



uOttawa

**The Impact of Exposure to Ambient Air Pollution on  
Educational Outcomes**

by

Yanjiao Zhang

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies  
in partial fulfillment of the requirements for the degree of

Master of Science in Environmental Sustainability

Supervisor

Dr. Nicholas Rivers

Institute of the Environment

University of Ottawa

© Yanjiao Zhang, Ottawa, Canada, 2017

## ABSTRACT

In this study, I examine if higher levels of ambient air pollution impact educational outcomes. According to the literature review, CO, O<sub>3</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> are found to be four pollutants that could have an impact on cognitive ability, so I focus on these air pollutants. I analyze provincial test results for the province of British Columbia, and Secondary School Literacy Test results from the province of Ontario (OSSLT) with air pollution and weather data corresponding to the locations and dates in which tests took place. A longitudinal approach is used, in which test results are compared within a school over time with a fixed effects model chosen to control for school and year fixed effects. Correlations are found among the four pollutants in the two provinces, therefore, an integrated Air Quality Index (AQI) is calculated to further examine the relationship between air pollution and educational outcomes. In British Columbia, I find that there is a negative impact of ambient air pollution on student's test results: a one standard deviation increase in AQI leads to a 0.23 percentage points decrease in student average grade. Furthermore, I find that in BC, air pollution's impact on students with special needs experience about 3.4 times of the average impact of other students. In Ontario, I do not find significant association between OSSLT results and the AQI, and this might be because of not having enough observations in Ontario school dataset and lots of missing data in air pollution dataset. However, the association between ambient air pollution and OSSLT results is found to be negative overall, congruent with results from BC.

## ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Nicholas Rivers, who has been a tremendous mentor for me. I would like to thank you for giving me continuous, patient, motivated support in learning of R studio, Econometrics and broaden my ways of thinking throughout my thesis research. I cannot think of a better supervisor to have. I also would like to thank my committee members, professors Scott Findlay and Frances Pick for serving as my committee members.

I would like to express my sincere appreciation to DataBC and BC Ministry of Education for providing the BC school data used in this study, especially Virginia Rego, Michael Korican and Morag Masterton who have been very helpful in providing and explaining the detailed information for the data. I would like to thank the Education Statistics & Analysis Branch as well as Education Quality and Accountability Office of Ontario, for providing the OSSLT data used in this research. The opinions presented in this paper are solely those of the author and do not necessarily reflect the opinions of the above departments.

I would like to express my thanks of gratitude to Institute of the Environment for giving me great help in adjusting to the new study in Canada.

Last but not least, I am really grateful to my husband Zizhao Jiang, who has been supportive and inspired all the time, and my son William, who gave me the big passion for new life and study, and my parents in law, who helped us taking care of William. In the end, I would like to express my deepest appreciation to my father Shijie Zhang, and my mother Xiucong Wang who loves me, believes in me and supports me all the time.

## TABLE OF CONTENTS

CHAPTER 1	INTRODUCTION .....	1
1.1	Air Pollution and Human Health .....	1
1.2	Cognitive Impacts of Air Pollution .....	2
1.3	Research in This Study .....	4
CHAPTER 2	LITERATURE REVIEW .....	9
2.1	Air Pollution Effects on Productivity: Pathological and Non-Pathological.....	9
2.2	Findings of Air Pollution’s Impact on Student’s Educational Performance.....	11
2.3	Findings of Air Pollution’s Impact on Employee Productivity .....	12
2.4	Compare and Conclusion of These Findings.....	13
CHAPTER 3	DATA.....	15
3.1	Panel Data .....	15
3.2	Data Sets used in this study .....	16
3.3	Processing of the Data Sets .....	19
3.4	Visual Overview of the Dataset .....	30
CHAPTER 4	METHODOLOGY.....	38
4.1	Introduction of Empirical Strategy.....	38
4.2	Fixed-effects model (with air pollutants as independent variables) .....	38
4.3	Air Quality Index.....	41
4.4	Logarithmic Transformation .....	43
4.5	Non-linear Relationship Exploration .....	44
CHAPTER 5	RESULTS AND DISCUSSIONS .....	45
5.1	Fixed-Effects Model With Pollutants as Independent Variables .....	45
5.2	Air Quality Index.....	48
5.3	Logarithmic Transformation .....	51
5.4	Non-linear relationship Exploration .....	52
5.5	Robustness Check .....	53
5.6	Heterogeneity.....	56
CHAPTER 6	CONCLUSIONS .....	58
BIBLIOGRAPHY	.....	61

## CHAPTER 1 INTRODUCTION

Over 80% of the global population lives in urban areas. The air pollution level in urban areas has increased by 8% from 2008 to 2013<sup>46</sup>. It is increasingly affecting human population in both visible and less visible ways.

### 1.1 Air Pollution and Human Health

Illness or disease is one of the main visible threats that air pollution poses for humans. Particulate matter, which is a mixture of solid and liquid particles in the air including pollutants such as sulfate, nitrates and black carbon etc., is one of the important components of air pollution. It reduces pulmonary and cardiovascular function, and causes diseases of asthma, heart attacks as well as problems of blood pressure changes, irritation in the ear, and headaches.<sup>1, 13, 33, 39</sup> Gases such as carbon monoxide (CO), which is odorless and colorless, restrains the oxygen in blood to be delivered to organs and tissues by binding to the iron in hemoglobin.<sup>34</sup> In addition, medical research has also found CO poisoning can trigger symptoms including headache, dizziness and confusion, which could make people performed at a lower efficiency or decide not to work or study.<sup>32</sup> Ozone(O<sub>3</sub>) is not emitted directly into the air, but forms from complex chemical reactions between nitrogen oxides (NO<sub>x</sub>) and volatile organic chemicals (VOCs) in the presence of heat and sunlight.<sup>17</sup> O<sub>3</sub> can lead to various health problems, such as shortness of breath, painful deep breath, cough, chronic obstructive pulmonary disease. People with lung diseases are more vulnerable to O<sub>3</sub>.

There are some statistical results which provide information on the cost of air pollution's impact on human health in Canada. The Ontario Medical Association (OMA) reported in 2005 that the pollution of ground-level ozone and particulate matter led to \$150 million health care cost in Toronto (not including visiting family doctors) and \$128 million productivity lost (people were too sick to go to work).<sup>28</sup> Moreover, the Canadian Medical Association (CMA) released that cardiopathy and respiratory disease caused by air pollution in 2008 resulted in 11,000 hospital admissions and 92,000 emergency department visits in Canada. The estimation of nationwide economic lost attributed to air pollution in 2008 was \$8 billion including health care cost, productivity lost.<sup>5</sup>

However, as demonstrated by some recent studies, air pollution could also affect people's cognitive performance, results in a potential decrease of educational outcomes or work efficiency.<sup>12, 14, 19, 34, 41, 42</sup> The economic damage of ambient air pollution might be underestimated if we do not include the cost of this impact.

## **1.2 Cognitive Impacts of Air Pollution**

With more and more studies done and published, the health impacts of air pollution are generally known by public. However, besides the obvious impacts on health, there are many less visible impacts that air pollution could have on humans and that are more ignored.

Cognitive ability is one of them, which is essential to student's performance and employee's productivity; as such, reduced cognitive ability can lead to huge loss to the whole economy.

In the pathology field, it is found that air pollution can lead to neuroinflammation which can result in neurodegenerative changes including diminished cognitive ability.<sup>37</sup> Some recent studies have found a close association between air pollution and student's performance as well as worker's productivity. It is found that ozone, CO, PM10 and PM2.5 have negative relationships with student's performance tests and even previous fetal exposure to air pollution may also impact test outcomes.<sup>14, 19, 34</sup> Also, students exposed to higher levels of pollution during test are found to have fewer years of post-secondary education and lower payment in adulthood.<sup>19</sup> Moreover, air pollution is found to have negative relationship with the productivity of workers, for example pear-packers, box-packers.<sup>12, 41, 42</sup>

Ignoring the above impacts could result in the underestimate of the cost of air pollution. The cost-benefit analyses that are currently conducted are missing some costs of air pollution or some benefits of air pollution reduction (productivity, education). For example, in the assessment of "Regulations Amending the Vessel Pollution and Dangerous Chemicals Regulations" in 2012, it is estimated that, by 2020, sulphur oxides emission would be reduced by 96%, nitrogen oxides emission reduced by 80%, and particulate matter reduced by 80%. In the cost-benefit analysis of the assessment, it is estimated that the total benefit derived from the reduction of the pollutants consists of a health benefit of \$9.81 billion, a GHG benefit of \$0.05 billion, as well as a non-monetary benefit. The non-monetary benefit is considered the pollution reduction in freshwater bodies and the improved health condition of vegetation and aquatic wildlife. However, the benefits derived from the potential increase

in student educational outcomes and employee productivity are not included, therefore the benefit of this regulation could be underestimated.\* Including these would lead to more stringent regulations.

The pollution level varies in different countries/regions and this difference could be very large. For example, Air Quality Index (AQI) varies from below 10 to above 300 in different regions of the world.† Even when the AQIs of different countries/regions are the same or very close, the kinds of pollutants and corresponding concentrations could be very different, since AQI is a comprehensive index integrated from several air pollutants and reflects the most detrimental impact among the pollutants. Hence, the association between air pollution and educational outcomes as well as productivity in one country/region may not apply to another country/region.

### **1.3 Research in This Study**

In this study, I examine if there is a detectable association between air pollution and student's test results, and if the association can be taken as a causal relationship.

Four pollutants are examined in this study: particulate matter less than 2.5 microns in diameter (PM<sub>2.5</sub>), particulate matter less than 10 microns in diameter (PM<sub>10</sub>), Carbon

---

\* Canada Gazette, Regulations Amending the Vessel Pollution and Dangerous Chemicals Regulations.

<http://www.gazette.gc.ca/rp-pr/p2/2013/2013-05-08/html/sor-dors68-eng.html>

The health benefit consists of benefits deriving from the reduction of mortalities, hospital admissions, emergency room visits, adult chronic bronchitis cases, child acute bronchitis episodes, asthma symptom days, minor restricted activity days, restricted activity days as well as acute respiratory symptom days.

