population Health Survey and Community Health and Social Survey, both of which contain information on whether

[3] Cuellar and Markowitz (2007) adopt a somewhat similar identification strategy, examining the effects of increases in access to medication that occurred as a result of expansions of Medicaid coverage on rates of suicide, injury, and crime in eligible populations.

[4] Quebec implemented a subsidized day care program in September of that same year. In the first few years the program focused on older children (4-6) and expanded to include younger children later on (Baker et al 2008). To ensure that our instrument is not conflating the two programs we replicate our estimates focusing on children who are older than the day care ranges by the time the daycare program took place. Our main results are quite similar in this specification.

or not individuals hold prescription drug insurance, [5] we calculate coverage rates in both Quebec and in the rest of Canada. Whereas the rate of drug insurance coverage pre-reform in 1996 was 55%, it jumped to 84% in 1998 and continued to rise to 89% by 2003. Drug coverage rates in the rest of Canada averaged 65% in 1996 and rose slowly over time to an average of 76% by 2003 (Table 2-1). Overall the jump in Quebec far exceeds the rise in coverage taking place in the rest of the country as Quebec was the only province that instituted a universal coverage mandate.

Table 2-1. Changes over time in prescription drug insurance rates in Canada versus the rest of Canada

	Pre-Reform	Post -Reform		
Year	1996	1998	2002	2003
Quebec	55%	84%	86%	89%
Rest of Canada	65%	72%	74%	76%

Our identification strategy, then, is to first explore the increase in the use of stimulants that accompanied the increase in drug coverage[6] and then to relate the increase in drug use to medium and long-run child outcomes. Since it is possible that there were divergent trends in outcomes in Quebec and Canada which were independent of the introduction of the prescription drug law, we focus on the effects of the law on children who had high levels of ADHD symptoms prior to the passage of

[5] The NPHS (1994, 1996 and 1998) and CCHS (2002, 2003) are both publicly available data sets that ask questions about prescription drug coverage. The NLSCY, the main source of data for our analysis does not ask questions on prescription drug coverage.

[6] Quebec's public plan formulary explicitly lists Ritalin as covered. The reimbursement for the drug the price for 100 20mg tablets was $53.06.

the law. The overall argument is that if an expansion in drug use is beneficial, then we should see an improvement in the performance of children with ADHD in Quebec relative to the rest of Canada.

Data

We use data from the NLSCY, a national longitudinal data set which began with an initial sample of children ages 0 to 11 and their families in 1994. In the second wave of data collection in 1996, 15,871 of these children were surveyed (a reduced sample due to budget restrictions). We use the children born in 1985 or later who appear in both the 1994 and 1996 surveys as the base sample for this study. Follow-up surveys were conducted biannually up to 2008, producing up to 8 potential survey responses for each child. For responses pertaining to children under age 16, the survey collected information from the person most knowledgeable (PMK) about the child, while older children (16 and older) were responsible for completing the survey themselves.

We employ distinct approaches to evaluating the medium and long-term effects of stimulant use, and our sample depends on the approach in question. To investigate medium-term outcomes, we exploit the panel nature of the NLSCY and restrict the sample to observations collected at ages 0 through 16. For the oldest children in the sample – those born in 1985 or 1986 – we are able to observe up to 3 observations per child, while we use up to 7 survey responses for the youngest children. Our medium-term outcomes are not collected for all ages, however, and we further restrict our medium-term base sample as data availability requires.[7] The data appendix (Appendix 2) provides information about the maximum number of observations potentially available for each measure, and the number actually available given attrition.

For the long-term analysis we focus on outcomes that are reported by the youths themselves at ages 16 and later. For the most part, these are measured only once for each child, like high school graduation. With the exception of the self-assessed depression score – which we construct by averaging all available scores for each child in order to better capture whether the child was ever depressed and the persistence of depression – variables are defined according to their last observed value. Our long-term outcomes sample therefore consists of children aged 0-9 in 1994 who remain in the sample until at least age 16, tracked through 2008, with one observation per child. Due mostly to attrition, the base long-term sample is composed of 8,643 children born in 1985 or later, surveyed in both 1994 and 1996, and followed thereafter.

We measure ADHD using questions that are asked to parents about symptoms of ADHD. ADHD is always diagnosed through the use of questions similar to those included in the survey. Parents are asked to rate on a scale of 0 to 2 how often their child demonstrates five behaviors common among those who suffer from ADHD. Answers to these five questions are summed to produce an ADHD score that ranges between 0 and 10, where higher scores indicate a higher level of ADHD symptoms.

[7] Most of the short-term behavioral outcomes are only collected at ages 2 to 11 years. The educational outcomes are only available for school-aged children, and thus are collected starting at age 6. The question assessing the quality of the child's relationships are asked for children aged 4-9.

The questions used are listed in the data appendix, along with the questions used to construct all outcome variables. One of the strengths of the NLSCY data for this analysis is that these screener questions are administered to all children aged 2 to 11 years old, rather than to only diagnosed cases. We use the ADHD screener score collected in 1996 as our measure of the child's ADHD symptoms. Using the 1996 measurement allows us to obtain a pre-policy measurement of the severity of any child's ADHD symptoms.

Our information on stimulant use for both the medium- and long-term analyses is derived from a survey question that asks whether the child takes, "any of the following prescribed medication on a regular basis: Ritalin or other similar medication." This question is asked about all children age 15 and younger. Approximately 9 percent of sample children in Quebec, and 5 percent of sample children in the rest of Canada report ever having used stimulants. Stimulant use has increased slowly in Canada from less than 2 percent in 1994 to around 4 percent in 2008. Figure 2-1 shows that in Quebec, stimulant use tracked the rest of Canada closely prior to the policy change, but began to increase significantly following the policy change in 1997.

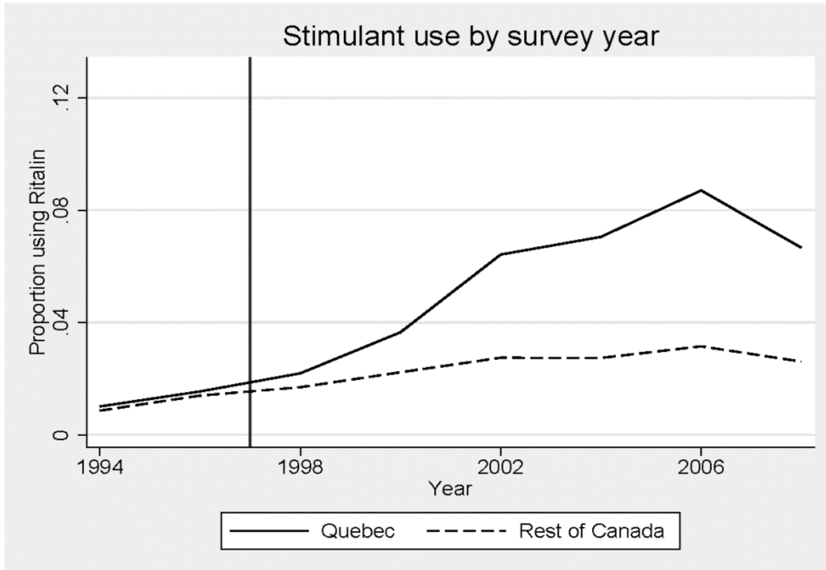

Figure 2-1. Stimulant use by survey year for Quebec versus the rest of Canada

Figure 2-2 provides additional evidence that the policy change led to significant increases in stimulant use. For all respondents who indicate ever using stimulants, we plot the fraction that commenced use in each cycle. If stimulant use is unrelated to outside factors, then uptake rates across survey cycles should exhibit a more-or-less smooth trend, with approximately equal proportions commencing use in any one year, peaking when the sample has the most children at peak diagnosis ages (6-10) and declining as the sample ages and diagnosis becomes less frequent. This is the pattern we see for children living outside of Quebec. For children in Quebec, however, there is a distinct spike in uptake rates in 2000 and 2002, following the policy change.

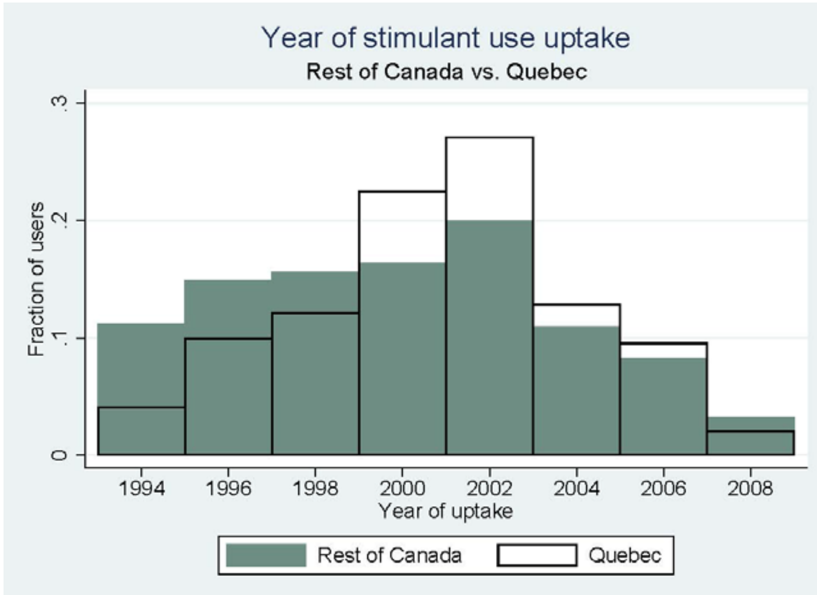

Figure 2-2. Stimulant uptake patterns in Quebec versus rest of Canada

The NLSCY also asks about other chronic conditions, some of which could also have been affected by increased drug coverage. Specifically, the survey asks whether, "a health professional has ever diagnosed any of the following long-term conditions…" where the listed conditions include: any type of allergy, bronchitis, heart conditions, epilepsy, cerebral palsy, kidney conditions, mental handicaps, learning disabilities, attention deficit disorder, emotional or psychological difficulties, eating disorders, autism, migraines, or any other chronic condition. We use these questions to test the robustness of our findings in two ways. First, we exclude children who had other (physical) chronic conditions from the sample and repeat our analyses.

Second, we examine children with asthma who may have gained access to, "Ventolin, inhalers or puffers for asthma" with expanded drug coverage. The increase in stimulant use was particularly pronounced relative to other medications such as the use of inhalers for asthma which did not increase disproportionately in Quebec relative to the rest of Canada (Figure 2-3). Thus, although the law was intended to increase access to all types of necessary medications, it seems to have had a disproportionate impact on prescriptions for stimulants.

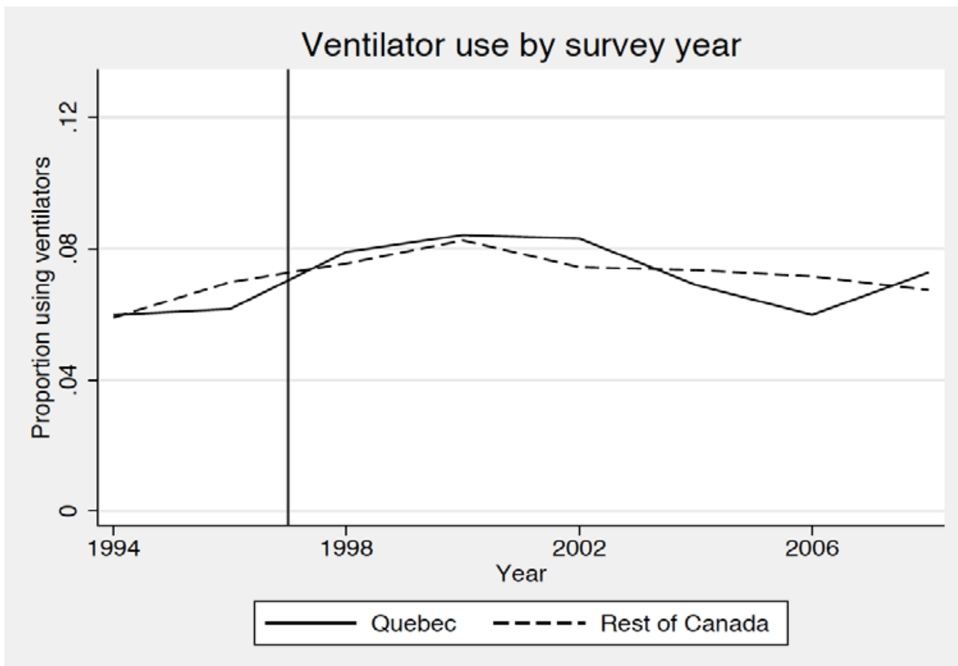

Figure 2-3. Ventilator use by survey year for Quebec versus the rest of Canada

We focus on outcomes that are intended to capture the child's behaviour, emotional state, and human capital accumulation in both the medium and longer run. The analysis of our medium-term outcomes involves a panel analysis of repeated observations over time for the same child They include: an unhappiness score, a rating of the child's relationship with his or her parents over the past 6 months[8], whether the child repeated a grade in the past two years and a mathematics score which is age-standardized to have a mean of zero and a standard deviation of one. Mathematics tests were administered in schools to children in grades two through ten and are based on the Canadian Achievement Tests.[9]

While the medium-term analysis is conducted using multiple outcome values for each child collected over time, the long-term analysis only employs one observation for each child. The long-term outcomes we consider include: an indicator for high school graduation, and indicator for having attended or graduated from a post-secondary institution, and a self-assessed depression score composed of six questions asked of all respondents aged 16 and older. In the case of the self-assessed depression score, we average all available self-assessed scores collected as of 2008 in order to construct an overall measure of the child's adolescent depression symptoms. The educational outcomes measure, by wave 8, whether the child had graduated from high school and whether he or she ever attended any post-secondary education.