The carbon price used in calculation of the GHG benefit was estimated based on the social cost of carbon estimated by published researches, the carbon price on international carbon markets as well as target prices announced by key jurisdictions.

† Real-time World Air Quality Index. <https://waqi.info/>

Monoxide (CO) and Ozone (O<sub>3</sub>). The following figure shows the trends of the four pollutants in BC and Ontario from 1980 to 2015.

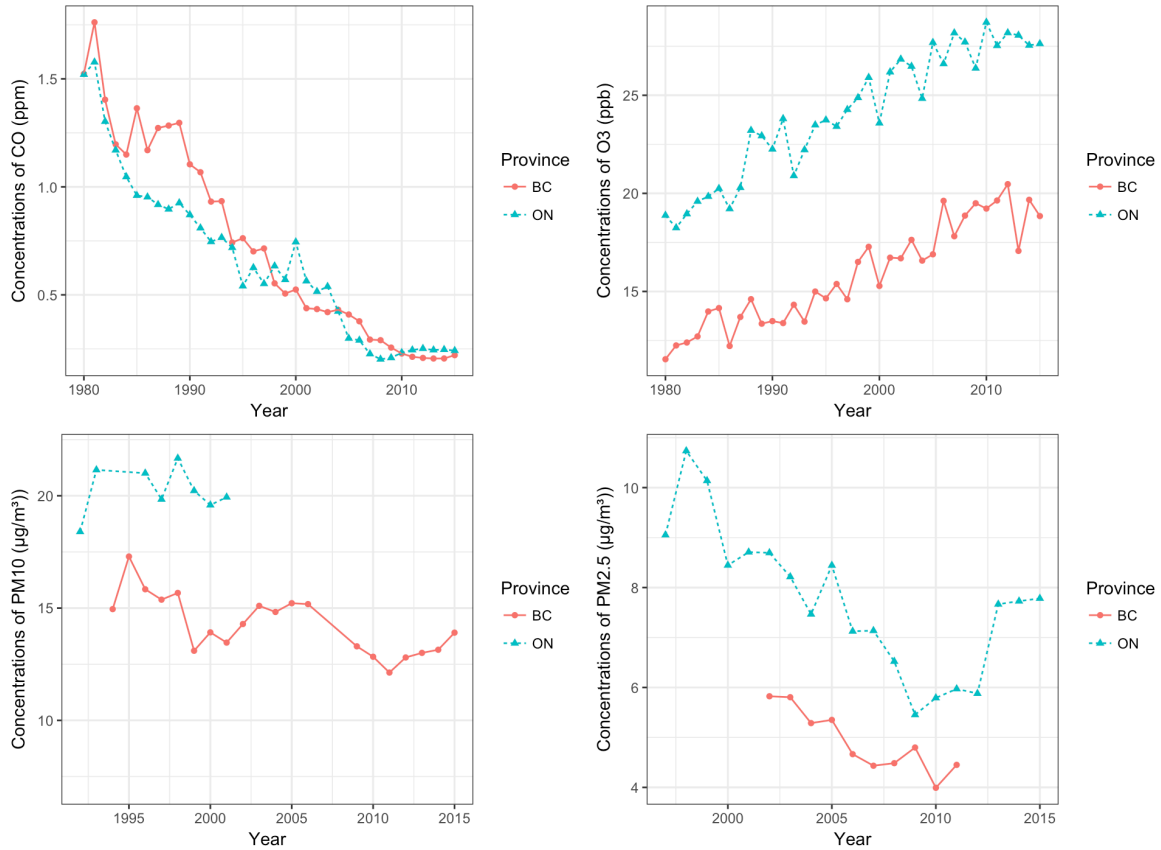


Figure 1-1 Air pollution of BC and Ontario from 1980 to 2015<sup>‡</sup>

Figure 1-1 shows that in the two provinces, the concentrations of CO, PM<sub>10</sub> and PM<sub>2.5</sub> all decreased over years, while ozone increased over years. The concentration of CO in Ontario is lower than BC before 2008 and is higher than BC after 2008. The concentrations of O<sub>3</sub> and PM<sub>2.5</sub> in Ontario are higher than BC. PM<sub>10</sub> in Ontario is higher than BC before 2002. It

<sup>‡</sup> The pollutants' concentration data was sourced from National Air Pollution Surveillance and some data were missing. <http://maps-cartes.ec.gc.ca/rnsps-naps/data.aspx?lang=en>  
 First, for each pollutant at each pollution monitor, I obtained the yearly averaged concentration by calculating the mean of the daily concentrations in each year. Second, for each pollutant, I calculated the mean of the yearly averaged concentration of all monitors in each province. Third, the results obtained in the second step were used to draw the variation trend lines in figure 1-1.

should be noted that in Ontario, all the PM10 data from 2002 to 2013 are missing and most of the PM10 data from 2014 to 2015 are missing; in BC, most of the PM10 data from 2007 to 2008 are missing, and all the PM2.5 data from 1980 to 2001, as well as from 2012 to 2015 are missing.

The school data sets in this study consist of BC provincial test results from the school years of 2006/2007 to 2013/2014 and Ontario provincial test results from school years of 2007/2008 to 2014/2015, as well as corresponding air pollution and weather data sets. Our datasets are panel data, which are “data for multiple entities in which each entity is observed at two or more time periods”.<sup>40</sup> The regression of panel data is a widely used analytical method in social science and econometrics. In our study, the entity is a combination of subject-grade-school in BC, since there are multiple subjects and grades; it is school in Ontario. The cross-sectional information of multiple schools can reflect the differences between the schools, while the time series information can reflect the changes within schools over time.

There are some challenges in testing this relationship. First, there are many other factors which could affect student test result, such as school quality, weather, difficulty levels of tests. Separating the air pollution’s impact from other factors, is one of the most crucial processes. A fixed-effects model will be built in this study and the subject-grade-school fixed effects can be controlled by including a set of dummy variables. The weather effects can be controlled by including weather in the regression as controlled variable. For difficulty level

of the test, both in BC and Ontario, test developers and administrators are required to ensure that the difficulty of provincial test is similar each year and the tests are administered under a rigorous standards-based guide and procedure.<sup>2, 10</sup> Therefore, the difficulty level of the provincial tests in this study can be considered time-invariant. And, even if there are some difficulty changes over time that affect all students equally, they can be captured by the time fixed-effects dummy variables included in the model. Secondly, compared with other factors such as school quality, the air pollution's impact on student's performance is relatively small, therefore, to test and draw a relationship (if there is one), it is crucial to have large size samples with large range in values. The provincial tests in the two provinces are taken by almost all the students in certain grades which makes the datasets have large amounts of observations and presumably makes the result of this study have validity. Thirdly, based on the literature review of the recent studies, it is found that air pollution can decrease student test results by diminishing their cognitive abilities.<sup>19, 32</sup> In this study, after regressing school test results on air pollution using the fixed-effects model, a robustness check is conducted to check whether the relationship remains robust. A heterogeneity check is conducted to check whether this impact differs between different sub-groups. However, there are still some weaknesses in our study that could hinder us from estimating a causal relationship. For example, using the air pollution data from the closest stations for schools rather than school specific air pollution data; using school aggregated data rather than student individual data; missing data in the air pollution datasets. (See chapter 6)

According to the analyzing results, I find that in British Columbia, there is a negative impact of ambient air pollution on student's test results: a one standard deviation increase in AQI leads to a 0.23 percentage points decrease in student average grade. In Ontario, I did not find a significant association, and this might be because of not having enough observations in school dataset and lots of missing data in air pollution dataset. Furthermore, it is also found that in BC, ambient air pollution's impact students with special needs experience about 3.4 times of the average impact of other students.

## CHAPTER 2 LITERATURE REVIEW

### 2.1 Air Pollution Effects on Productivity: Pathological and Non-Pathological

Industrial pollution together with meteorological conditions has resulted in several severe air pollution episodes that have caused significant loss of life, including 1930 Meuse Valley Fog in Belgium, 1948 Donora Smog in the US and 1952 London Smog in UK.<sup>§</sup> After these severe incidents, air pollution's impact on human health has received attention from scientists, environmentalists and governments. Consequently, many studies have explored the relationship between air pollution and human health, and the public has a greater understanding of the detrimental impacts of air pollution.

A large body of evidence from toxicology and epidemiology shows that exposure to CO is harmful to human's health. CO inhibits the oxygen in blood to be delivered all over the body, especially to vital organs and tissues such as heart and brain, which will result in death by asphyxia or brain hypoxia. CO also causes cognitive sequelae in human, which cannot be prevented by hyperbaric oxygen therapy or any other therapy to date.<sup>32</sup> Moreover, lower pulmonary function has been linked to CO in the uterus or early childhood, and this lower

---

<sup>§</sup> The 1930 Meuse Valley Fog resulted in more than sixty of mortality and thousands of illness between December 3<sup>th</sup> and December 5<sup>th</sup> in 1930 (Firket 1931). The 1938 Donora Smog occurred on October 27<sup>th</sup>, 1948, caused twenty of mortality and more than seven thousand of illness, which was one third to a half of local population at that time (Ciocco et al. 1961). The 1952 London Smog lasted from December 5<sup>th</sup> to December 9<sup>th</sup> in 1952, resulted in four thousand of mortality and ten thousand of illness (Logan et al. 1953).

pulmonary function could also lead to lower educational performance and employee's productivity in the following decades.<sup>6, 18, 23, 34</sup>

Particulate matter primarily reduces human body's pulmonary and cardiovascular function.<sup>39</sup> These diseases mainly express as asthma, heart attacks and other cardiovascular events, which eventually lead to hospitalizations and even mortality.<sup>9, 33</sup> Besides the above serious health problems, PM2.5 also results in subtle effects including changes in blood pressure, irritation in the ear, and mild headaches.<sup>1, 13, 33</sup> These subtle effects are very likely to be unnoticed because they usually do not result in clinic visits or hospitalization.

However, these effects could also decrease student performance and employee productivity in many sectors of economy. Moreover, particulate matter, especially PM2.5 does not break down rapidly after penetrating buildings, so it's not very helpful to reduce people's exposure to PM2.5 by being indoors.<sup>30, 43, 45</sup>

Ozone constraints air in the alveoli via muscles in the airways, which results in various illnesses and health problems, including: difficulty in breathing deeply and vigorously, coughing, sore throat, aggravation of lung diseases, chronic obstructive pulmonary disease.<sup>44</sup>

To summarize, air pollution can affect student's performance and employee's productivity by way of illness, negatively affecting people's brain activity, etc. People with these illnesses or health problems could perform worse in study and work than usual.

Furthermore, besides health problems, people may decide not to go to school and work or reduce studying and working hour because of some problems caused by air pollution, such as slow traffic or traffic jam caused by smog, or bad mood caused by air pollution.

## **2.2 Findings of Air Pollution's Impact on Student's Educational Performance**

A few studies have examined the relationship between air pollution measures and student's performance. An association is found in most of the papers, and some of them can be interpreted as a causal effect.<sup>19</sup> However, the impact level varied in these studies. Ham et al. (2011) found a reduction in air pollution in California generally increases academic performance on standardized tests by a significant amount. The effects were strongest for ozone, PM2.5, and PM10, whereas effects for CO were insignificant in most of the cases.<sup>14</sup> Lavy et al. (2014) found that both PM2.5 and CO exhibit a robust negative relationship with test outcomes in Israel. PM2.5 had a larger effect on groups with higher rates of asthma, while CO's effect is more homogenous across different demographic groups.<sup>19</sup>

In addition to air pollution on the exam day, previous fetal exposure to air pollution may also have impact on their test outcomes. Bharadwaj et al. (2014) examined this relationship using birth data, subsequent test outcomes, demographic information, air pollution and meteorological data in Santiago, Chile. This study extracted the fixed locational effect and time-invariant family characteristics by comparisons across siblings. And a significant negative effect from fetal exposure to CO on test outcomes was found.<sup>34</sup> Similarly, Lavy et al. (2014) also examined the long-term impacts in Israel. Datasets of PM2.5 measures,

student's matriculation scores, subsequent post-secondary school performance and corresponding payment in adulthood were used to analyze air pollution's long-term impact. A significantly negative relationship was found, which is that students exposed to higher levels of pollution during test were found to have fewer years of post-secondary education and lower income in adulthood.<sup>20</sup>

### **2.3 Findings of Air Pollution's Impact on Employee Productivity**

Employee productivity is similar with student test result and they both belong to cognitive performance. I also did a research of air pollution's impact on employee productivity. Air pollution's impact on employee productivity has been examined in recent studies using data from various kinds of plant. Chang et al. (2014) examined the effect of air pollution on pear-packer's productivity using three years' data from a pear-packing factory in California. A threshold impact was found that when PM<sub>2.5</sub> exceeded 15 µg/m<sup>3</sup>, there was a negative relationship between PM<sub>2.5</sub> measures and worker's packing speed, which is that an increase of 10 µg/m<sup>3</sup> in PM<sub>2.5</sub> resulted in a decrease of \$0.41/h in worker's packing. But ozone was found have no impact on worker's productivity in this study.<sup>42</sup> However, another study done by Zivin et al., also located in California, using two years' box-packing data from a farm, showed different results for ozone. It was found that ozone pollution had a significantly negative impact on worker's productivity when ozone level went well below US federal air quality standards: a 10 ppb decrease of ozone resulted in a 4.2% increase of worker's productivity.<sup>12</sup> The two studies have similar type of work and locate in the same

state, however, the ozone's impact on worker's productivity is quite different. This could be because that the ozone's concentrations are quite different in the two studies (with a mean of 31.6 in Chang's study and a mean of 47.77 in Zivin's study). And this difference in ozone's concentration might be that the time periods in the two studies are in different seasons and ozone's concentration is related with sunlight which varies with season.

China's air pollution level has been increasing over the past decades and is very high in recent years. Studies using Chinese data can uncover the relationship pattern between air pollution and worker's productivity under high pollution level. A study was done by Li et al. in 2015 using data from a textile plant in Hebei province of China, which is very close to Beijing. The PM<sub>2.5</sub> concentration ranges from 10 to 773  $\mu\text{g}/\text{m}^3$ , which is very wide. It was found that with the increase of PM<sub>2.5</sub> concentration from 10 to 200  $\mu\text{g}/\text{m}^3$ , the productivity decreased by 15%, but the leveling drops off with further increase in PM<sub>2.5</sub> pollution and this relationship is non-linear. It was also found that the amount of defective textile increased as PM<sub>2.5</sub> increases.<sup>41</sup>

## **2.4 Compare and Conclusion of These Findings**

The findings among these studies indicate that it is possible that a causal relationship between air pollution and educational outcomes exists. And the difference of results for the same pollutant in different countries/regions could be due to different pollution levels, different response levels to air pollution for sub-populations, different test/work patterns, different education/work mechanisms, etc. Moreover, the existing literature on air

pollution's impact in Canada so far is mainly focused on health impact.<sup>4, 7, 8, 11, 28, 31, 38</sup>

Examining this relationship in Canada can help to provide evidence for decision making when designing regulations that affect air pollution levels. It can also contribute to providing more evidence by exploring broader data in this field.

## CHAPTER 3 DATA

### 3.1 Panel Data

The dataset used in this study is panel data. Panel data have many advantages over the time-series and cross sectional sets. First, there are two indexes (i and t), which help provide large numbers of observations ( $i \cdot t$ ). Large size of samples give large degrees of freedom and will reduce issues associated with the collinearity among the variables. However, it should be noted that more observations do not always mean more information (if repeated observations are not independent), and there could be unobserved heterogeneity, which is a form of omitted variable bias and can be controlled by fixed-effects model used in this study (Chapter 4). Second, the cross-sectional information and time series information of panel data can observe the complexity of human behavior in a comprehensive way. For instance, the example stated in Hsiao's book, married women were found to have an average yearly labor-force participation rate of 50%. This could be interpreted in two ways: each married woman has 50% chance to work each year; or 50% of married woman always goes to work and the other 50% never goes to work.<sup>16</sup> Panel data analysis can give more accurate interpretation of “between” difference and “within” changes, because it includes the information of woman’s work participation rate in different subintervals of their life cycle.<sup>3, 16</sup> Third, panel data controls for omitted variables. Panel data has time invariant features because it observes the same entities over time. Therefore, some of the omitted variables can be controlled without observing them.

## **3.2 Data Sets used in this study**

### **3.2.1 Educational Data**

Both BC and Ontario educational datasets we obtained are the test results averaged on each subject of each grade of each school. We were not able to obtain data for individual students.

#### **3.2.1.1 British Columbia Provincial Test Data**

In British Columbia, grades 10, 11 and 12 are required to take provincial test for certain subjects, and the test results contribute to the final grade of these subjects. There are six exam sessions each year before the school year of 2013/2014, which are January, April, May, June, August, November with June being the largest followed by January. Since 2013/2014, another exam session of October has also been added. For each subject exam, one student could participate in the exam more than once and the highest score is recorded as the final result. The difficulty level of the provincial test does not change, which is appropriate for statistical analysis. The dataset is the average results at the school level, district level and provincial level. It contains 529 schools consisting of 135 independent schools and 394 public schools. Six sub-population groups are included in this dataset, which are “All students”, “Female”, “Male”, “Aboriginal”, “English learners” and “Students with special needs”.

#### **3.2.1.2 Ontario Secondary School Literacy Test Data**

The Ontario Secondary School Literacy Test (OSSLT) is a provincial literacy skills test for grade 10. Successful passing OSSLT is one of the 32 requirements for students to obtain the

Secondary School Diploma in Ontario. The difficulty level of the OSSLT is comparable year by year as well, which is one of the requirements for designing this test at the beginning.

The OSSLT has rigorous scoring procedure to make sure that the test results are accurate and valid. First, the scorers are required to receive professional training and pass a qualification test. Second, the choice items are scored by machine; and the open-response items are scored twice by two different scorers; whenever there is a discrepancy between the two scores, it will be scored for the third time by an expert scorer.<sup>10</sup>

The dataset we obtained is the percentage of students who successfully passed the exam and were eligible to participate in the OSSLT for the first time.\*\* It contains 581 schools, consisting of public and Roman Catholic secondary schools. The test result is the aggregated result on school level and three sub-population groups were included in the dataset, which are “all students”, “male” and “female”.

### **3.2.2 Air Pollution Data**

The air pollution data used in this study is sourced from National Air Pollution Surveillance Program (NAPS) of Environment and Climate Change Canada. The NAPS was established to monitor the ambient air pollution in populated areas across Canada in 1969. BC and Ontario are two of the earliest provinces that joined in NAPS in 1969. The NAPS has developed from a small program with only 36 monitoring sites to a broad, mature air pollution surveillance

---

\*\* It needs to be clarified that in Ontario’s school dataset we obtained, only students who were eligible to take OSSLT for the first time were included in our dataset. Students who were previous eligible, were not included in the dataset.

program with more than 700 monitoring stations in 286 sites in 203 communities located in every province and territory of Canada. The measurements were taken at regular frequencies (hourly, daily, etc.).<sup>49</sup> The map of the monitoring stations is as follows:



Figure: 3-1 NAPS monitoring station location (Sourced from NAPS 2013)

The daily measurements of CO, O<sub>3</sub>, PM<sub>10</sub> and PM<sub>2.5</sub> matched with the exam dates of the educational data are used in this study.

### **3.2.3 Weather data**

Except for air pollution, weather is another factor that could also affect student test result, for example, there are some studies found that temperature and heavy rainfall could have an impact on student performance.<sup>19, 51, 52</sup> However, weather changes with time and

location, therefore, it cannot be controlled by entity fixed effects or time fixed effects. The weather is included as controlled variable in our model.