Descriptive statistics for stimulant use, the outcome variables, and key independent variables for both the medium and long-term samples are shown in Table 2-2 (referred to as *Samples 1* and *2*, respectively). The table shows means separately for Quebec and the rest of Canada. The increase in stimulant use in Quebec vs. the rest of Canada is apparent in the first half of the table, and the second half of the table

---

[8] The relationship questions are indicators that equal 1 if the PMK has reported that the child has gotten along with the person in question "quite well" or "very well" over the previous six month period.

[9] The NLSCY began collecting a reading test score in its first three cycles but dropped this measure in subsequent cycles.

shows that a much larger number of children had ever used stimulants in Quebec by the end of the sample period. It is apparent that there are some differences in mean outcomes in Quebec vs. the rest of Canada, though the baseline child and family characteristics are fairly similar.

Table 2-2. Stimulant use, ADHD symptoms, and child and family characteristics

Sample 1 - Outcomes observed before age 16		
Variable	**Quebec**	**Rest of Canada**
ADHD score in 1996	2.83 (2.42)	2.65 (2.31)
Stimulants, waves 1 and 2	0.016	0.014
Stimulants, waves 3 and up	0.049	0.023
Medium-term Outcomes		
Unhappiness Score (6 pt)	0.47 (0.80)	0.53 (0.89)
Parent relationship	0.93 (0.26)	0.85 (0.36)
Standardized Math Score	10.02 (4.49)	8.27 (4.37)
Repeat a grade since last interview	0.07	0.02
Child and Family Characteristics		
Child is male	0.51	0.51
Child is first born	0.55	0.50
Permanent Household income	$58,958(33087)	$64,518(36938)
Two-parent household	0.82	0.84
Family size	3.99 (0.93)	4.12 (0.93)
Mother age at birth	27.98 (4.78)	28.14 (5.13)
Mother high school grad	0.85	0.90
Mother is working	0.71	0.75
PMK is an immigrant	0.04	0.09
Number of children in sample 1	2,478	10,471
Number of obs. In sample 1	10,622	44,617

Table 2-2. Cont.

Sample 2 - Outcomes observed after age 16		
Variable	**Lives in Quebec (cycle 1)**	**Rest of Canada (cycle 1)**
ADHD score in 1996	2.80(2.44)	2.54(2.32)
Ever Stimulants	0.09	0.05
# Surveys used Stimulants, given ever used	2.21(1.26)	2.06(1.26)
Post-1997 Years used Stim., given ever used	1.91(1.30)	1.56(1.22)
Long-term outcomes		
Self-assessed depression score (36 pt)	5.84(4.73)	6.53(4.99)
High school grad	0.85	0.90
Some post-secondary	0.77	0.65
Child and Family Characteristics		
Child is male	0.50	0.50
Child is first born	0.53	0.46
Permanent household income	$58,711 (34,333)	$64,669 (37,075)
Two-parent household in 1994	0.89	0.88
Family size in 1994	3.93 (0.89)	4.06 (0.90)
Mother age at birth	27.77 (4.59)	27.96 (5.09)
Mother high school grad in 1994	0.82	0.87
PMK is an immigrant	0.07	0.05
Number of children in sample 2	1,654	6,989

Standard errors of continuous variables in parentheses.

Methods

We begin by estimating the effect of the policy change on the use of stimulants in a difference-in-difference framework. The estimating equation takes the form:

$$Stim_{it} = \alpha + \beta Post_{it-1} + \lambda Que_{it-1} + \varphi Post_{it-1} * Que_{it-1} + X_{it}\Pi + \tau_t + p_i + \varepsilon_{it} \qquad (1)$$

where $Stim_{it}$ is a dichotomous variable equal to one if the PMK reports that child *i* is currently taking stimulant medication in year *t*, τ are survey year fixed effects and *p* are province fixed effects. $Post_{it}$ is a variable that identifies those survey responses collected from children after 1996, Que_{it} identifies responses from children in Quebec, and their interaction indicates the treatment group. In this specification, we compare children in Quebec to children in other provinces, before and after the policy change. The vector *X* includes family income, whether the person most knowledgeable about the child is an immigrant, whether the person most knowledgeable about the child (the survey respondent) is male or female, the sex of the child, birth order, family size, whether there are two parents present in the family, the mother's age at birth, whether the mother had a teen birth, and child-age dummies. To allow for delayed uptake in medication treatment, as well as time for the medication to take effect, we lag the policy change variable by one period (both the province of residence and the indicator for being post policy change).We expect a positive coefficient estimate on the $Post_{it}$ * Que_{it} interaction term, implying that increased access in post-reform Quebec led to expanded use of stimulant medication.

A limitation of the difference-in-differences approach is that there may be post 2007 differences in outcomes between Quebec and the rest of Canada for other reasons. Therefore, we focus on a triple difference specification that focuses on those children most likely to benefit from increased stimulant use in response to the policy change: Those with the worst initial ADHD symptoms. These models add an additional level of interaction terms to equation (1) – the ADHD score for the child between the ages of 2 to 11, measured in 1996 (pre policy change) – in order to estimate a difference-in-difference-in-difference (DDD) model, comparing children with worse underlying ADHD symptoms (measured before the reform) in post-reform Quebec to other children. This model is specified as:

$$\begin{aligned} Stim_{it} &= \alpha + \beta Post_{it-1} + \lambda Que_{it-1} + \gamma ADHD96_i \\ &+ \eta Que_{it-1} * ADHD96_i + \varphi Post_{it-1} * Que_{it-1} + \delta Post_{it-1} * ADHD96_i \\ &+ \boldsymbol{\theta Post_{it-1} * Que_{it-1} * ADHD96_i} + X_{it}\Pi + \tau_t + p_{it} + \varepsilon_{it}, \end{aligned} \tag{2}$$

where $ADHD96_i$ is the child's 1996 ADHD symptom score.[10] Using this approach, we are able to isolate the effect of the reform on stimulant use among children with worse ADHD symptoms, net of any pre-existing differences in stimulant use across time, geography, and severity of symptoms. In this specification, we expect that the estimate of θ should be positive.

When we examine medium term outcomes, we focus on versions of equations (1) and (2) that include child specific fixed effects. In these models, the effects are identified through changes in stimulant use for the same child before and after the policy change. The ability to control for child fixed effects obviates concerns about possible changes in the sample of children over time.

We use the same DDD framework (equation (2)) to examine the effect of the policy change on outcomes: if stimulant use improves outcomes, and children with

[10] Currie and Stabile (2007) show non-parametric Lowess plots which indicate that short-term test scores and grade repetition vary approximately linearly with ADHD scores, and that the 90th percentile of the ADHD score (which corresponds approximately to a threshold for diagnosis) is similar in Canada and the U.S . We therefore use linear average ADHD scores in our analysis.

worse symptoms are more likely to be treated post reform, then children with worse ADHD symptoms in post-reform Quebec – should demonstrate post-reform improvements in outcomes relative to their peers.

In order to examine the longer-term effects of an increase in stimulant use, we next use the sample with one long-term observation per child and estimate a quasi-first stage regression where the dependent variable is an indicator that equals 1 for children who ever reported using stimulant medication between ages 0 and 15 (*EverRit)*. We construct a policy exposure variable intended to capture the number of years the child was eligible for the new prescription drug regime: The total number of under age 16 survey responses for the child that occurred post 1996 (*PostYrs*). We then interact this lifetime exposure window variable with a Quebec indicator and the 1996 ADHD symptom score to create a parallel to (2):

$$EverRit_i = \alpha + \beta PostYrs_i + \lambda Que94_i + \gamma ADHD96_i$$
$$+\eta Que94_i * ADHD96_i + \varphi PostYrs_i * Que94_i + \delta PostYrs_i * ADHD96_i \quad (3)$$
$$+\boldsymbol{\theta PostYrs_i} * \boldsymbol{Que94_i} * \boldsymbol{ADHD96_i} + X94_i\Pi + p94_i + \varepsilon_i$$

Equation (3) is estimated using one observation per child and includes measures that are constructed at different periods in the child's life. Here the vector *X* includes controls measured as of 1994. The maximum number of years that a child can be treated depends on his or her age in year 1 of the survey (1994). We include age/cohort dummies to control for the fact that different children will be observed for different lengths of time. After estimating the relationship between lifetime stimulant use and exposure to the policy, we use equation (3) to examine the relationship between stimulant use and long term outcomes.

We perform a number of robustness checks to control for other health and learning disabilities that the child may have, as well as to specifically control for other contemporaneous policy changes that occurred in Quebec over this period. We discuss these checks following the presentation of our main results.

Results

We first examine the effect of the policy change on the probability of stimulant use in our sample as well as the relationship between exposure to the policy change and the number of years that a child used stimulants. Table 2-3 presents the results. Columns 1 and 2 report the difference-in-differences results without and with child fixed effects. In both cases we see an increase in the probability of using stimulants of approximately 2.5 percentage points for children in Quebec after the policy change. Columns three and four of Table 2-3 report the triple difference estimates (the D-D interacted with the child's 1996 ADHD score). Here the preferred fixed effect estimate suggests an increase of approximately 0.43 percentage points with each one unit increase in ADHD scores, which is quite similar to the OLS estimates without fixed effects of 0.48 percentage points. At the average ADHD score, this is a 1.15 percentage point change in stimulant use compared to the average baseline number of children on stimulants of 2 percent.

Table 2-3. Effects of law change on stimulant use

	Outcome: Uses Stimulants					Outcome: Ever Used Stimulants		
	(1)	(2)	(3)	(4)			(5)	(6)
	DD - FE	DD - No FE	DDD - FE	DDD - No FE			DD	DDD
After 1997	-0.0072**	-0.0092**	-0.0172**	-0.0133**		U16 Survey years after 1997	0.0003	0.0078**
	(0.0027)	(0.0036)	(0.0031)	(0.0040)		*(Elig Yrs)*	(0.0036)	(0.0033)
Quebec	-0.0118	0.0154**	-0.0256	0.0101**		Quebec in 1994	-0.0120	0.0181
	(0.0209)	(0.0027)	(0.0243)	(0.0039)			(0.0121)	(0.0253)
After 1997 * Quebec	**0.0247****	**0.0287****	**0.0123****	**0.0159****		**Elig Yrs * Quebec**	**0.0196****	**0.0073**
	(0.0046)	**(0.0062)**	**(0.0030)**	**(0.0052)**			**(0.0034)**	**(0.0085)**
1996 ADHD Score	-	-	-	0.0105**		1996 ADHD Score	-	0.0403**
				(0.0006)				(0.0047)
After 1997*1996 ADHD Score	-	-	0.0039**	0.0014		Elig Yrs*1996 ADHD Score	-	-0.0038**
			(0.0008)	(0.0010)				(0.0012)
Quebec*1996 ADHD Score	-	-	0.0051	0.0004		Quebec*1996 ADHD Score	-	-0.0201**
			(0.0083)	(0.0009)				(0.0098)
Aft. 1997*Que.*ADHD Sc.	**-**	**-**	**0.0043****	**0.0048****		**EligYrs*Que.94*96 ADHD Sc.**	**-**	**0.0056**
			(0.0021)	**(0.0018)**				**(0.0032)**
N	55,239	55,239	55,239	55,239		N	8,643	8,643
Age Range	2-15	2-15	2-15	2-15		Age Range	0-9 in 1994	0-9 in 1994

Notes: Controls include: Year-of-birth fixed effect, age fixed effect, province fixed effect, family permanent income, indicator of pmk immigrant, male, first born, log family size, indicator for two-parent family, mother's age at birth, mother teen birth, indicator if pmk is male. Controls measured in each survey wave in columns 1-4, and in 1994 in columns 5 and 6. Standard errors in columns 1-4 are in brackets and are clustered at the province-year level. Standard errors in columns 5 and 6 are clustered at the cohort-province level. ** indicates significant at 95%.

We also estimate a similar "first stage" model for our longer-term analysis by examining the relationship between exposure to the policy change and ever taking stimulants, as described in equation (3). These results are presented in columns 5 and 6 of Table 2-3. The DDD estimate suggests a 0.56 percentage point increase in the probability of ever taking stimulants on a baseline of 4 percent, however, the coefficient is imprecisely estimated and is significant only at the 90 percent level of confidence. Again, while this is a fairly small overall change in stimulant use it reflects a large change relative to baseline.

Having established that the policy change resulted in a reasonably large change in the use of stimulants we now turn to examining both the medium and longer term consequences of this change. Table 2-4 presents the estimates for medium term outcomes. All columns include child specific fixed-effects. The difference-in-differences estimates suggest consistently worse outcomes post policy change in Quebec though, even with the inclusion of child fixed effects, these differences could possibly reflect divergent trends in Quebec and the rest of Canada. Therefore, we prefer to focus on the DDD estimates. These also suggest a significant negative effect of the policy change in terms of grade repetition, math scores, and relationships with parents. For example, the coefficient on the triple interaction in the "Did not repeat grade" model suggests that for each one unit increase in ADHD scores, the probability that a child progressed normally through school between waves fell post policy change by 1.28 percentage points on a baseline 93 percent progression rate.