The weather data used in this study were obtained from Environment and Climate change Canada. Daily average temperature and daily precipitation was chosen as the two weather items in this study, which could have an impact on air pollution and student's performance in test. It contains 237 weather stations in BC and 130 weather stations in Ontario.<sup>50</sup>

### 3.3 Processing of the Data Sets

To analyze the potential relationship between air pollution and student's test results, we need to merge school, air pollution and weather datasets together by some links (common variables) between these datasets, and obtain a final dataset including the test result, air pollution, and weather data for each test taken at each school in each year.

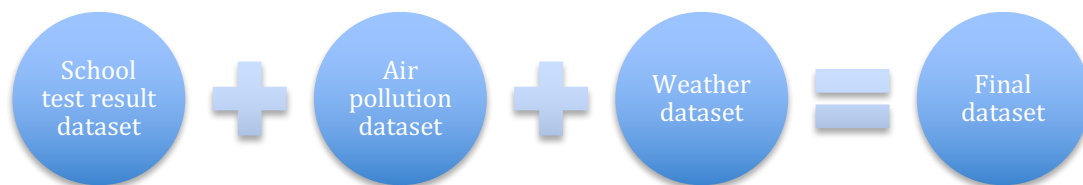


Figure 3-2 Merging school, air pollution and weather datasets

Prior to merging data, we conduct some cleaning to preprocess each dataset.

#### **3.3.1 Preprocessing of School Test Data**

➤ Figure 3-3 shows the summary of preprocessing BC school dataset:

In BC, the time period of the original school dataset is from school year of 1996/1997 to 2013/2014. Dataset from 1996/1997 to 2005/2006 was dropped because of no

exam schedule available or the quantity of students for each exam session (six/seven exam sessions each year, as stated in the above) was not available. Only school level data was selected whereas district level and province level data was dropped since we need more individual level data. Only exam marks were chosen whereas course marks

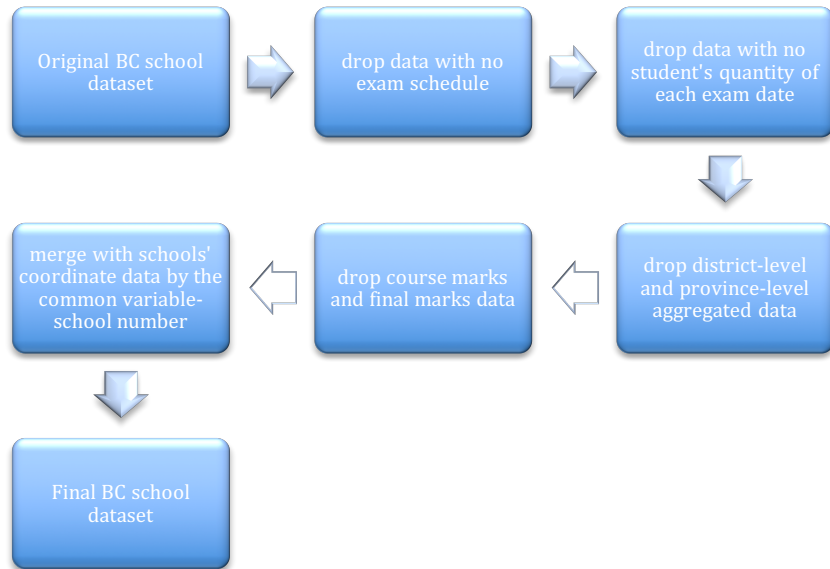


Figure 3-3 Preprocessing of BC school dataset

and final marks data were dropped, since that we focus on air pollution's impact on test result. Afterwards, the school test result data were merged with school coordinates dataset by the common item: school ID. Figure 3-4 explains the process of "merging two datasets":

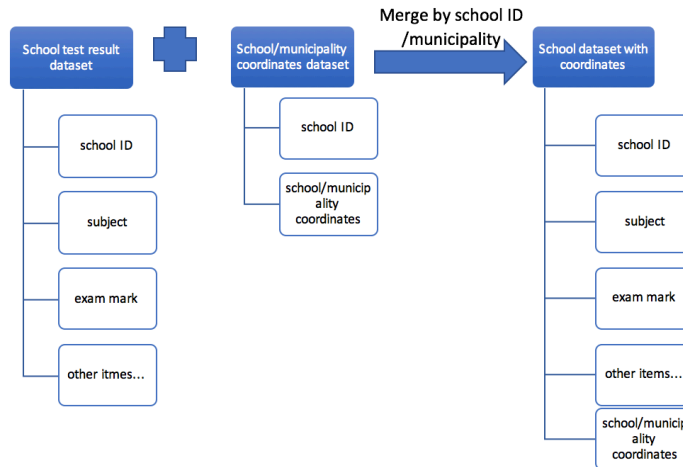


Figure 3-4 Merge school test result with school coordinates

- In Ontario, the time frame of the dataset was from school year of 2007/2008 to 2015/2016. The data of 2015/2016 was dropped because of the lack of air pollution data. Moreover, neither the school coordinates nor the school names are available in Ontario dataset. Therefore, instead of the school coordinates, the coordinates of the municipality for each school were merged with test results data by the common item: municipality<sup>††</sup>.

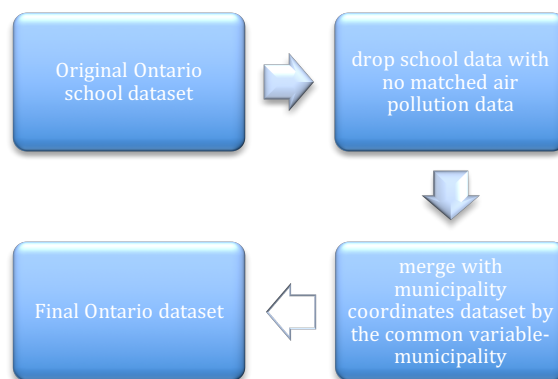


Figure 3-5 Preprocessing of Ontario school dataset

<sup>††</sup> For each school, input the municipality information in the website of Distancesto (<https://www.distancesto.com/coordinates.php>), then the coordinates will be obtained. The merging process is similar with figure 3-4.

### 3.3.2 Preprocessing of Air Pollution Data

The following flow chart summarizes the preprocessing of air pollution data:

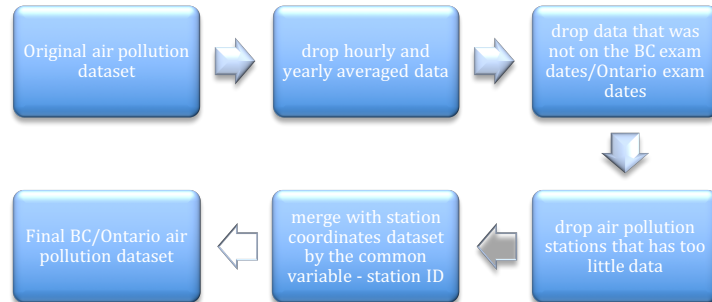


Figure 3-6 Preprocessing of air pollution dataset in BC and Ontario

- Because the schedule of the exam is recorded by day, so daily air pollution data were chosen, while hourly and yearly data were dropped.
- Only air pollution data on exam days (222 exam days for BC dataset, 8 exam days for Ontario dataset) was selected and all the other data was dropped.

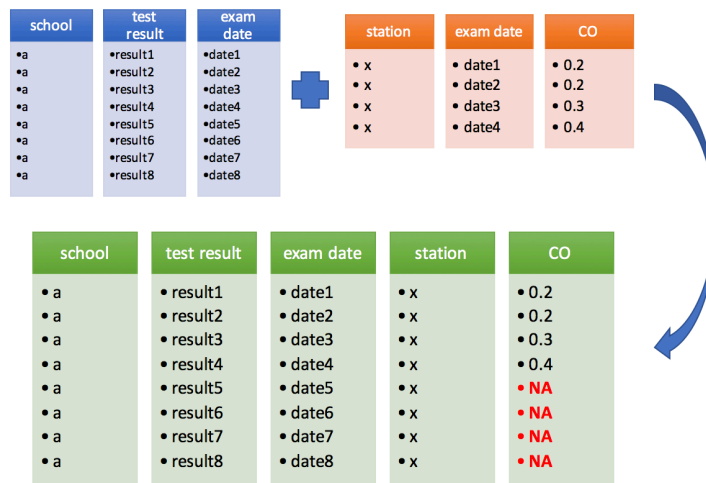


Figure 3-7 Merging school data and air station which have very few data

There are stations that only have data on some of the exam dates and have no data on other exam dates, for example in figure 3-7, station x in Ontario only have data on 4

exam days, whereas has no data on the other 4 exam days. This will produce lots of NAs when merging with school dataset.

To avoid many NAs being produced, we examined the amount of data for each station.

Figure 3-8 and 3-9 show the histograms of NAPS stations with data on 1 exam-day, 2 exam-days, ....., and the total exam-days in BC and Ontario. It shows that the number of stations with data on more exam-days (in short, stations with more data) is much more than stations with data on less exam-days (in short, stations with less data). Therefore, if we drop stations with less data, the rest of the stations' quantity is still large. Hence, to avoid producing too many NAs when merging it with school data, air pollution stations with data less than 180 exam-days in BC and 7 exam-days in Ontario was dropped.##

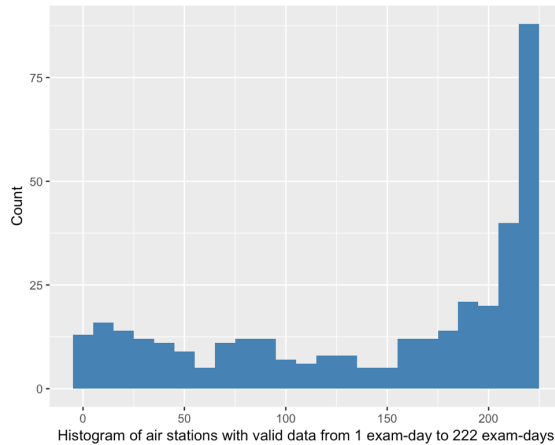


Figure 3-8 Histogram of air stations with valid data from 1 to 222 exam-days in BC

## In the following section 3.3.4, for school x, if its nearest pollution station is y, and y is "dropped" because of less valid data, the next closest pollution station will be matched for school x.

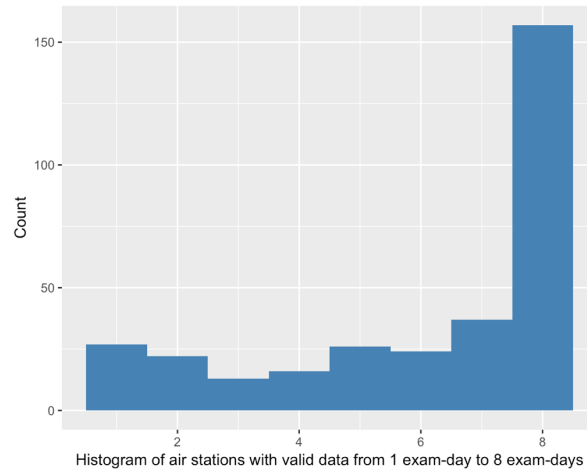


Figure 3-9 Histogram of air stations with valid data from 1 to 8 exam-days in Ontario

- The air pollution dataset does not include station’s location information, which will be needed for merging with school data in section 3.3.4. Hence, the pollution dataset was merged with the station location dataset by the common item: station ID<sup>§§</sup>.

### **3.3.3 Preprocessing of Weather Data**

The preprocessing of weather data is similar with air pollution data, and figure 3-10

summarizes this process:

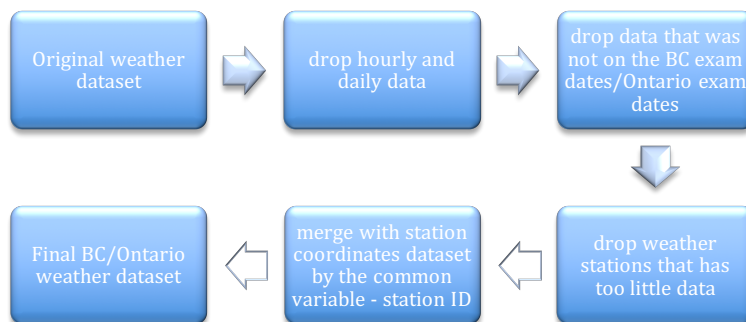


Figure 3-10 Preprocessing of weather dataset in BC and Ontario

<sup>§§</sup> The merging process is similar with figure 3-4.

- Only daily data was chosen, which are daily mean of temperature and daily total precipitation.
- Only weather data on exam days (222 exam days for BC dataset, 8 exam days for Ontario dataset) was selected and all the other data was dropped. To avoid producing too many NAs when merging with school data, weather stations with less data (less than 180 exam-days in BC and less than 7 exam-days in Ontario) were dropped.
- The weather data was merged with weather station coordinates dataset by the common item: weather station ID\*\*\*.

#### **3.3.4 Merge School Data, Air Pollution and Weather Data**

- Merge the closest air pollution station and weather station for each school. Since the process of merging closest pollution stations for schools and merging closest weather stations for schools are the same, we'll take the air pollution dataset as an example in this section, and the steps for weather dataset will be the same. The following functions are used to calculate the distance between each school and each air pollution station.

$$\text{tr} = \pi/180 \tag{3-1}$$

$$a = \sin^2(\Delta\varphi/2 \times \text{tr}) + \cos(\varphi_1 \times \text{tr}) \times \cos(\varphi_2 \times \text{tr}) \times \sin^2(\Delta\lambda/2 \times \text{tr}) \tag{3-2}$$

$$d = R \times 2 \times \text{atan2}(\sqrt{a}, \sqrt{1-a}) \tag{3-3}$$

---

\*\*\* The merging process is similar with figure 3-4.

where:

$\pi$ : the mathematical constant (approximate 3.14);

$\text{tr}$ : transfer degrees to radians;

$\varphi_1$ : the latitude of school;

$\varphi_2$ : the latitude of air pollution station;

$\Delta\varphi$ : the difference of  $\varphi_1$  and  $\varphi_2$ ;

$\Delta\lambda$ : the difference of school's longitude and air pollution station's longitude;

$R$ : the earth's radius (mean radius = 6,371km);

$d$ : distance of school to air pollution station.

Suppose the total number of schools is  $n$ . For one school, the distances to all the air pollution stations are calculated. Then compare the distances and assign the station with minimum distance for this school. Then do the same calculation for the second school, the third school, ....., the  $n^{\text{th}}$  school.

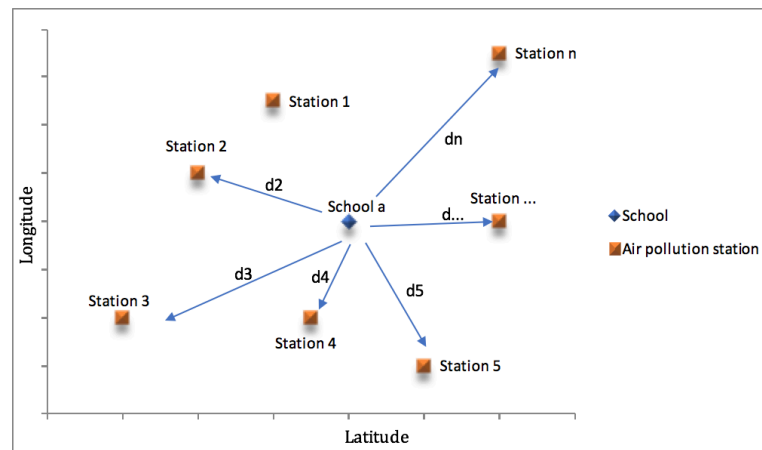


Figure 3-11 Distances between one school and air pollution stations

For example, as shown in figure 3-11, for school a, calculate its distances to all the

stations and recorded as d1, d2, d3, d4, d5, ....., dn. After that, compare and choose the minimum distance, for example d4, then assign station 4 for school a.

- Referring to “Representativeness and classification of air quality monitoring stations” of European Commission, PM10 and O3 will be transformed into other chemical species if the transport distance exceeds 100km. In this study, schools with distance to the closest air pollution station larger than 100 km were dropped. The following figures show the distribution of distances between schools and NAPS stations in BC and Ontario.

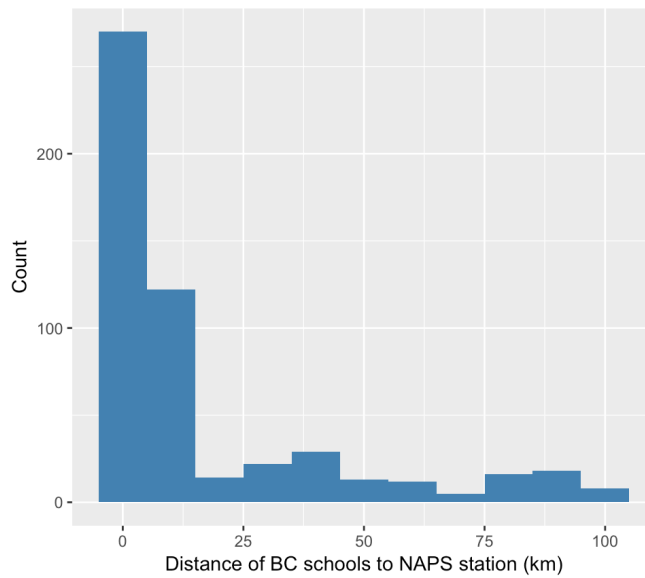


Figure 3-12 Distance from schools to NAPS stations in BC

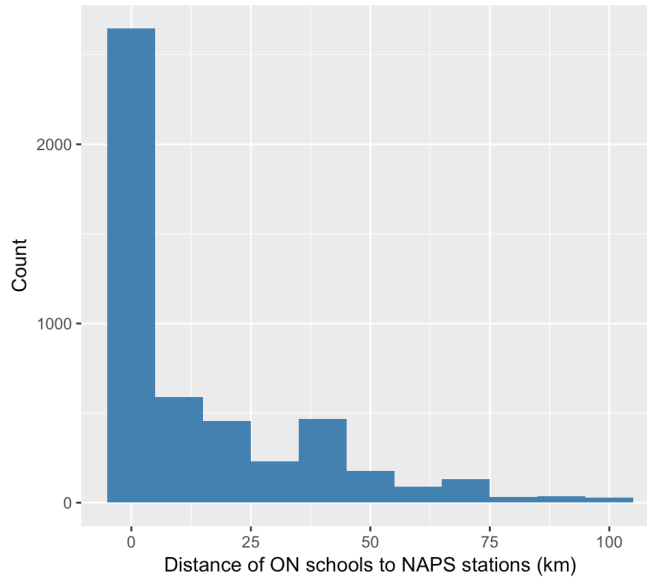


Figure 3-13 Distance from schools to NAPS stations in Ontario

- After the above processes, the BC educational dataset includes the test results for each subject of each grade of each school, as well as the closest air pollution station ID and weather station ID. We want to merge corresponding air pollution data and weather data for each test result by exam date. However, in BC, there were six exam sessions annually from 2006/2007 to 2012/2013 and seven exam sessions for 2013/2014. Each exam session includes many subject tests, which means that for each subject, the exam happened more than once and less than six or seven times annually. For example, in school year of 2006/2007, the average grade of English for grade 10 in school x is 77, which is the average result of exams on “2006-11-15”, “2007-01-29”, “2007-04-19”, “2007-06-20” and “2007-08-10”. How should we match the air pollution data and weather data on multiple dates with one test result? The following figure shows the method we used for air pollution data and the steps for weather data will be the same.

We assign a weight for air pollution data of each exam date and calculate the weighted air pollution data. Afterwards, match the school data and weighted air pollution data together, as shown in figure 3-14.