Table 2-4. Child fixed effects estimate of exposure to policy on contemporaneous outcome

Dependant Variable:	**Did Not Repeat Grade**		**Math score**		**Unhappiness**		**Relationship w Parents**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After 1997	0.0247**	0.0207**	-0.3179**	-0.2836**	-0.2846**	-0.1977**	0.0227**	-0.0015
	(0.0063)	(0.0061)	(0.1072)	(0.1129)	(0.0176)	(0.0285)	(0.0094)	(0.0109)
Quebec	0.0578**	0.0225	0.2730	-0.1511	-0.0023	0.2054	0.0480	-0.0665
	(0.0282)	(0.0388)	(0.2410)	(0.2828)	(0.1133)	(0.1280)	(0.1157)	(0.1330)
After 1997 * Quebec	**-0.0581****	**-0.0228****	**-0.1883**	**-0.0694**	**0.1232****	**0.0769**	**-0.0353****	**-0.0014**
	(0.0068)	**(0.0076)**	**(0.1927)**	**(0.1929)**	**(0.0152)**	**(0.0475)**	**(0.0073)**	**(0.0099)**
After 1997*1996 ADHD Score	-	0.0016	-	-0.0141**	-	-0.0326**	-	0.0088**
		(0.0009)		(0.0056)		(0.0076)		(0.0021)
Quebec*1996 ADHD Score	-	0.0117	-	0.1369**	-	-0.0894**	-	0.0789**
		(0.0141)		(0.0557)		(0.0366)		(0.0348)
Aft. 1997*Que.*ADHD Sc.	**-**	**-0.0128****	**-**	**-0.0403****	**-**	**0.0172**	**-**	**-0.0124****
		(0.0016)		**(0.0062)**		**(0.0145)**		**(0.0033)**
N	44,968	44,968	32,515	32,515	36,458	36,458	22,554	22,554
Age Range	4-15	4-15	5-15	5-15	2-11	2-11	4-9	4-9

Notes: See Table 3. Models include child fixed effect. Standard errors clustered at the province-year level. ** indicates significance at the 95% level.

Turning to the long term outcomes, Table 2-5 shows estimates of equation (3). In the triple difference framework, the estimates suggest that the only long term effect is on unhappiness – there is no statistically significant effect of exposure to the policy on high school completion or post-secondary schooling among those with higher ADHD scores.

Table 2-5. Effects of exposure to the policy on long-term outcomes

Dependant Variable:	**Depression Score**		**High School grad**		**Some Post-sec**	
	(1)	(2)	(3)	(4)	(5)	(6)
U17 Survey years after 1996 *(Elig Yrs)*	0.3696** (0.0793)	0.4226** (0.1149)	-0.0028 (0.0047)	-0.0066 (0.0061)	-0.0005 (0.0096)	-0.0096 (0.0122)
Quebec in 1994	0.6756 (0.3817)	1.6972** (0.4795)	-0.2191** (0.0307)	-0.2430** (0.0448)	-0.1265** (0.0377)	-0.2068** (0.0662)
Elig Yrs * Quebec	**-0.0876 (0.0883)**	**-0.3073** (0.1413)**	**0.0498** (0.0097)**	**0.0675** (0.0164)**	**0.0666** (0.0142)**	**0.0868** (0.0219)**
1996 ADHD Score	-	0.3414** (0.0923)	-	-0.0206** (0.0062)	-	-0.0447** (0.0089)
Elig Yrs*1996 ADHD Score	-	-0.0288 (0.0274)	-	0.0019 (0.0018)	-	0.0046 (0.0026)
Quebec*1996 ADHD Score	-	-0.4340** (0.1242)	-	0.0125 (0.0133)	-	0.0382** (0.0163)
EligYrs*Que.94 *96 ADHD Sc.	**-**	**0.0867** (0.0402)**	**-**	**-0.0067 (0.0043)**	**-**	**-0.0084 (0.0048)**
N	6,493	6,493	4,676	4,676	4,676	4,676

Notes: Sample includes children 0-9 in 1994. See Table 2-3 notes. Standard errors clustered at the province-cohort level. ** indicates significance at the 95% level.

These estimates cast doubt on the idea that the diffusion of stimulant use improved academic outcomes among those with ADHD, and raise the possibility that children were actually harmed. There are several possible mechanisms that could be at work. First, many of the known side effects of stimulant use have to do with children's emotional wellbeing; direct effects on unhappiness or depression may therefore not be surprising. It is also possible that stimulants have direct effects on children's cognitive abilities, particularly if dosages are not optimized for the individual child. A second possible mechanism is that stimulant use might crowd-out other therapies or learning strategies that could be more beneficial to the child. A third possibility is that stigma associated with an ADHD diagnosis and stimulant use is harmful to the child. In order to further assess these possibilities we turn to a separate analysis by gender.

Estimates by Gender

There are well-documented differences in ADHD prevalence and in the use of stimulants between boys and girls: For example, Schwarz and Cohen (2013) find that 15% of U.S. boys and only 7% of U.S. girls have ever been diagnosed with ADHD. Figure 2-4 plots stimulant use rates for Quebec versus the rest of Canada separately for boys and girls. Due to NLSCY data release rules, we have pooled observations by two-survey year time periods. Thus, the first point in the graphs shows the rate of stimulant use indicated in 1994 and 1996 survey responses and it is our pre-policy observation; the remaining points represent stimulant use rates for post-policy years. What is clear is that while both boys and girls increased stimulant use substantially after the policy change, the effect is much larger among boys.

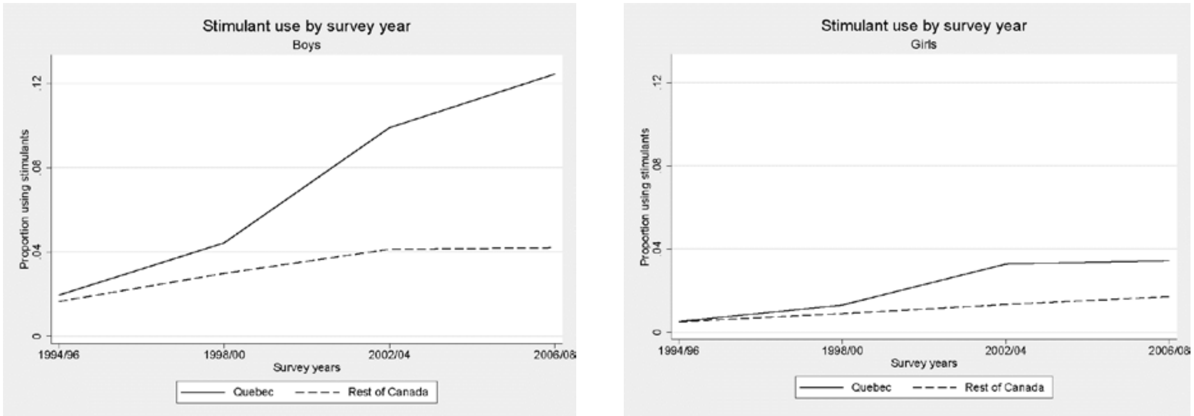

Figure 2-4. Trends in stimulant use by gender, Quebec vs. Rest of Canada

Table 2-6 shows our estimates of the effect of the policy change on take up of stimulants by gender. Column (1) shows that among boys, there was a strong increase in the use of stimulants in Quebec post policy change. However, column (2) shows that there was no differential impact among children with higher ADHD scores. In other words boys with low levels of ADHD symptoms were just as likely to take up stimulants as those with high ADHD scores post policy change, suggesting that the marginal boy taking stimulants had lower levels of ADHD symptoms post policy change. Columns (3) and (4) show the comparable estimates for girls. The story for girls is quite different, suggesting that the increase in stimulant use post policy change was concentrated among girls with high ADHD scores, and that there was no increase in usage among girls with low ADHD scores.

Table 2-6. Effects of the law change on stimulant use by gender

Outcome: Uses Stimulants			**Outcome: Ever Used Stimulants**		
	(1) Boys	(2) Girls		(3) Boys	(4) Girls
After 1997	-0.0255** (0.0052)	-0.0077** (0.0022)	U16 Survey years after 1997 *(Elig Yrs)*	0.0107** (0.0052)	0.0032 (0.0038)
Quebec	0.0070 (0.0325)	-0.0651 (0.0391)	Quebec in 1994	-0.0499 (0.0513)	0.0580** (0.0182)
After 1997 * Quebec	**0.0299** (0.0058)**	**-0.0031 (0.0042)**	**Elig Yrs * Quebec**	**0.0254 (0.0160)**	**-0.0043 (0.0040)**
1996 ADHD Score	-	-	1996 ADHD Score	0.0525** (0.0061)	0.0228** (0.0071)
After 1997*1996 ADHD Sc.	0.0049** (0.0012)	0.0022** (0.0009)	Elig Yrs*1996 ADHD Score	-0.0050** (0.0015)	-0.0018 (0.0016)
Quebec*1996 ADHD Score	0.0009 (0.0117)	0.0154 (0.0105)	Quebec*1996 ADHD Score	-0.0057 (0.0178)	-0.0293** (0.0092)
Aft. 1997*Que.*ADHD Sc.	**0.0004 (0.0025)**	**0.0078** (0.0021)**	**EligYrs*Que.94*96 ADHD Sc.**	**0.0021 (0.0056)**	**0.0076** (0.0021)**
N	27,971	27,268		4,333	4,310
Age Range	2-15	2-15		0-9 in 1994	0-9 in 1994

Notes: See Table 2-3. Columns 1 and 2 include child fixed effect and standard errors clustered at the year-province level are reported in parentheses. Standard errors in columns 3 and 4 are clustered at the cohort-province level. ** indicates significance at the 95% level.

Table 2-7 shows estimates of the medium-term impacts of the policy on boys and girls. The estimates for girls suggest that the negative effects of the policy change are confined to girls who had initially high ADHD scores, which makes sense, since these are the only girls who increased stimulant use as a result of the policy change. Among these girls, there are increases in unhappiness, deteriorations in relations with parents, and reductions in math scores.

Table 2-7. Child fixed effects estimate of exposure to policy on contemporaneous outcomes, by gender

Dependant Variable:	Unhappiness	Rel. w. Parents	Did Not Rep. Gr.	Math Sc.
Boys	(1)	(2)	(3)	(4)
After 1997 * Quebec	**0.1300***	**-0.0106**	**-0.0439****	**-0.1299**
	(0.0666)	**(0.0136)**	**(0.0091)**	**(0.1712)**
Aft. 1997*Que.*ADHD Sc.	**-0.0047**	**-0.0120****	**-0.0142****	**-0.0257**
	(0.0203)	**(0.0057)**	**(0.0018)**	**(0.0134)**
N	18,484	11,457	22,719	16,191
Girls				
After 1997 * Quebec	**0.0258**	**0.0090**	**-0.0124**	**-0.0063**
	(0.0344)	**(0.0152)**	**(0.0128)**	**(0.2027)**
Aft. 1997*Que.*ADHD Sc.	**0.0430****	**-0.0127****	**-0.0058**	**-0.0588****
	(0.0120)	**(0.0060)**	**(0.0031)**	**(0.0210)**
N	19,974	11,097	22,249	16,324
Age Range	2-11	4-9	4-15	5-15

Notes: See Table 2-4. Models include child fixed effect. Standard errors clustered at the province-year level. ** indicates significance at the 95% level.

For boys, the DD coefficient estimates on the Quebec post-policy indicator suggest that the policy change is associated with an increase in grade repetition among all boys; however, this result could be part of a general trend towards greater use of grade repetition among boys in Quebec. Among boys with higher ADHD scores, there are deteriorations in relations with parents and an even larger increase in grade repetition post policy change. These estimates suggest that the upswing in stimulant use following the policy change had larger negative effects on boys with ADHD than on those without, even though stimulant use increased for boys with *and* without ADHD symptoms. It is possible that the negative effects of increased stimulant use – for example the crowding out of other types of intervention –were greater for boys with more severe ADHD symptoms since they had greater need for these interventions.

Turning to the results for long-term outcomes which are shown in Table 2-8, the estimates suggest that the policy impacted girls with ADHD but not boys. Specifically girls with higher initial ADHD scores were more likely to have suffered from depression, and less likely to have any post-secondary education, the more they were exposed to the post-policy change regime.

Table 2-8. Effects of exposure to the policy on long-term outcomes

Dependant Variable:	Depression Score		High School grad		Some Post-sec	
Boys	(1)	(2)	(3)	(4)	(5)	(6)
Elig Yrs * Quebec	**-0.0223**	**0.1049**	**0.0591****	**0.0836****	**0.0735****	**0.0744****
	(0.0841)	**(0.1932)**	**(0.0171)**	**(0.0242)**	**(0.0187)**	**(0.0271)**
EligYrs*Que.94*96 ADHD Sc.	**-**	**-0.0317**	**-**	**-0.0091**	**-**	**-0.0017**
		(0.0630)		**(0.0047)**		**(0.0058)**
N	3,213	3,280	2,259	2,259	2,259	2,259
Girls						
Elig Yrs * Quebec	**-0.1428**	**-0.6068****	**0.0439****	**0.0549****	**0.0619****	**0.0942****
	(0.1450)	**(0.2221)**	**(0.0137)**	**(0.0210)**	**(0.0195)**	**(0.0237)**
EligYrs*Que.94*96 ADHD Sc.	**-**	**0.2025****	**-**	**-0.0041**	**-**	**-0.0150****
		(0.0541)		**(0.0049)**		**(0.0050)**
N	3,280	3,280	2,417	2,417	2,417	2,417

Notes: See Table 2-5. Standard errors clustered at the province-cohort level. ** indicates significance at the 95% level.