Figure 3-14 School test result matching with weighted air pollution

And to do this, we obtained another dataset from BC government, including the number of students who took the exam on each exam date. The number of students taking the exam on each exam date was used to calculate the weight for each exam date. The functions are as follows:

$$W_{ijdy} = N_{ijdy} / N_{ijy} \quad (3-4)$$

Where:

$W_{ijdy}$ : the weight for the test result of subject  $j$  at school  $i$ , on exam date  $d$ , in year  $y$ ;

$N_{ijdy}$ : the number of students taking exam of subject  $j$  at school  $i$ , on exam date  $d$ , in year  $y$ ;

$N_{ijy}$ : the total number of students taking exam of subject  $j$  at school  $i$ , in year  $y$ .

The weighted air pollution data is calculated with air pollution data on each exam date and the corresponding weight. The functions are as follows:

$$A_{ijy} = A_{ijdy} \times W_{ijdy} \quad (3-5)$$

Where:

$A_{ijy}$ : the weighted air pollution data for the exam of subject  $j$  at school  $i$ , in year  $y$ ;

$A_{ijdy}$ : air pollution data for the exam of subject  $j$  at school  $i$ , on exam date  $d$ , in year  $y$ ;

$W_{ijdy}$ : the corresponding weight for the test result of subject  $j$  at school  $i$ , on exam date  $d$ , in year  $y$ .

However, there are some missing data in air pollution dataset and weather dataset, which may lead to the shrinking of the weighted values. To avoid this shrinking, the weighted result will be divided by the total weights that have valid air pollution/weather data (no NAs). For example, if CO's monitoring values on the exam days are 0.2, 0.3, NA and NA, and the corresponding weights are 50%, 10%, 20% and 20%, then the weighted CO is:  $0.2 \times 50\% + 0.3 \times 10\% + NA \times 20\% + NA \times 20\% = 0.13$ , which is smaller than both 0.2 and 0.3. And if we use 0.13 divided by (50%+10%), we will get 0.217, which makes more sense.

### 3.4 Visual Overview of the Dataset

Figure 3-15 is a map showing the locations for schools, air pollution stations and weather stations in BC and Ontario.

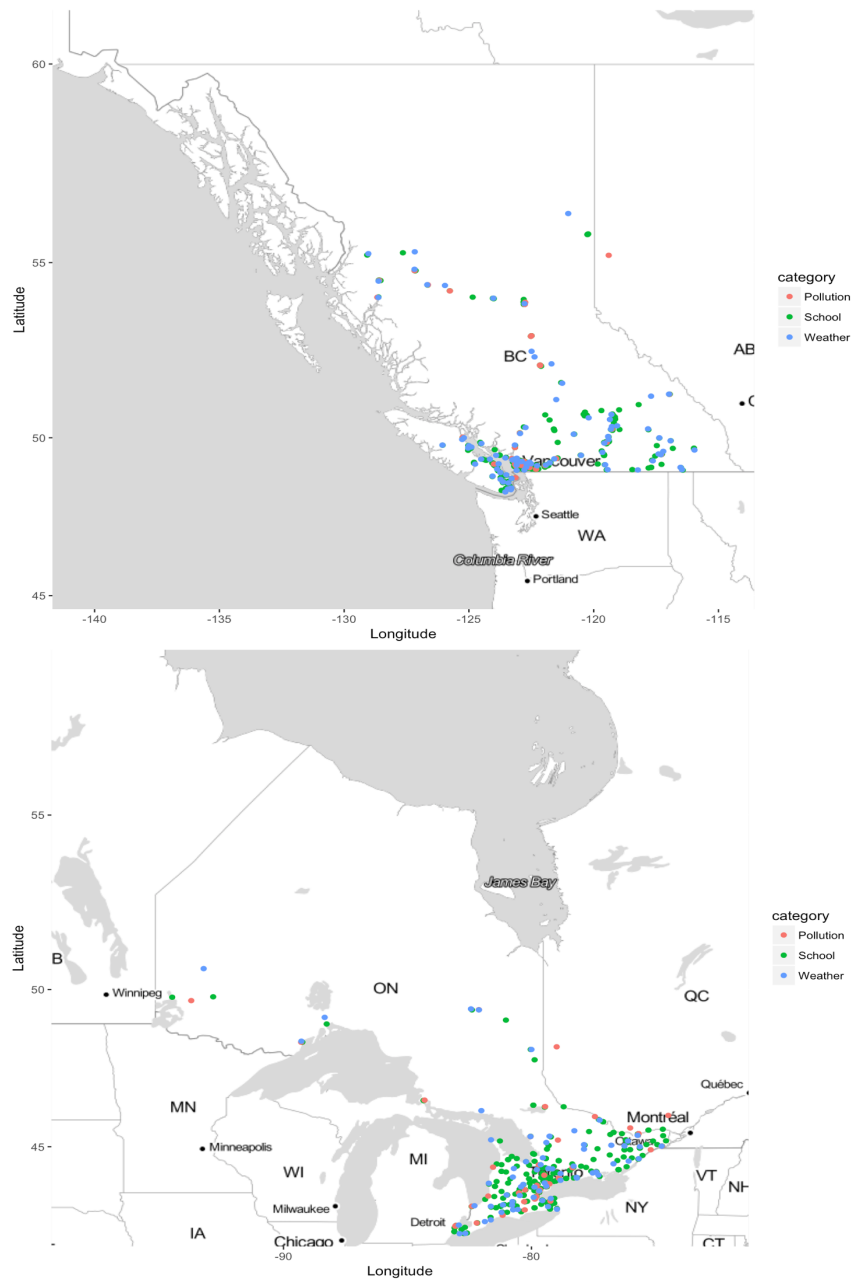


Figure 3-15 Distribution of schools, pollution and weather stations

### **3.4.1 School Test Result Data**

Before fitting the data in the fixed-effects model, I conducted a visual overview of the data distribution. Figure 3-16 shows the spread out of the average grade for different subjects/schools over years in BC. The blue points are the distribution of average grades for

each subject of each grade of each school in each year, while the green points are the yearly averaged values of blue points. To overview the distribution of the samples better, the points has been “jittered” in the following figures from figure 3-16 to figure 3-23. The darker the color is, the more samples there are.

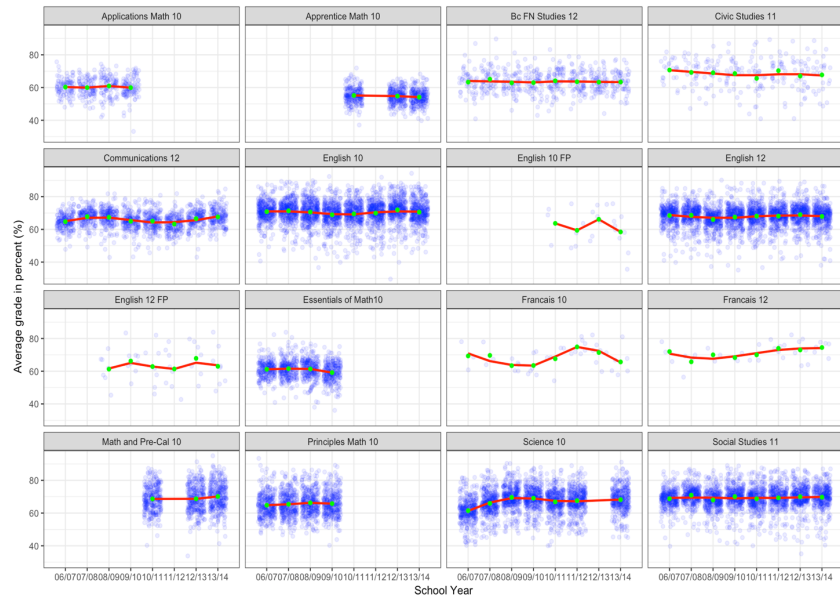


Figure 3-16 School test results by subjects over years in BC

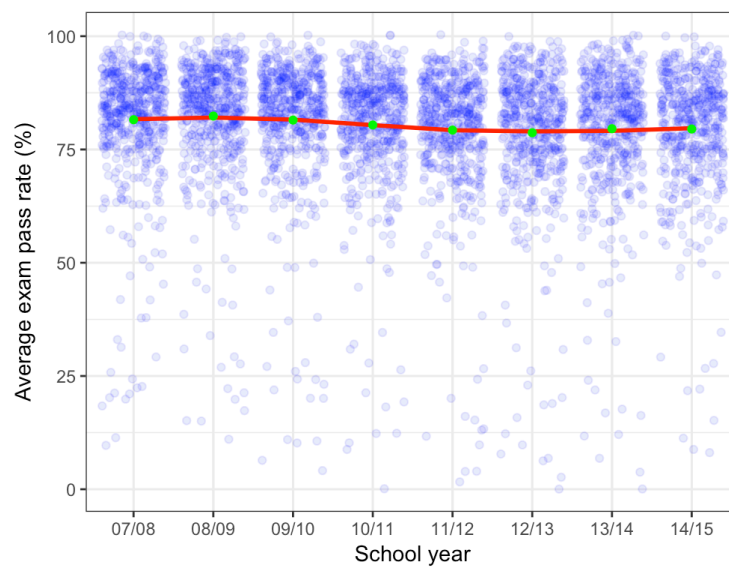


Figure 3-17 Average exam pass rate for schools in Ontario

Figure 3-16 and 3-17 show that there is some fluctuation in the test results each year, and air pollution could be one of the factors that caused this variation. Moreover, for overall looking, the yearly average test result didn't vary drastically, which means that the difficulty level of the exams is relatively stable over these years. Figure 3-16 shows that English 10, English 12, Science 10 and Social Studies 11 have the most samples in BC dataset.

### **3.4.2 Air Pollution Data**

The following figures show the variation of the pollutant's concentrations on the exam dates over years.