Robustness checks

We performed a number of specification checks to test the robustness of our findings. First we re-estimated the triple difference models excluding children with physical chronic conditions. These children may have benefited from increased access to other medications, which could have affected outcomes as well. However, the estimates are quite similar in this sub-sample. Estimates are shown in Appendix 2, Table 4. Since asthma is the most common physical chronic condition among the children in our sample, we also asked whether there was an increase in ventilator use following the policy change. Such an improvement in the treatment of asthma could have had independent effects on children's outcomes. We find insignificant coefficients on the DDD estimates for an increase in ventilator use, unlike our estimates for increases in the use of stimulants.[1] The results are reported in Appendix 2, Table 4.

A second possible concern is that our triple difference, despite focusing on the children who were most likely to benefit from stimulant use, could be picking up the effect of contemporaneous policy changes. One important policy change that happened around the same time was the introduction of subsidized day care in Quebec. Baker et al (2008) find negative effects of exposure to subsidized day care programs in Quebec on a number of child outcomes. To make sure that we are not confounding these two policy changes, we re-estimated our models limiting the sample to children born in 1991 or earlier – that is, to those unaffected by the childcare policy change. Although this restriction greatly reduces the sample size, we continue to find negative effects on math scores and grade repetition. These estimates are reported in Appendix 2, Table 4.

We have focused above on unhappiness and depression given that these are the most prevalent mental health conditions (besides ADHD) in our sample. However, given that other measures of mental health are available, we also created a composite mental health measure. Using the unhappiness score, along with similarly constructed scores measuring anxiety and physical aggression, we standardized and then averaged the scores to construct an overall composite mood score. We continued to find positive

[1] We use an indicator for asthma diagnoses as the third difference in this robustness check.

and significant coefficient estimates (reflecting an increase in mood and behavioral problems) for girls. These results are also reported in Appendix 2, Table 4.

Finally we conduct a series of placebo tests using data excluding observations from Quebec. We define placebo policy change dates every two years, from 1995 to 2005, and policy change regions in Ontario, British Columbia, the prairie provinces (Alberta, Saskatchewan and Manitoba), and the maritime provinces (Newfoundland, Nova Scotia, New Brunswick and Prince Edward Island). We then estimate equations (2) and (3) for each placebo year-region combination, resulting in a total of 24 placebo DDD coefficient estimates for each model. We plot the distributions of these estimates in Appendix 2, Figures 1 and 2; in both cases, the vertical line denotes the DDD coefficient estimate derived from the model estimation using the true policy change in Quebec in 1997. The figures reveal that the true coefficient estimates fall in the tails of the placebo distributions, suggesting that if the policy had been ineffective – as the placebo changes are by definition – we would be very unlikely to have generated estimates as large in magnitude as those that we find. The lack of any systematic or robust relationship between the experiment and the stimulant use outcomes in the placebo context provides some confidence that we are not picking up a spurious correlation in the true policy experiment setting.

Discussion and Conclusions

This paper examines the effect of a "natural experiment" in Quebec that greatly expanded access to stimulant medication, and the take up of stimulants among children with ADHD. One might have anticipated that increases in access to medication would be associated with improved outcomes among these children. Instead, we actually find some evidence of negative effects. Some of these negative effects are consistent with known possible side effects of stimulant medication, especially depression.

We find little evidence of positive effects on academic outcomes or schooling attainment. In fact, we find deterioration in important academic outcomes including grade repetition and math scores. When we examine the effects of the policy by gender, we find that stimulant use among boys increased greatly, but that it increased equally among boys with high and low levels of initial ADHD symptoms. Among girls, the increase in stimulant use was more concentrated among children with initially high levels of ADHD symptoms. However, the increase in stimulant use among girls with ADHD was associated with increases in unhappiness and the probability that a girl had depression, decreases in math scores, and a decline in the probability of having any post-secondary education.

Our findings of potentially negative effects associated with the increase in stimulant medications use raise the question of mechanisms. How is it possible that an increase in the utilization of medication for ADHD could be associated with worse academic performance?

One possibility is that an increase in the availability of stimulants makes it more likely that a child will be treated for ADHD and that treatment triggers harmful social stigma or other consequences, such as being placed in special education.[2] A second possibility is that medication is a substitute for other types of cognitive and

[2] The NLSCY dataset does not include information on special education.

behavioral interventions that might be necessary to help the child learn. By making children less disruptive, ADHD medication could decrease the attention that they receive in the average classroom and reduce the probability that the child receives other needed services. A third possibility is that the medication itself, particularly if the dosage is not appropriately tailored, could have negative effects on emotional wellbeing and learning.

It is important to acknowledge that this is an ecological study which does not shed light on the question of whether optimal medication use could be beneficial. It is clear that many children use stimulant medication in a haphazard manner. For example, on average, among those who ever report using stimulants in our data, children use stimulants for about 30% of the survey years we observe them. Moreover, the average child who is ever reported to use stimulants switches twice over the observation period (between the time they are ages 4-7 and age 15, depending on how old they were in 1994). While it is possible that some of this churning is measurement error, recall that in the MTA most children had stopped taking medications 6 to 8 years after follow up. In addition, while we have no information about dosage, it seems likely that many children are taking doses of ADHD that are not calibrated to achieve optimal results, even in terms of short-term behavioral effects.

What our results do speak to, is the effect of a large increase in the use of ADHD medications in a community, given the usual standard of care available to Quebec children. In Quebec, as in the U.S., any doctor can prescribe stimulants, and it is not necessary to have expertise treating ADHD in order to do so. Hence, it is not surprising that some use is sub-optimal. Our results suggest that observers of the large increases in the use of medication for ADHD in Canada, the U.S., and other countries are right to be concerned.

REFERENCES

Aizer, A. (2009). Peer Effects, Institutions, and Human Capital Accumulation: The Externalities of ADD. *National Bureau of Economic Research Working Paper # 14354.*

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th Ed*.

Baker, M., Gruber, J. & Milligan, K. (2008). Universal Child Care, Maternal Labor Supply, and Family Well-Being. *Journal of Political Economy,*116, 709-745.

Barbaresi, W.J., Katusic, S.K., Colligan, R.C., Weaver, A.L., & Jacobsen, S.J. (2007). Modifiers of Long-Term School Outcomes for Children with Attention-Deficit/Hyperactivity Disorder: Does Treatment with Stimulant Medication Make a Difference? Results from a Population Based Study. *J Dev Behav Pediatr*, 28, 274-287.

Bedard, A, Jain, U., Johnson, S.H., & Tannock., R. (2007). Effects of Methylphenidate on Working Memory Components: Influence of Measurement. *J Child Psychol Psychiatry*, 28, 872-880.

Cuellar, A., & Markowitz, S. (2007). Medicaid Policy Changes in Mental Health Care and Their Effect on Mental Health Outcomes. *Health Economics, Policy and Law*, 2, 23-49.

Currie, J., & Kahn, R. (2012). Children with Disabilities: Introducing the Issue. In *Children with Disabilities*. Princeton: The Future of Children.

Currie, J., & Stabile, M. (2006). Child Mental Health and Human Capital Accumulation: The Case of ADHD, *Journal of Health Economics,* 25, 1094-1118.

Currie, J., & Stabile, M. (2009). Mental Health in Childhood and Human Capital. In J. Gruber (Ed.) *Perspectives on the Problems of Disadvantaged Youth*. Chicago: University of Chicago Press for NBER.

Centers for Disease Control and Prevention. (2005). Mental Health in the United States: Prevalence of Diagnosis and Medication Treatment for Attention-Deficit/Hyperactivity Disorder: United States, 2003. *MMWR Morb Mortal Wkly Rep.,* 54, 842-847.

Dalsgaard, S., Nielson, S. H., & Simonsen, M. (2013). Long-term Consequences of ADHD Medication Use for Children’s Human Capital Development, *Aarhus University Working Paper*.

Douglas, V.I. (1999).Cognitive Control Processes in Attention-Deficit/Hyperactivity Disorder. In H.C. Quay & A.E. Hogan (Eds.), *Handbook of Disruptive Behavior Disorders*. New York, NY: Plenum Press.

Evans, S.W., Pelahm, W.E,. Smith, B.H., Bukstein, O., Gnagy, E.M., Greiner, A.R., Altenderfer, L., & Baron-Myak, C. (2001). Dose-Response Effects of Methylphenidate on Ecologically Valid Measures of Academic Performance and Classroom Behavior in Adolescents with ADHD. *Exp Clin Psychopharmacol*, 9, 163-175.

Fletcher, J., & Wolfe, B. (2008). Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited. *Journal of Health Economics*, 27, 794-800.

Fletcher, J., & Wolfe, B. (2009). Long-Term Consequences of Childhood ADHD on Criminal Activities. *The Journal of Mental Health Policy and Economics*, 12, 119-138.

Gouvernment du Quebec. (2007*). La Politique du Medicament, Gouvernment du Quebec.* Retrieved from www.msss.gouv.qc.ca.

Griffin, S., Weatherly, H., Richardson, G., & Drummond, M. (2008). Methodlogical Issues in Undertaking Independent Cost-Effectiveness Analysis for NICE: The Case of Therapies for ADHD. *The European Journal of Health Economics*, 9, 137-145.

Joshi, S.V., & Adam, H.M. (2002). ADHD, Growth Deficits, and Relationships to Psychostimulant Use. *Pediatr Rev*, 23, 67-68.

Miech, R. , Avshalom Caspi, Terrie Moffitt, Bradley Entner Wright, and Phil Silva. "Low Socioeconomic Status and Mental Disorders: A Longitudinal Study of Selection and Causation During Young Adulthood," American Journal of Sociology, 104, 1999, 1096-1131.

Molina, B.S., Hinshaw, S.P., Swanson , J.M., et. al. (2009). The MTA at 8 Years: Prospective Follow-Up of Children Treated for Combined Type ADHD in a Multisite Study. *J Am Acad Child Adolesc Psychiatry*, 48, 484-500.

Morgan, S. (1998). Quebec's Drug Insurance Plan: A Prescription for Canada? *Centre for Health Services and Policy Research, University of British Columbia, Working Paper HPRU:98:2D.*

The MTA Cooperative Group. (1999). A 14-Month Randomized Clinical Trial of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder. *Arch Gen Psychiatry*, 56, 1073-1086.

Nagin, D., & Tremblay. R.E. (1999). Trajectories of Boys' Physical Aggression, Opposition, and Hyperactivity on the Path to Physically Violent and Nonviolent Juvenile Delinquency. *Child Development*, 70, 1181-1196.

Nissen, S.E. (2006). ADHD Drugs and Cardiovascular Risk, *N Engl J Med*, 354, 1445-1448.

Pelham, W.E., Carlson, C., Sams, S.E., Vallano, G., Dixon, M.J., Hoza, B.. (1993). Separate and Combined Effects of Methylphenidate and Behavior Modification on Boys with Attention Deficit-Hyperactivity Disorder in the Classroom. *J Consult Clin Psychol* 61, 506-515.

Polanczyk, G., De Lima, M.S., Horta, B.L., Biederman, J., Rohde, L.A. (2007). TheWorldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis. *American Journal of Psychiatry*, 164, 942-948.

Pomey, M.P., Forest, P.G., Palley, H., & Martin, E. (2007). Public/Private Partnerships for Prescription Drug Coverage: Policy Formulation and Outcomes in Quebec's Universal Drug Insurance Program, with Comparisons to the Medicare Prescription Drug Program in the United States. *The Millbank Quarterly*, 85, 469-498.

Schachter, H.M., Pahm, B., King, J. et al. (2001). How Efficacious and Safe is Short-Acting Methylphenidate for the Treatment of Attention-Deficit Disorder in Children and Adolescents? A Meta-Analysis, *CMAJ*, 165, 1475-1488.

Scheffler, R. M., Brown, T., Fulton, B., Hinshaw, S., Levine, P., & Stone, S. (2009). Positive Association Between Attention-Deficit/Hyperactivity Disorder Medication Use and Academic Achievement During Elementary School. *Pediatrics*, 123, 1273-1279.

Schwarz, A., & Cohen, S. (2013). ADHD Seen in 11 percent of U.S. Children as Diagnoses Rise. *New York Times*, March 31.

Therapeutics Initiative. (2008). What is the Evidence for Using CNS Stimulants to Treat ADHD in Children? *Therapeutics Letter*, 69, March-May.

U.S. National Institute of Mental Health. (2012). Attention Deficit Hyperactivity Disorder. Retrieved from http://www.nimh.nih.gov/health/publications/attention-deficit-hyperactivity-disorder/medications.shtml.

Wigal, T., Swanson, J., Regino, R., Lerner, M., Soliman, I., Steinhoff, K., Gurbani, S., & Wigal, S. (1999). Stimulant Medications for the Treatment of ADHD: Efficacy and Limitations. *Mental Retardation and Developmental Disabilities Research Reviews*, 5, 215-224.

Zoega, H., Rothman, K., Huybrechts, K., Olafsson, O., Baldursson, G., Almarsdottir, A., Jonsdottir, S., Haldorsson, M., Hernandez-Diaz, S., Valdimarsdottir, U. (2012). A Population Based Study of Stimulant Drug Treatment of ADHD and Academic Progress in Children. *Pediatrics*, 130, e53-e62.

CHAPTER 3

SUCCESSFUL SCIENTIFIC REPLICATION AND EXTENSION OF LEVITT (2008): CHILD SEATS ARE STILL NO SAFER THAN SEAT BELTS AND IMPROPER USE IS LIFE THREATENING

WITH NICOLAS ZIEBARTH

Introduction

In the US, as well as in over 90 countries worldwide, traffic safety regulations require use of specific approved child safety seats for children in automobiles (WHO, 2013). Statutory age and weight regulations have increased over time. Currently, all US states mandate the use of child safety seats. Forty-three states require the use of child safety seats until at least age four, but the variation in state regulations ranges from mandatory use until age 3 up to age 7. For first time offenders, the variation in fines for no use ranges from as low as $20 (West Virginia) up to $500 (Nevada), with a mass point around $100 (IIHS, 2013).

Despite the prevalence of these laws, there is evidence that child seats may not be any more effective than seatbelts at preventing children's death and injury. Using 1975 to 2003 data from the US Fatality Analysis Reporting System (FARS), Levitt (2008) shows empirically that the use of child safety seats does not significantly reduce the probability of a fatality in accidents relative to traditional seatbelts.[22] This finding casts serious doubts on the effectiveness of child safety seats, despite the common acceptance and large support of this traffic safety regulation. If child safety

[22] Levitt and Doyle (2010) show that this result holds not only for fatalities but also for injuries.

seats do not provide any safety improvement over standard seatbelts, then laws that require citizens to buy and use child seats, along with the costs of their enforcement, represent a costly and welfare-decreasing state regulation.

Because of the high practical relevance of this question this note has the following purpose: *(a)* to replicate Levitt's (2008) original findings, and *(b)* to test whether Levitt's (2008) findings also apply to a more recent time period, from 2004 to 2011. The latter contribution is important since major potential safety improvements have been generated by manufacturers and consumer trends in the last decade, such as the popularity of SUVs or increases in child safety seat quality. Also traffic regulation and speed limits have become more restrictive over time (IIHS, 2013). One would expect that the increases in vehicle safety conditions could strengthen Levitt's findings. On the other hand, simultaneous developments may reduce or offset these effects. Changes in the number of car owners and miles travelled or changes in driving behavior may all affect the utility of restraint types.

Thus, we update and enrich Levitt's (2008) analysis on two dimensions: First, the SUV "arms race"—which makes roads less safe since accidents involving passenger cars and SUVs have an increased fatality probability—could impact the effectiveness of child safety seat (White, 2004; Daly et al., 2006, Small and Van Dender, 2007; Li, 2012; Klier and Linn, 2012; Busse et al. 2013). Second, the misuse of child safety seats potentially impacts their effectiveness (Howland et al., 1965; Bull et al., 1988). It is estimated that more than half of child safety seats are improperly used (Children's Safety Network, 2005). We investigate how these two traffic phenomena—the SUV arms race and improper use of child seats—could mitigate or strengthen Levitt's (2008) findings on the effectiveness of child safety seat vs. traditional seat belt use. To our knowledge, this study is one of first to formally estimate the dangers of improper restraint use for children involved in fatal accidents.

Scientific Replication and Extension of Levitt (2008)

Replication of Levitt (2008): Data and Methods Used

Levitt (2008) makes use of US FARS data from 1975 to 2003. The dataset includes the universe of all accidents in which at least one person died. Moreover, it includes information on the type of restraint used by each vehicle occupant.

We did not have access to the program code used by Steven Levitt, nor did we have access to the specific dataset used in Levitt (2008). This replication is solely based on the descriptions and explanations in Levitt (2005) and Levitt (2008). After accessing the FARS data, we followed the description of how the author restricted the data as closely as possible.[23] Due to the number of restrictions imposed, we were unable to exactly replicate the working dataset. While our total sample has 38,456 observations, his has only 37,635.

As explained in Levitt (2008), the econometric approach employed regresses a binary indicator of whether a child died in a crash or not on the main variables of interest. The main variables of interest consist of the following set of dummies for restraint use: *(i)* no restraint, *(ii)* child safety seat, *(iii)* lap-only belt, and *(iv)* lap and

[23] Levitt (2008) writes that he drops crashes in which the only fatalities were pedestrians, motorcyclists, or occupants of nonstandard vehicles. Furthermore, he limits the analysis to occupants of automobiles, minivans and SUVs with model years older than 1969. Next, he discards observations with missing values on relevant variables and cases in which the occupant did not sit in the first three rows of the vehicle. Finally, he restricts the sample to children between the age of two and six.

shoulder belt. In addition, a rich set of vehicle and driver characteristics are used as controls. All models are linear probability models.

This simple regression intends to explain the statistical relationship between the type of restraint use and the probability that a child dies in a fatal car accident, controlling for observables. The accidents included in the FARS data, however, do not provide a random sample of American vehicles and occupants. Since restraint use may affect the probability of dying in a crash, and since both restraint use and accident fatalities may be related to a third, unobserved variable, the probability of being included in the FARS data is not independent of restraint use. Levitt (2008) adopts the Levitt and Porter (2001a) approach to correct for this sample selection issue. The simple idea is to restrict the sample to two-car crashes where an accident death occurs in the *second* vehicle involved. The sample selection issue is then resolved under the assumption that child restraint use in vehicle A does not affect the probability that an occupant dies in vehicle B, given both vehicles are involved in an accident.[24] Columns (1) through (4) of Table 3-1 below shows the exact replication of Table 3 in Levitt (2008). However, in the interest of space, we only show results from the fully controlled model and the specification without any controls (columns (1), (3), (4), and (6) in the Levitt paper).[25]

---

[24] Levitt (2008) points out that while the selection correction method employed remedies bias due to the dependence of restraint use and own-car fatality risk, it produces a sample of less severe crashes. Thus the results using the selection-bias corrected sample may not necessarily carry over to the universe of fatalities.

[25] The results for columns (2) and (5) are very similar and available upon request. The only difference to columns (3) and (6) is that they solely use a subset of control variables instead of the full set of controls.

Table 3-1. Replication and extension of Levitt (2008), Table 3 (Columns 1, 3, 4 and 6)—Impact of Child Restraints on Probability of Fatality

	Levitt replication: 1975-2003			
	Without Sample Selection Correction		With Sample Selection Correction	
	(1)	**(2)**	**(3)**	**(4)**
Child seat	-0.1168***	-0.1144***	-0.0462***	-0.0458***
	(0.0053)	(0.0068)	(0.0055)	(0.0076)
Lap and Shoulder Seat	-0.1046***	-0.1290***	-0.0470***	-0.0524***
	(0.0060)	(0.0072)	(0.0058)	(0.0080)
Lap-only belt	-0.1245***	-0.1080***	-0.0512***	-0.0480***
	(0.0061)	(0.0068)	(0.0058)	(0.0073)
Controls				
Position of child in car	No	Yes	No	Yes
Gender, age of child, driver belted	No	Yes	No	Yes
Car, model year, vehicle weight, type of crash	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
Other controls in Levitt (2008)	No	Yes	No	Yes
R^2	*0.0195*	*0.0810*	*0.0130*	*0.0496*
N	**38,456**	**38,456**	**10,330**	**10,330**

Note: * p<0.05, ** p<0.01, *** p<0.001; data from the Fatality Analysis Reporting System (FARS) for the years 1975-2011. Values in the table show the change in probability of dying in the crash associated with each restraint-type, relative to being unrestrained. Results in columns (1) and (2) are obtained from analyses using the sample of all 2 to 6 year-olds involved in a fatal crash; results in columns (3) and (4) are obtained from analyses using the sample of all 2 to 6 year-olds involved in 2-car fatal crash where someone died in the other car. "Other controls in Levitt (2008)" include the difference in weight of the cars, indicators for missing vehicle weight, of whether the driver had any major violations, of whether the speed limit on the road was less than or equal to 55 mpg, of whether the crash occurred on a rural road, or of whether the crash occurred on a weekend, at night (8pm to midnight), in the early morning (1am to 5am). The position of child in car variable indicates where the child was seated in the car relative to the back, middle. The child age categories are defined relative to 2 year-olds. All reported regressions are linear probability models. Standard errors are clustered at the vehicle level, and are reported in parentheses.

Columns (1) and (2) in Table 3-1 make use of the full sample without the sample selection correction, while columns (3) and (4) solely focus on the subset of two-car crashes with a death in the other car. Columns (1) and (3) regress the child death probability solely on the set of restraint use dummies, while columns (2) and (4) additionally control for a wide range of background information as indicated in the rows of Table 3-1. In all models employed, the reference category is *no restraint*, such that the estimates for each restraint type indicate the statistical death probability relative to being unrestrained. Given a child is unrestrained and involved in an accident with a fatality, the baseline probability that the child dies is 27%.

Although Levitt's (2008) full sample has 800 fewer observations than ours, the point estimates for all coefficients displayed in the first two columns of our Table 3-1 are almost identical to Levitt's (2008) Table 3 coefficients. For example, in column (1) of his table, Levitt (2008) reports a point estimate for *child seat* of -0.112, which is significant at the 1% level. Our equivalent estimate in column (1) of Table 3-1 shows a coefficient of -0.1144, also significant at the 1% level. Results of a formal test of the statistical difference between ours and Levitt's (2008) restraint use coefficients in the simple model with all covariates are reported in Table 3-2. The differences are negligible.

One can summarize the results displayed in the first two columns of Table 3-1 as follows: *(i)* Using a child safety seat, a lap and shoulder belt, or a just a lap belt reduces the likelihood that a child dies in a fatal accident by about 10 ppt. or 30% as compared to being unrestrained; *(ii)* controlling for a wide range of background information barely changes the point estimates; and *(iii)* the differences between our point estimates and Levitt's (2008) are very small and not statistically significant. Our replication results confirm Levitt's main finding: child safety seats provide no additional safety benefit as compared to shoulder and lap or lap-only belts.

Table 3-2. Formal t-test of differences in covariate estimates

	Levitt (2008)	Jones and Ziebarth (2013)	Difference	p-value
Simple				
Child seat	-0.132	-0.114	-0.018 (0.010)	0.0651
Lap belt	-0.108	-0.108	0 (0.010)	1.00
Lap/Shoulder belt	-0.132	-0.129	-0.003 (0.010)	0.7747
Selection-corrected				
Child seat	-0.054	-0.046	-0.008 (0.011)	0.4501
Lap belt	-0.046	-0.048	0.002 (0.010)	0.8414
Lap/Shoulder belt	-0.052	-0.052	0 (0.011)	1.00

The models in columns (3) and (4) of Table 3-1 only use a subset of observations to correct for sample selection and use about 2,000 fewer observations than Levitt's (2008) sample.[26] Consequently, the point estimates differ slightly, but the main findings are again very robust: sitting in a child safety seat or wearing a shoulder or lap or a lap-only belt reduces the probability that a child dies in a severe accident,

[26] Levitt (2005) writes that "for the sample selection correction, we created a dummy variable equal to one if someone died in another vehicle involved in the crash." This implies that all observations in the selection corrected sample should derive from crashes with at least two vehicles involved. However, in Levitt's (2008) results, a coefficient estimate is reported for the *one-car crash* variable. We followed Levitt (2005) in defining the selection corrected sample and therefore have no one-car crash victims included. The deviation in sizes between Levitt (2008) and our samples is likely due to the extra inclusion of one-car crash victims in Levitt's sample.

where an occupant dies in the other car, by about 5ppt. – about a 60% reduction.[27]

Table 3-3. Extension of Levitt (2008)

	Levitt extension: 1975-2011			
	Without Sample Selection Correction		Without Sample Selection Correction	
	(1)	(2)	(3)	(4)
Child seat	-0.1168***	-0.1168***	-0.0462***	-0.0473***
	(0.0053)	(0.0053)	(0.0055)	(0.0075)
Lap and Shoulder Seat	-0.1046***	-0.1046***	-0.0470***	-0.0524***
	(0.0060)	(0.0060)	(0.0058)	(0.0079)
Lap-only belt	-0.1245***	-0.1245***	-0.0512***	-0.0476***
	(0.0061)	(0.0061)	(0.0058)	(0.0073)
Child seat* Post2003	-0.0818***	-0.0818***	-0.0646**	-0.0439
	(0.0131)	(0.0131)	(0.0249)	(0.0237)
Lap and shoulder belt * Post2003	-0.0936***	-0.0936***	-0.0559*	-0.0347
	(0.0145)	(0.0145)	(0.0254)	(0.0243)
Lap belt * Post2003	-0.0495*	-0.0495*	-0.0550*	-0.0372
	(0.0193)	(0.0193)	(0.0268)	(0.0255)
Post2003	0.0662***	0.0662***	0.0569*	0.0293
	(0.0113)	(0.0113)	(0.0244)	(0.0451)
Controls				
Position of child in car	No	Yes	No	Yes
Gender, age of child, driver belted	No	Yes	No	Yes
Car, model year, vehicle weight, type of crash	No	Yes	No	Yes
Year Fixed Effects	No	Yes	No	Yes
Other controls in Levitt (2008)	No	Yes	No	Yes
R^2	*0.0248*	*0.0824*	*0.0165*	*0.0524*
N	**48,203**	**48,203**	**13,550**	**13,550**

Note: * p<0.05, ** p<0.01, *** p<0.001; See Table 3-1.

[27] As compared to columns (1) and (2), the point estimates slightly decrease in columns (3) and (4). However, since the mean fatality rate among unrestrained children are also smaller for this subsample (7%), the restraint-related safety increase in percent increases to 60%.