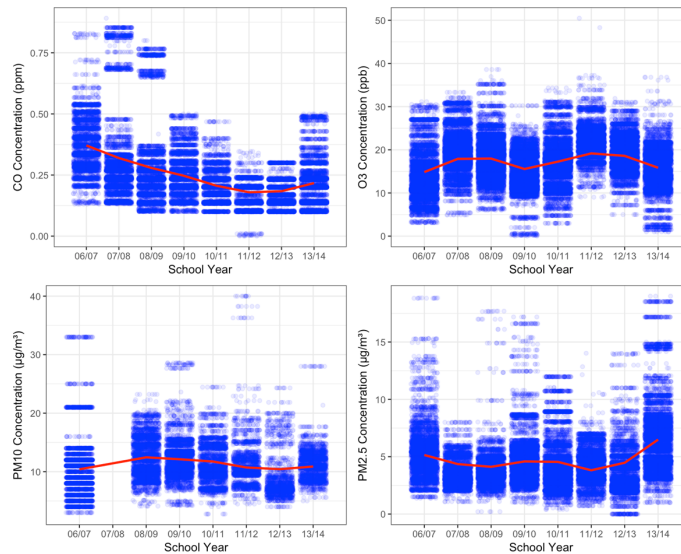


Figure 3-18 Pollutant concentrations on exam dates over time in BC

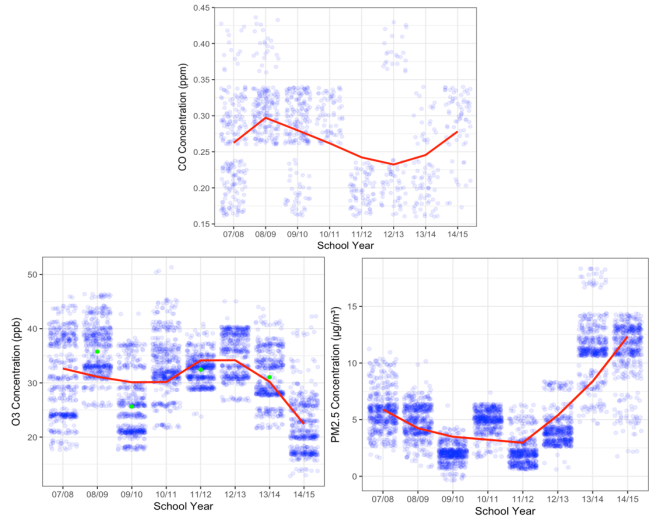


Figure 3-19 Pollutant concentrations on exam dates over time in Ontario

Figure 3-18 and 3-19 show that pollutant concentrations differ across these years. The PM10 and PM2.5 curves within BC do not look quite similar, it might be because that some parts of the PM10 data are missing, and if more data were available, the graph may look different. It should be noted that we were not able to obtain PM10 data in Ontario and there are lots of missing data in Ontario's CO data. And this might arise bias in Ontario's analysis result, which will be discussed in Chapter 5.

### 3.4.3 Weather Data

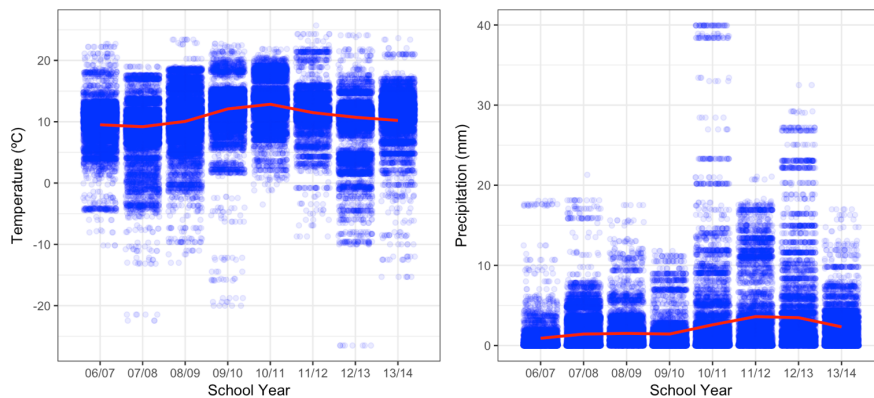


Figure 3-20 Weather data varies over time in BC

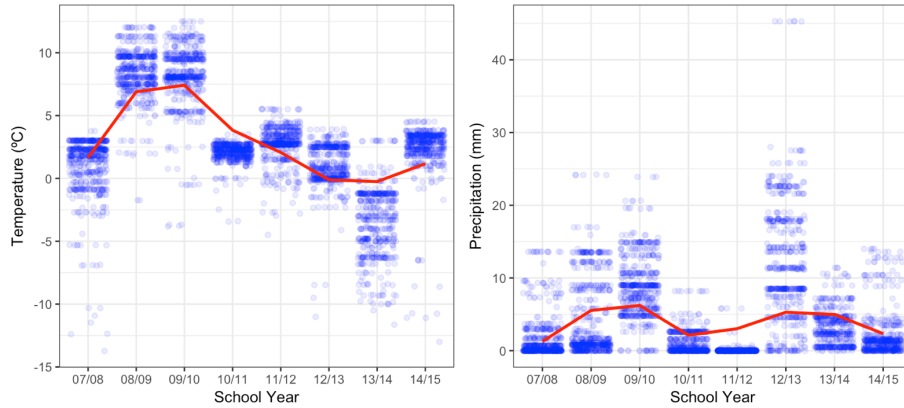


Figure 3-21 Weather data varies over time in Ontario

Overall, the weather data in the two provinces is comparatively stable but with variation in between. And in Ontario, the temperature of exam day in school year of 2008/2009 and 2009/2010 is very high, while 2013/2014 is very low.

**3.4.4 School test result varies with air pollutant concentrations across years**

**3.4.4.1 BC**

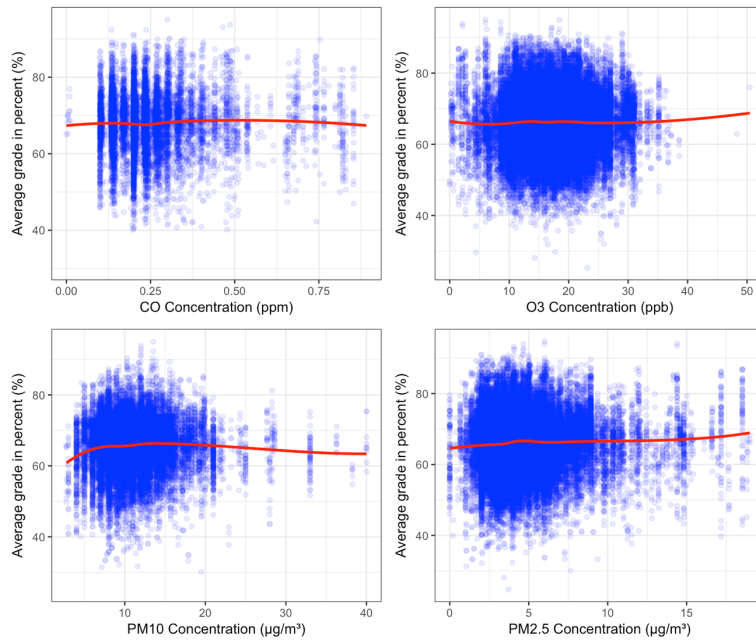


Figure 3-22 School test result varies with air pollutant concentrations in BC

Figure 3-22 shows the simple plot of BC school test results with air pollutant concentrations without any fixed-effects control.<sup>†††</sup> It shows that with the increasing of CO and PM10 concentration, the average test grade goes down overall, with fluctuations in between. While for O3 and PM2.5, with the increasing of these two pollutants, the average grade seems going up overall, with fluctuations in between. However, the variation trends between school test results and the four pollutants are not very clear. There are other factors which can also affect school test results, such as weather factors, school fixed effects, time fixed effects. Therefore, it would make more sense to explore this relationship controlled for the fixed effects (Chapter 4).

#### 3.4.4.2 Ontario

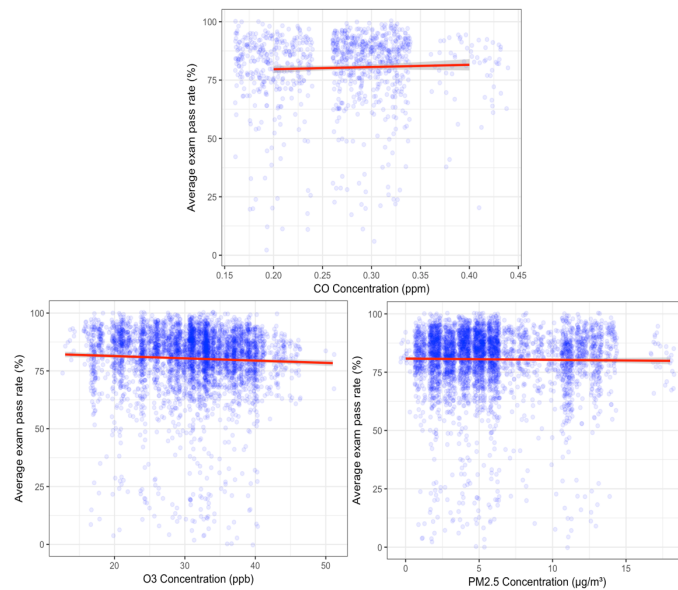


Figure 3-23 OSSLT result varies with air pollution data in Ontario

<sup>†††</sup> The figure is a simple visual overview of school test results and air pollution, without control any fixed-effects. It is not derived from the fixed effects model.

Figure 3-23 shows OSSLT results with air pollutants' concentration without any fixed-effects control.<sup>###</sup> The exam pass rate of OSSLT decreases with the increase of O<sub>3</sub> concentration. However, same with in BC, the variation trends for all the pollutants are not very clear because the fixed effects are not controlled in the figure.

---

<sup>###</sup> The figure is a simple visual overview of school test results and air pollution, without control any fixed-effects. It is not derived from the fixed effects model.

## CHAPTER 4 METHODOLOGY

### 4.1 Introduction of Empirical Strategy

In this study, we examine whether there is an impact of air pollution on school test results using BC and Ontario's school test result data, corresponding air pollution and weather data. To do this, a fixed-effects model controlled for entity fixed effects and time fixed effects will be built in this study. First, we use air pollutants as independent variables in the fixed effects model to examine if there is a relationship between school test results and air pollutants. Second, since the air pollutants may affect students synergistically and they might be correlated with each other, we calculate an Air Quality Index (AQI) and examine the relationship between AQI and school test results. Third, to make the interpretation of the research results straightforward, logarithmic transformation is conducted. Fourth, the AQI is broken into several sub-groups, to test the magnitudes of the coefficients under different levels of AQI, i.e. non-linear relationship.