Extension of Levitt (2008): Do the Results Hold Up in the "Arms Race" Era 2004-2011?

The results in Table 3-3 extend Levitt's (2008) analysis by adding the years 2004 to 2011. The full sample has now 48,203 observations. We add a *post2003* dummy to the analysis and interact it with all restraint use variables of interest to identify whether restrain effectiveness has changed in the post-2003 period. One notes the following:

*(i)*The coefficient estimates indicate that in the modern era, relative to unrestrained children, children in safety belts and child seats appear even less likely to die, i.e., restraint use in general seems to have become more effective.

(ii) In the post-2003 period, correction for observables does matter. When controlling for a wide range of background characteristics, the coefficients significantly decrease in size. Still, the effects are significant and large, given that the mean fatality rate in the modern era for unrestrained children is about 33%: for a 2 to 6 year old child, both traditional seat belts and child safety seats reduce the probability of dying in a fatal accident by about 50% relative to being unrestrained.

(iii) The selection-corrected models in columns (3) and (4), with just 13,550 observations, lack statistical power when differentiating between the pre- and post-2003 time periods. However, when we partition the data and estimate models on the 2004 to 2011 selection corrected sample (results not reported here), we find that restraints reduce the likelihood of dying in a crash by about a 70%.[28]

[28] The coefficient estimates for the selection corrected models estimated on the 2004-2011 data only are -0.0828*** for *child seat* and -0.0759*** for *lap and shoulder belt*.

Finally, *(iv)* the main finding and conclusion of Levitt (2008) also holds in more recent years under changing traffic conditions: when it comes to preventing fatalities, child safety seats are *not* more effective than simple lap and shoulder restraints.

Restraint Use and Effectiveness in the Modern Era

Increased Effectiveness of Restraint Use: An Artifact of a More Negatively Selected Group of Parents Who Do Not Restraint Their Kids?

Figure 3-1 plots the rates of restraint use over time for children between 2 and 6 years, given they were involved in a fatal accident. The most striking observation is the strong, almost linear, decline in the share of children who are not restrained. In 1980 almost 100% of all 2 to 6 year olds in the sample were unrestrained. This proportion had only dropped to about 50% by the mid-1990s and to below 20% by 2010. It is obvious that, in the modern era, the group of children who remain unrestrained are a selective sample and the restraint use patterns are driven by a select group of parents or guardians.[29] If the marginal parent to take up restraint use is arguably less safety-conscious than the average restraint user, average driver quality among non-restraint users decreased over time.

[29] Among the selection corrected sample, only about 40% of children were unrestrained in the mid-1980s. The fraction of children who were unrestrained dropped to 20% in the mid-1990s, and to only about 5% by 2010.

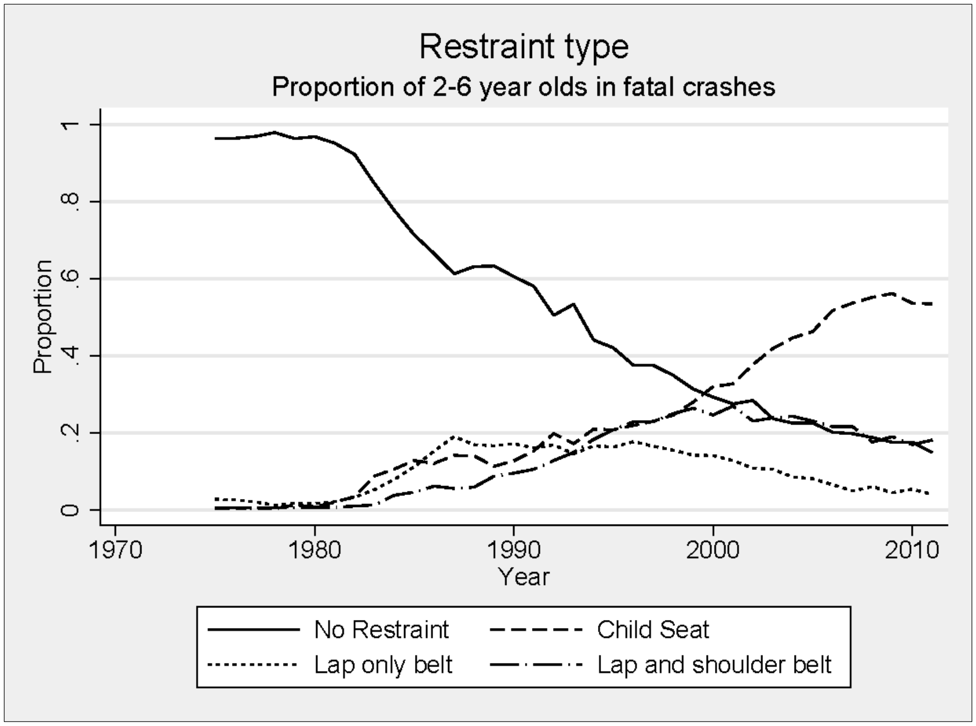

Figure 3-1. Development of restraint use (1975-2011)

Figure 3-2 provides evidence that the characteristics of drivers using different child restraint technologies is changing over time. For four characteristics, the figure plots the proportion of drivers using each restraint type with each characteristic for the child in their car.[30] The figure shows that, over time, drivers who do not restrain their child have become younger, more likely to have consumed alcohol or have a previous major violation, and more likely to be unrestrained themselves. This reinforces the hypothesis outlined above and suggests that the quality of drivers who do not restrain

[30] The characteristic rates among each restraint use category have been normalized to the characteristic rate among all drivers in the accident year, so that a value of 0 indicates no deviation from the average rate in the sample of all drivers. To normalize by trend in the characteristic prevalence in the entire sample, we calculate the proportion of all drivers involved in a fatal crash with a child who have the given characteristic in the given year. We then subtract this proportion from the proportion of drivers in sample using each restraint method that have the given characteristic.

their child has decreased over time, a result that may explain the finding of increased restraint effectiveness after 2003.

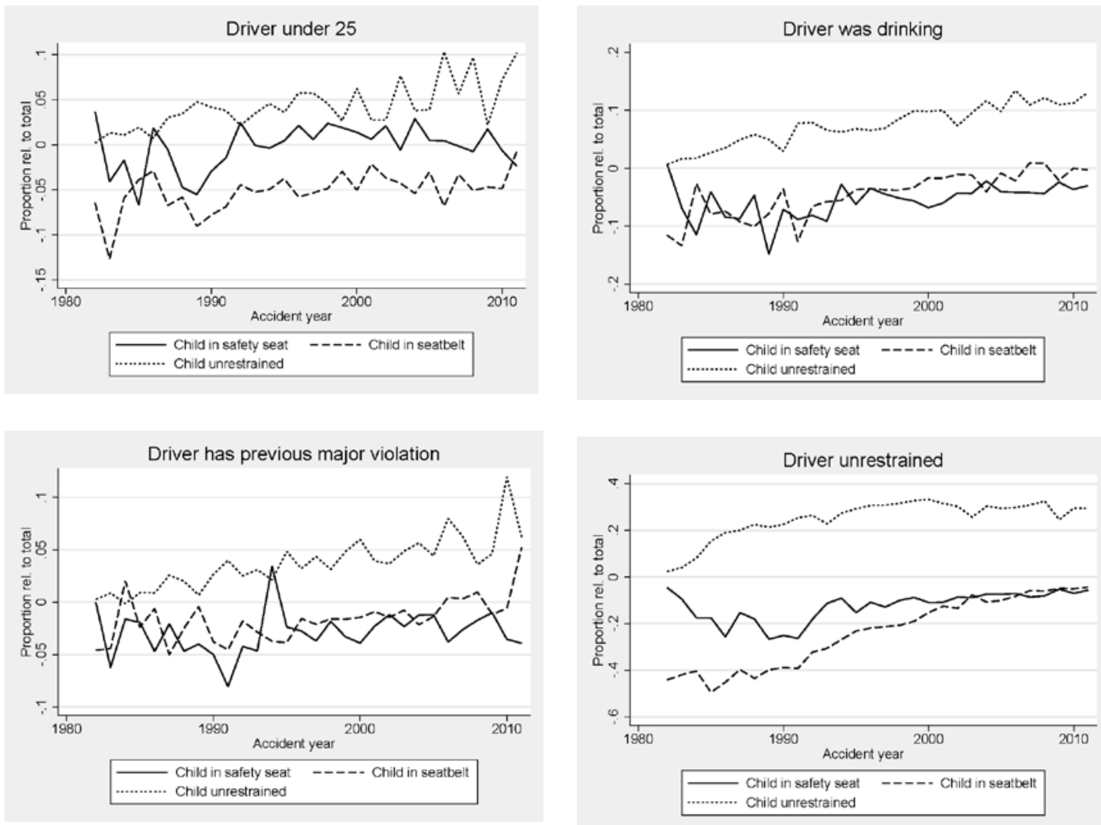

Figure 3-2. Driver characteristics by child restraint type, relative to overall driver characteristics, 1982-2011

Note that a more negatively selected sample of parents who do not restraint their kids poses *per se* no threat to the general empirical approach since the effectiveness of restraint use is always benchmarked against the unrestraint category, no matter how selective this reference sample is.

Restraint use development: Implications for child safety seat vs. seat belt effectiveness

Figure 3-1 shows that the usage of both restraint types strongly increased from 0% in 1980 to about 25% in 2000. Since then, however, traditional lap and shoulder belt use remained stable, even decreased slightly, while child safety seat use strictly increased to more than 60% in 2010. These divergent trends may reflect a selection story, which is supported by the fact that post-2003 era effectiveness estimates significantly decrease once controls for observables are added. One could assume that more "responsible" parents are better drivers, more likely to use child safety seats over seatbelts, and less likely to have fatal accidents.

Selection on unobservables between child safety seat and traditional belt users may introduce two potential sources of bias: First, child safety seat users would be less likely than seatbelt users to be included in the FARS fatality sample. However, to the extent that the probability of having a fatal accident is determined by the second car causing the accident in two-car crashes, this sample selection issue is taken care of by the Levitt and Porter (2001b) correction. Second, if the driving quality among the sample of child safety users improved relative to seatbelt users, the econometric approach *over*estimates the effectiveness of child safety seats relative to traditional belts. Positive selection into child seat use implies that child death would be less likely among the sample of child safety seat users relative to seat belt users, regardless of restraint type use.

Figure 3-2 helps dispel some of the concern about selection between seatbelt users and child safety seat users. With the exception of age, the driver characteristic trends among child safety seat users match the trends among seat belt users quite closely. Further, since positive selection into child safety seat use would result in an overestimate of their true effectiveness, such a selection pattern would not jeopardize the main finding. Finally, it is likely that much of the driver age variation in seatbelt use versus child seat use is due to variation in state-level laws regulating the age until which child safety seat use is mandatory. Indeed, the average age of children restrained in child seats is about 3, while those restrained by traditional seatbelts are about 4.5 years old. [31]

The Role of SUVs and Improper Restraint Use

Figure 3-3 investigates another recent development in road safety conditions: the share of traditional passenger cars versus SUVs on the road. Passenger car use in the sample declined from about 70% in 1980 to about 40% in 2010. Today, about 50% of the sample is riding in minivans or SUVs at the time of the crash. One observes a particularly strong increase in the use of SUVs since the year 2000—a near doubling from 17 to 31%.

[31] We tried partitioning the sample by age and re-estimating the models to determine whether child seat versus seatbelt effectiveness depends on child age. For children who are 2 or 3 at the time of the accident, seatbelts and child seats appear equally effective in preventing death; for 4, 5 and 6 year olds, our results (available upon request) suggest that seatbelts might be slightly more effective than child seats at preventing death.

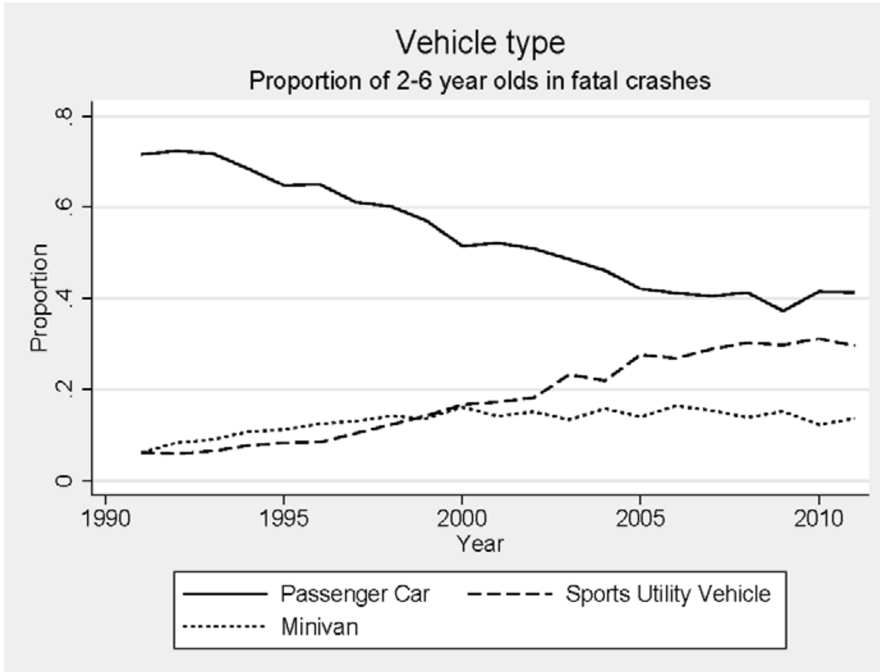

Figure 3-3. Development of vehicle type (1992-2011)

Table 3-4. Impact of child restraints, their improper use, and their use in SUVs on probability of fatality

	Dependent Variable=1 if Fatal Injury, 0 Otherwise			
	Without Sample Selection Correction		With Sample Selection Correction	
	(1)	(2)	(3)	(4)
Child seat*SUV	-0.0470**		-0.0226	
	(0.0150)		(0.0276)	
Seatbelt*SUV	-0.0162		-0.0218	
	(0.0157)		(0.0279)	
SUV	0.0105		0.0251	
	(0.0136)		(0.0275)	
Child seat improperly used		0.3014***		0.1447**
		(0.0247)		(0.0504)
Seatbelt improperly used		0.1522***		0.1119
		(0.0370)		(0.0726)
Child seat	-0.1310***	-0.1566***	-0.0594***	-0.0618***
	(0.0072)	(0.0077)	(0.0100)	(0.0099)
Seatbelt	-0.1441***	-0.1538***	-0.0632***	-0.0594***
	(0.0066)	(0.0071)	(0.0088)	(0.0095)
Controls				
Position of child in car	Yes	Yes	Yes	Yes
Gender, age of driver, driver belted	Yes	Yes	Yes	Yes
Car, model year., vehicle weight, type of crash	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Other controls in Levitt (2008)	Yes	Yes	Yes	Yes
R^2	0.0861	0.0980	0.0347	0.0449
N	**33,140**	**25,622**	**10,497**	**8,264**

Note: * p<0.05, ** p<0.01, *** p<0.001; data from the Fatality Analysis Reporting System (FARS) for the years 1991-2011 for columns (1) and (3), and for the years 1994-2007 for columns (2) and (4). Values in the table show the change in probability of dying in the crash associated with each restraint-type, relative to being unrestrained. Results in columns (1) and (2) are obtained from analyses using the sample of all 2 to 6 year-olds involved in a fatal crash; results in columns (3) and (4) are obtained from analyses using the sample of all 2 to 6 year-olds involved in 2-car fatal crash where someone died in the other car. "Other controls in Levitt (2008)" include the difference in weight of the cars, indicators for missing vehicle weight, of whether the driver had any major violations, of whether the speed limit on the road was less than or equal to 55 mpg, of whether the crash occurred on a rural road, or of whether the crash occurred on a weekend, at night (8pm to midnight), in the early morning (1am to 5am). The position of child in car variable indicates where the child was seated in the car relative to the back, middle. The child age categories are defined relative to 2 year-olds. All reported regressions are linear probability models. Standard errors are clustered at the vehicle level, and are reported in parentheses.

Table 3-4 formally investigates how the interplay between SUV and restraint use affects safety in the 1991 to 2011 period. Interestingly, we do not find evidence that SUV use alone reduces the probability that a child dies in a crash—in general and in the selection correction approach in column (3). However, in the naïve model in column (2), there is some evidence that child safety seat use *in combination* with SUV use reduces child fatalities by 5ppt. or 25%, while seatbelt use does not appear additionally effective in SUVs.[32] However, once selection into the sample of fatalities is corrected for, the child seat safety premium associated with SUV use disappears. In column (3), the coefficient estimates on both the seat belt and child seat interaction terms are small in magnitude, almost identical, and not significantly different from 0. Thus, overall, there is no evidence that SUVs prevent fatalities better than other cars, even with the use of restraints.

Lastly, we make use of an explanatory factor that was included in the survey between 1994 and 2007: *improper* child seat and seat belt use. [33] Column (2) of Table 3-4 illustrates that the safety gains from using lap and shoulder belts are completely offset by their improper use, such that improper use is as dangerous as no use. Strikingly, improper child safety seat use appears significantly *less* safe than no restraint. While child safety seat use is associated with a 15 ppt. decrease in the probability of death, the effect of improper use completely overwhelms the safety gain, resulting in a net *increase* in death probability of about 14 ppt. On a base fatality

[32] For all results presented in Table 4, we collapse the lap and shoulder belt category and the lap-only belt category due to the relatively few children using lap-only belts in more recent years.
[33] Thus, the models in column (2) and (4) only make use of the years 1994 to 2007 and have only 25,622 and 8,264 observations, respectively.

rate of about 20%, this amounts to about a 75% increase in the likelihood of death associated with improper use of child seats.

Of course, the question of selection bias is again an important one in this discussion. There is reason to believe that parents who use restraints improperly may differ significantly from correct users on dimensions that affect probability of crash and crash severity. The results, however, persist with the inclusion of covariates and when the model is estimated using the selection corrected sample. Indeed, the effect of improper use continues to subsume and overpower the safety benefits of child seat use in the selection corrected model (column (4)). On a lower base fatality probability of about 4% in the selection corrected sample, the net effect of improper child seat use nearly triples the risk of death relative to no restraint.

Discussion and Conclusion

In this paper, we replicate the results in Levitt (2008) nearly perfectly. According to these findings, child safety seats provide no additional safety advantage over traditional lap and should seat belts.

We additionally extend Levitt's (2008) analysis and show that the results also hold in the new millennium despite some remarkable developments on Americans roads. For example, child seat safety use has strictly increased while the prevalence of unrestrained children has strictly decreased. We thus provide a careful analysis of changing driver characteristics among differently restrained children which may account for conflicting findings in past studies of restraint effectiveness (e.g. Elliot et al. (2006)).

Further analyses show that the SUV safety premium disappears once selection into a crash is accounted for. This indicates that the SUV safety premium is due to selection rather than true differences in safety, which is an important finding, given the commonly-held belief that SUVs are safer. For example, a 2005 National Highway Traffic Safety Administration report using the same FARS data found that properly restrained children in SUVs are significantly less likely to die in a fatal crash (Starnes, 2005). It is also important from a welfare perspective since passengers in cars involved in a crash with an SUV are significantly more likely to die. If SUVs do not provide additional safety benefits to occupants, and endanger passengers in other car types, their increasing prevalence on American roads is cause for concern (White, 2004; Anderson, 2008; Li, 2012).

Finally, our analysis shows that improper use of child safety seats provides less protection from child deaths than either traditional lap and should belt use or unrestraint. This effect persists even once selection is accounted for. Given the prevalence of improper use—the Children's Safety Network (2005) estimates that more than half of all child seats are improperly used—and the lack of evidence for their effectiveness beyond seatbelts, laws enforcing their use might be welfare reducing.

REFERENCES

Anderson, Michael (2008): “Safety for whom? The effects of light trucks on traffic fatalities,” *Journal of Health Economics*, 27(4: 973-989.

Bull, Marilyn J. and Stroup, Karen Bruner and Gerhart, Susan (1988): “Misuse of Car Safety Seats,” *Pediatrics*, 81(1): 98-101.

Busse, Meghan R. and Christopher R. Knittel and Zettelmeyer, Florian (2013): “Are Consumers Myopic? Evidence from New and Used Car Purchases,” *American Economic Review*, 103(1): 220-256.

Children’s Safety Network (2005): “Child Safety Seats: How Large are the Benefits?” available at http://www.childrenssafetynetwork.org/sites/childrenssafetynetwork.org/files/child_safety_seats_childhood_injury_cost_prevention.pdf, last accessed on October 7, 2013.

Daly, Lauren, Kallan, Michael J., Arbogast, Kristy B. and Durbin, Dennis R. (2006): “Risk of Injury to Child Passengers in Sport Utility Vehicles,” *Pediatrics*, 117(1): 9-14.

Elliot, Michael R., Kallan, Michael J., Dennis R. Durbin and Winston, Flaura K. (2006): “Effectiveness of Child Safety Seats vs Seat Belts in Reducing Risk for Death in Children in Passenger Vehicle Crashes,” *Pediatrics*, 160(6): 617-621.
Howland, Willard J., Curry, Joseph L. and Buffington, Caroll B. (1965): “Fulcrum fractures of the lumbar spine: Transverse fracture induced by an improperly placed seat belt,” *JAMA*, 193(3): 240-241.

Insurance Institute for Highway Safety (IIHS) (2013): “State Laws”. Available at http://www.iihs.org/iihs/topics/laws/safetybeltuse?topicName=child-safety, last accessed on Nov. 10, 2013.

Klier, Thomas and Linn, Joshua (2012):”New-vehicle characteristics and the cost of the Corporate Average Fuel Economy standard,” *RAND Journal of Economics*, 43(1): 186-213.

Levitt, Steven D. (2005): “Evidence that Seat Belts are as Effective as Child Safety Seats in Preventing Death for Children Aged Two and Up,” *mimeo*, Data Appendix, available at: http://pricetheory.uchicago.edu/levitt/Papers/levitt_carseats_farsdata.pdf, last accessed on September 18, 2013.

Levitt, Steven D. (2008): “Evidence that Seat Belts are as Effective as Child Safety Seats in Preventing Death for Children Aged Two and Up,” *The Review of Economics and Statistics*, 90(1): 158-163.

Levitt, Steven D. and Doyle Junior, Joseph J. (2010): "Evaluating the Effectiveness of Child Safety Seats and Seat Belts in Protecting Children from Injury," *Economic Inquiry*, 48(3): 521-536.

Levitt, Steven D. and Porter, Jack (2001a): "Estimating the Effect of Alcohol on Driver Risk Using only Fatal Accident Statistics," *The Review of Economics and Statistics*, 83(4): 603-615.

Levitt, Steven D. and Porter, Jack (2001b): "How Dangerous are Drinking Drivers," *Journal of Political Economy*, 109(6): 1198-1237.

Li, Shanjun (2012): "Traffic safety and vehicle choice: quantifying the effects of the 'arms race' on American roads," *Journal of Applied Econometrics*, 27(1): 34-62.

Starnes, Marc (2005): "Child Passenger Fatalities and Injuries, Based on Restraint Use, Vehicle Type, Seat Position, and Number of Vehicles in the Crash," Mathematical Analysis Division, National Center for Statistics and Analysis, National Highway Traffic Safety Administration, available at http://www-nrd.nhtsa.dot.gov/Pubs/809784.pdf, last accessed on Oct. 7, 2013.

Small, Kenneth A. and Van Dender, Kurt (2007): "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect," *The Energy Journal*, 28(1): 25-52.

White, Michelle J (2004): """The Arms Race" on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety," *Journal of Law and Economics*, 47(2): 333-355.

World Health Organization (2013): "Global Status Report on Road Safety". Available at http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/index.html, last accessed on Nov. 10, 2014.

APPENDIX 1

Childhood Sub-County Imputation Details

First, I know whether a respondent has ever moved. For those who have not (N=10, 350), I set birth sub-county as current sub-county. I further know whether respondents live on-reserve. Since there is very little mobility onto reserves, I assume that among the remaining respondents, those respondents who currently live on-reserve (N=14, 960) were likely born on-reserve and I set their current sub-county of residence as their sub-county of birth. For the remaining 9,000 respondents, I predict the most likely sub-county of birth given birth province and band membership.

1991 Aboriginal Peoples Survey - Relevant Questions

Residential School Attendance

Who did you live with while you were going to [elementary/high] school? Was it … with your family? at a residential school? with a non-Aboriginal family? with an Aboriginal family? with someone else?

Culture

Do you speak and Aboriginal language well enough to carry on a conversation? Did you ever speak an Aboriginal language?

How much of the meat, fish, poultry which you eat is obtained through hunting and fishing by you, members of your family, or friends?

Census question: What is this person's religion? Specify one religion or denomination only.

Health

Have you been told by a health care professional that you have … diabetes? high blood pressure? arthritis or rheumatism? heart problems? bronchitis? emphysema or shortness of breath? asthma? tuberculosis, that is, T.B.? epilepsy or seizures?

In the past twelve months, how often on average did you drink beer, wine, liquor, or home brew?

Do you now smoke cigarettes … daily? occasionally? not at all?

How tall are you when you are not wearing shoes? How much do you weigh?

Social

Census question: Legal marital status … legally married (and not separated)? legally married and seperated? divorced? widowed? never married?

How many liveborn babies have you had?

In your opinion, are any of the following a problem for Aboriginal people in the community or neighbourhood where you are living now … suicide?

Table A1-1. Results of models including linear and quadratic cohort trends

	North American Indians				
	Panel A				
		Catholic	Hunts	Culture	
Exposure Years					
County-level linear cohort trends		0.0027**	-0.0020**	0.0050**	
		(0.0007)	(0.0006)	(0.0011)	
County-level quad. cohort trends		0.0028**	-0.0024**	0.0066**	
		(0.0009)	(0.0007)	(0.0013)	
Demographic controls		X	X	X	
N		33,830	31,020	31,020	
	Panel B				
	Married	Babies	Suicide	Drinks	Smokes
Exposure Years					
County-level linear cohort trends	-0.0033**	-0.0127**	0.0042**	0.0015**	0.0026**
	(0.0008)	(0.0037)	(0.0008)	(0.0004)	(0.0006)
County-level quad. cohort trends	-0.0025**	-0.0081**	0.0051**	0.0020**	0.0022**
	(0.0007)	(0.0041)	(0.0009)	(0.0005)	(0.0006)
Demographic controls	X	X	X	X	X
N	33,830	17,340	25,100	33,480	33,480

Data are the 1991 Aboriginal Peoples Survey. Models estimated using responses from those who identify as Native, who indicate North American Indian status, and who were born between 1942 and 1971. Dependent variables in Panel A are an indicator for having graduated high school, for receipt of welfare, for Catholic religion, for obtaining at least half of meat from hunting; Culture is a combined measure of Catholicism, Hunting and Speaking Ab that ranges from 0 to 3, where someone who is Catholic, not a hunter and does not speak and Aboriginal language receives a score of 3. Dependent variables in Panel B are indicators for being married, worrying that suicide is a problem in the community, for drinking at least weekly and for smoking daily; Babies is the number of liveborn babies among women. Estimated models are linear probability. Demographic controls include gender, an indicator for Official Indian Status, an indicator for multiple ethnic origin, and indicator for non-Canadian birth, and 3 indicators for geographic region: North, mid-North or South. Reported coefficient estimates derive from models that include the stated county-level years-of-birth trends. Standard errors clustered at the county-cohort level reported in brackets. *p <0.05; **p <0.01

APPENDIX 2

Sample and Attrition

Base Sample: Children who were between the ages of 0 and 9 in Cycle 1 (1994). These children were between the ages of 14 and 23 in Cycle 8 (2008). Table 1 of the appendix shows the number of children in our base sample surveyed in each cycle of data collection. Cycle-to-Cycle loss of respondents is due to attrition, with the exception of the large decline in sample size after the initial year of data collection; the sample size was purposefully reduced after Cycle 1 due to budgetary restrictions.

Table A2-1. Number of children surveyed in each Cycle of data collection

Cycle	Number of children
1	19,397
2	13,189
3	12,793
4	11,321
5	10,753
6	9,848
7	9,581
8	8,861

Children who stayed in the survey sample until Cycle 8 – whom we call "stayers" – did not exhibit different ADHD symptoms than attriters, as measured by the ADHD screener questions in Cycle 1 of data collection. However, attriters were more likely to report being on Ritalin in Cycle 1 than stayers. Attriters were also more likely than stayers to be male, to come from lower income households, to come from single parent homes and to have mothers with a high school education or less. Appendix Table 2 compares the number of observations we would have in the case of no attrition with the actual numbers, for each outcome.

Table A2-2: Effects of attrition for each outcome

Outcome	Expected number of Observations	Actual number of observations
Medium-term outcomes		
On stimulants (age 2-15)	72,084	55,239
Repeat Grade (age 4-15)	68,278	44,968
Unhappiness Score (age 2-11)	44,858	36,458
Relationship with Parent (age 4-9)	27,379	22,554
Math Score (age 5-15)	64,788	32,515
Long-term outcomes		
Ever took stimulants	9,747	8,643
Depression Score	9,747	6,493
Completed High School	6,819	4,676
Some Post-Secondary	6,819	4,676

Variable Construction

Mental Health Variables

The mental health score variables are all constructed from questions that ask the respondent to rate the frequency of certain behaviors on a scale from 0 to 2. Scores are constructed by summing the frequency values for appropriate questions. Higher scores imply more severe behavior. The section below indicates which questions were combined to create each behavior score.

1. Short-term Hyperactivity Score:

a) HOW OFTEN WOULD YOU SAY THAT %FNAME%: Can't sit still, is restless or hyperactive?

*b)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Is distractible, has trouble sticking to any activity?

*c)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Can't concentrate, can't pay attention for long?

*d)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Cannot settle to anything for more than a few moments?

e) HOW OFTEN WOULD YOU SAY THAT %FNAME%: Is inattentive?

2. **Short-term Anxiety and Depression Score:**
*a)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Seems to be unhappy, sad or depressed?**

*b)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Is not as happy as other

*c)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Is worried?

*d)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Is nervous, high-strung or tense?

*e)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Has trouble enjoying him/herself?**

f) HOW OFTEN WOULD YOU SAY THAT %FNAME%: Is too fearful or anxious?

** Questions marked with asterisks were used to construct the depression score, while non-marked questions were used to construct the anxiety score.

3. Short-term Physical Aggression Score:

*a)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Gets into many fights?
*b)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: When another child accidentally hurts %him/her% (such as by bumping into %him/her%), assumes that the other child meant to do it, and then reacts with anger and fighting?
*c)*HOW OFTEN WOULD YOU SAY THAT %FNAME%: Kicks, bites, hits other children?

4. Self-assessed over-16 Anxiety and Depression Score:
How often have you felt or behaved this way during the past week (7 days)?
(a) I did not feel like eating; my appetite was poor.
(b) I felt I could not shake off the blues even with help from my family or friends.
(c) I had trouble keeping my mind on what I was doing.
(d) I felt depressed.
(e) I felt that everything I did was an effort.
(f) I felt hopeful about the future. **
(g) My sleep was restless.
(h) I was happy. **
(i) I felt lonely.

(j) I enjoyed life. **
(k) I had crying spells
(l) I felt people disliked me.

** 0 to 2 scales for the marked questions were reversed when calculating the overall score.

5. Ever Diagnosed with a Mental or Psychological Disorder: We construct this indicator from a question asked of all youths age 16 and over: "Has a health professional ever diagnosed you with an emotional, psychological or nervous disorder?" Youths who indicated ever having a diagnosed disorder were given a 1 for this indicator variable.

EDUCATION VARIABLES

1. Standardized Math Score: The mathematics test was administered in school to children in grade 2 or higher and was composed of 15questions drawn from the Canadian Achievement Test (CAT2). The difficulty of the questions increased as the child advanced in school, meaning that the age-specific average score did not differ substantially from the overall average score. We therefore standardized the score irrespective of age.

2. Repeated Grade: Parents of all children up to age 15 were asked whether the child had repeated a grade since the previous interview 2 years prior. We used the answers to these questions to create an indicator that equals 1 if the child *has not* repeated a grade in the previous 2 years.

3. Age-15 Standardized Math Score: Using the same mathematics test score that we employ in the short-term analysis, we identify the final math score recorded for each child, which is recorded at age 15.

4. Completed High School, Some Post-Secondary: These variables are constructed from the NLSCY education status variables. We begin with Cycle 8 data and observe whether the youth has graduated high school, begun post-secondary education or completed post-secondary education. High school graduates and those pursuing or having completed post-secondary education receive a 1 for the High School Graduation variable, while those who indicate not having completed high school receive a 0. Similarly, those pursuing or having completed post-secondary education get a 1 for the Some Post-Secondary indicator, while high school drop outs and high school graduates who did not continue their education receive a 0.

If the Cycle 8 education status variable is missing, we look back to the most recent Cycle of data collection with a non-missing education status variable. We assign missing values for both indicator variables for youths who, at last contact, were still in high school or whose education status is unknown and have never reported completing high school.

Quebec Insurance Program Detail
Costs of the Basic Public Plan for people 18-64 not covered by Private Insurance

Table A2-3: Quebec insurance rates

Year	Yearly Premium	Co-Insurance Rate %	Monthly Deductible	Maximum Yearly Out of Pocket Contribution
1997	175	25	8.33	750
2002	422	27.4	9.13	822
2003	460	28	9.6	839
2004	494	28.5	10.25	857
2005	521	28.5	11.90	857
2006	538	29	12.10	857

Children up to age 17 of people insurance under the public plan are eligible for free prescription medication.

Source: Gouvernment du Quebec, 2007

Table A2-4: Robustness Checks

Robustness Check	Ventilator Use	Kids w/o other chronic	Kids born 1985-1991	Composite Mood Score	
	DDD	DDD	DDD	Boys - DDD	Girls - DDD
Dependent Variable	(1)	(2)	(3)	(4)	(5)
On Ventilator	0.001	-	-	-	-
Ages 0-15	(0.001)				
N	55,239				
On Ritalin	-	0.007**	0.002	-	-
Ages 0-15		(0.003)	(0.002)		
N		38,314	28,819		
Unhappiness Score	-	0.027	-0.027	-	-
Ages 2-11		(0.017)	(0.019)		
N		26,011	16,765		
Relationship With Parent	-	-0.021**	0.016	-	-
Ages 4-9		(0.005)	(0.016)		
N		16,192	16,765		
No repeated grade	-	-0.008**	-0.011**	-	-
Ages 4-16		(0.002)	(0.002)		
N		30,368	27,158		
Math Score	-	-0.034**	-0.037**	-	-
Ages 4-16		(0.014)	(0.007)		
N		21,692	19,238		
Mood Score	-	-	-	-0.0126	0.0353**
Age 2-11				(0.0228)	(0.0156)
N				18,484	17,974

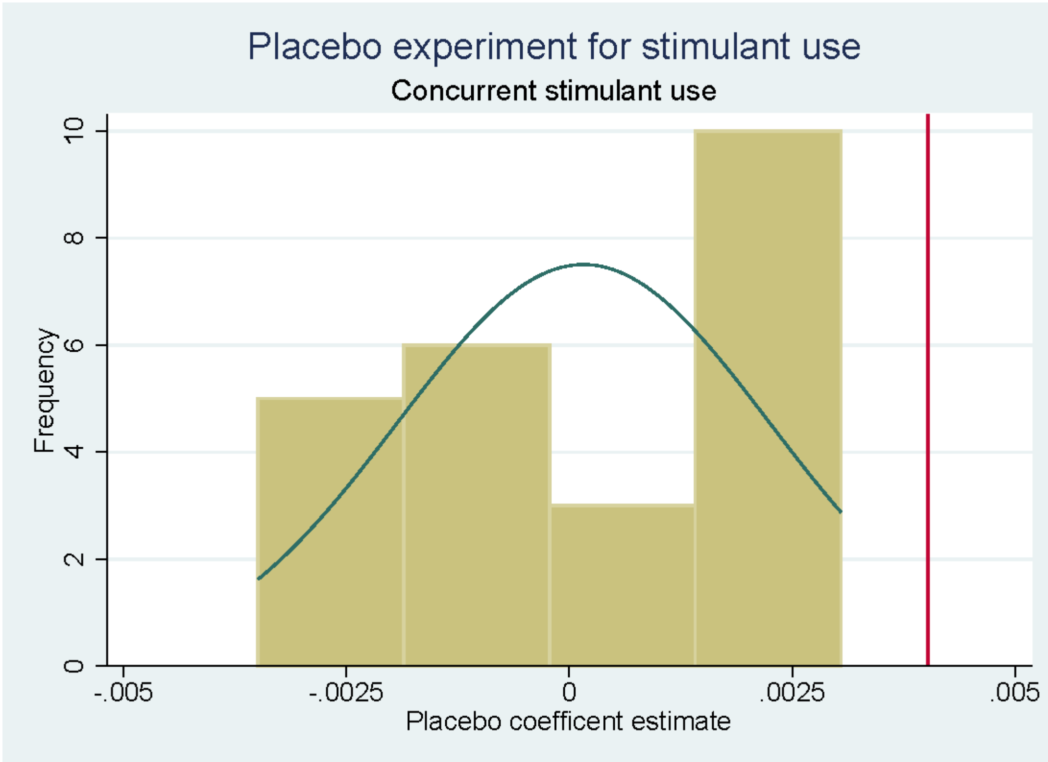

Figure A2-1: Placebo test results for contemporaneous stimulant use

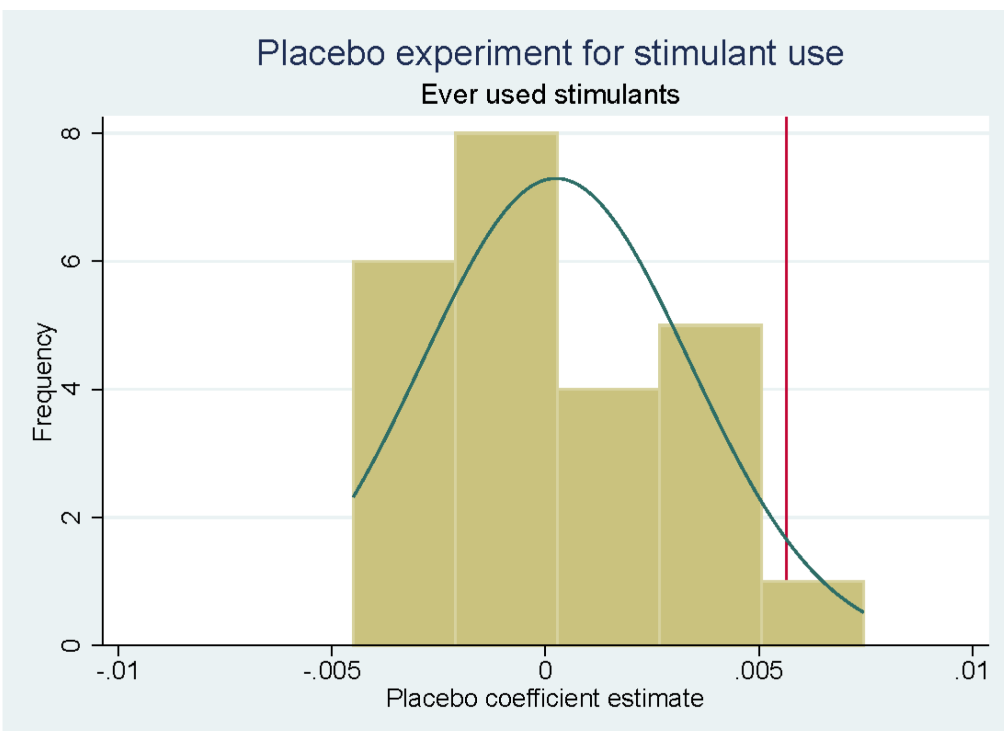

Figure A2-2: Placebo test results for ever used stimulants