However, there are some weaknesses that might hinder us from observing the causal relationship, such as the test results are school aggregated not individual level, the air pollution data are station specific not school specific, remaining unobserved variables.

### 4.2 Fixed-effects model (with air pollutants as independent variables)

There are fixed-effects model and random-effects model to choose to do the panel data analysis. Both fixed-effects model and random-effects model can remove the omitted variables bias in panel data by measuring the variation within groups. However, one

important assumption for using random-effects model is that the unobserved effect is uncorrelated with the independent variable. In our study, we cannot be sure about this, for example, the class attendance rate of students could affect a school's test result, but it could also be correlated with air pollution (students may decide not to go to school because of air pollution). Furthermore, Hausman test can be used to decide which of two models to choose.<sup>47</sup> The null hypothesis of Hausman test is preferring random-effects model to fixed-effects model. The Hausman test result for our datasets is as follows:

	<b>p-value</b>	<b>chisq</b>	<b>df</b>
<b>BC data</b>	0.0006	23.72	6
<b>Ontario data</b>	2.003e-12	63.78	5

Table 4-1 Hausman test result

The p-values in both BC and Ontario, i.e. 0.0006 and 2.003e-12, are less than 0.05 (even less than 0.001), so the null hypothesis is rejected and fixed effects model is chosen for this study.

As stated in the above, fixed-effects model can control for omitted variables in panel data which vary across entities but are constant over time. It assigns one intercept for each entity and all the intercepts are a set of dummy variables, which absorb the omitted variable's impacts that vary across entities but do not change over time. Similarly, fixed-effects model can also control for unobserved effects that are constant across entities but evolve over time, which are called time fixed effects. For example, when national

educational policies evolve, they help to improve educational outcomes of all subjects in schools. The time-fixed effects can be controlled by including another set of dummy variables in the fixed-effects model. Except for these fixed effects, there are some variables that change with time and locations, and could also affect the school test results, for example, temperature and precipitation.<sup>19</sup> Therefore, these two factors were included as controlled variables in our model. Ultimately, the following model is built:

$$S_{ijgt} = \beta_0 + \beta_1 CO_{ijgt} + \beta_2 O3_{ijgt} + \beta_3 PM10_{ijgt} + \beta_4 PM2.5_{ijgt} + \beta_5 Temp_{ijgt} + \beta_6 Prec_{ijgt} + \gamma_2 D2_{ijg} + \gamma_3 D3_{ijg} + \dots + \gamma_n Dn_{ijg} + \delta_2 B2_t + \delta_3 B3_t + \dots + \delta_T BT_t + \epsilon_{ijgtp} \quad (4-1)$$

Where:

$i, j, g, t$ : indices for school, subject, grade, and year;

$S_{ijgt}$ : student's test result, the dependent variable;

$CO_{ijgt}$ : monitoring value of CO concentration, independent variable;

$O3_{ijgt}$ : monitoring value of O3 concentration, independent variable;

$PM10_{ijgt}$ : monitoring value of PM10 concentration, independent variable;

$PM2.5_{ijgt}$ : monitoring value of PM2.5 concentration, independent variable;

$Temp_{ijgt}$ : monitoring value of temperature, controlled variable;

$Prec_{ijgt}$ : monitoring value of precipitation, controlled variable;

$D2_{ijg}, D3_{ijg}, \dots, Dn_{ijg}$ : dummy variables controlled for entity fixed effects; the entity is a combination of subject-grade-school in BC, since there are multiple subjects and grades; it is school in Ontario;

$B2_t, B3_t, \dots, BT_t$ : dummy variables controlled for time fixed effects;

$n$ : the quantity of entities;

$T$ : the quantity of years;

$\beta_0$ : the intercept of the regression;

$\epsilon_{ijgt}$ : the error term of the regression;

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \gamma_2, \dots, \gamma_n, \delta_2, \dots, \delta_n$ : coefficients.

### 4.3 Air Quality Index

There might be some correlation among the four pollutants since that one pollutant's emission source may also be the emission source for other pollutants. To test this, Pearson test will be done in this step. Based on the Pearson test result and the regression result of the fixed-effects model (4-1), i.e. if there is a correlation among the independent variables, to avoid multi-collinearity<sup>§§§</sup>, the Air Quality Index (AQI) will be calculated to integrate the four variables together. The AQI will be calculated referring to Quebec government's calculation method and US EPA's reference value<sup>\*\*\*\*</sup>. The functions are as follows:

$$\text{Sub-AQI}_i = (\text{Pollutant}_i / \text{Ref-value}_i) \times 50 \quad (4-2)$$

$$\text{AQI} = \text{Max} (\text{Sub-AQI}_i) \quad (4-3)$$

---

<sup>§§§</sup> In statistics, multi-collinearity (also collinearity) is a phenomenon in which two or more independent variables in a multiple regression model are highly correlated, which means that one independent variable can be linearly predicted from the others. In this case, the estimated coefficients of the multiple regression may change erratically due to small changes in the model or the data.

<sup>\*\*\*\*</sup> [http://www.iqa.mddefp.gouv.qc.ca/contenu/calcul\\_en.htm](http://www.iqa.mddefp.gouv.qc.ca/contenu/calcul_en.htm)  
<https://www.epa.gov/criteria-air-pollutants/naaqs-table>

Where:

Sub-AQI<sub>i</sub>: the calculated Sub-AQI value for each pollutant;

Pollutant<sub>i</sub>: the monitoring concentration for each pollutant;

Ref-value<sub>i</sub>: the pollutant's reference value, which is the concentration at which the air quality is considered harmful to public's health. For CO, O<sub>3</sub>, PM10, PM2.5, the reference values are 30 ppm, 82 ppb, 150 µg/m<sup>3</sup> and 35 µg/m<sup>3</sup> respectively.

First, the Sub-AQI<sub>i</sub> is calculated for each pollutant on each test day, then the maximum Sub-AQI<sub>i</sub> among the four pollutants is chosen as the AQI for the day. From the calculation method, we can see that the AQI represents the pollution level of the most polluted pollutant on each test day.

During the calculation, we found that on most of the exam dates, the pollutant that is most responsible for AQI in BC and Ontario is O<sub>3</sub> with all other pollutants within lower pollution level.

The model (4-1) is transformed as follows:

$$S_{ijgt} = \beta_0 + \beta_1 AQI_{ijgt} + \beta_2 Temp_{ijgt} + \beta_3 Prec_{ijgt} + \gamma_2 D2_{ijg} + \gamma_3 D3_{ijg} + \dots + \gamma_n Dn_{ijg} + \delta_2 B2_t + \delta_3 B3_t + \dots + \delta_T B T_t + \epsilon_{ijgt} \quad (4-4)$$

Where:

AQI<sub>ijgt</sub>: the integrated air quality index for exam of subject j of grade g, at school i, in year t.

#### 4.4 Logarithmic Transformation

Making natural logarithmic transformation on both dependent variables and independent variables (Log-Log transformation) can make the interpretation straightforward, because the coefficients of the Log-Log model represent the elasticities of the dependent variable with regard to the independent variables. To make the impact straightforward to interpret and to test an alternative non-linear functional form of the relationship, the logarithmic transformation is conducted in this step. The Log-Log transformation model is as follows:

$$\begin{aligned} \text{Ln}(S_{ijgt}) = & \beta_0 + \beta_1 \text{Ln}(AQI_{ijgt}) + \beta_2 \text{Temp}_{ijgt} + \beta_3 \text{Prec}_{ijgt} + \gamma_2 D2_{ijg} + \gamma_3 D3_{ijg} + \dots + \gamma_n Dn_{ijg} \\ & + \delta_2 B2_t + \delta_3 B3_t + \dots + \delta_T B T_t + \epsilon_{ijgt} \end{aligned} \quad (4-5)$$

Where:

$\text{Ln}(AQI_{ijgt})$ : the natural logarithm of AQI.

$\text{Ln}(S_{ijgt})$ : the natural logarithm of  $S_{ijgt}$ .

According to the mathematical theory of Log-Log transformation, the coefficient can be interpreted like this: when AQI increases by 1%, the school test result S approximately increases (if  $\beta_1$  is positive)/decreases (if  $\beta_1$  is negative) by  $\beta_1\%$ . The interpretation in percentage explains the association better and make public easy to understand.

#### 4.5 Non-linear Relationship Exploration

In addition to a logarithmic transformation, to test if there were a non-linear relationship between AQI and school test results, the AQI is broke into several bins and a different coefficient is fit for each bin.<sup>42</sup> The model becomes:

$$S_{ijgt} = \beta_0 + \beta_1(\text{AQI\_cut})_{ijgt} + \beta_2\text{Temp}_{ijgt} + \beta_3\text{Prec}_{ijgt} + \gamma_2D2_{ijg} + \gamma_3D3_{ijg} + \dots + \gamma_nDn_{ijg} \\ + \delta_2B2_t + \delta_3B3_t + \dots + \delta_TBT_t + \epsilon_{ijgt} \quad (4-6)$$

where:

AQI\_cut: a non-parametric estimator, bins of AQI cutting from AQI, such as AQI (0~5),

AQI(5~10).

## CHAPTER 5 RESULTS AND DISCUSSIONS

### 5.1 Fixed-Effects Model With Pollutants as Independent Variables

The summary statics of our final dataset is as follows:

	Mean	Std. Dev	Min.	Max.	Observations
<b>Panel A: BC</b>					
<u>Pollutants</u>					
CO(ppm)	0.25	0.13	0	0.89	32,484
O3(ppb)	17.34	5.4	0	50.47	52,375
PM10( $\mu\text{g}/\text{m}^3$ )	11.33	4.09	2.75	40	19,206
PM2.5( $\mu\text{g}/\text{m}^3$ )	4.67	2.41	0	19	42,859
AQI	10.71	2.78	5	27	10,669
<u>Weather</u>					
Temperature( $^{\circ}\text{C}$ )	10.61	4.61	-26.5	25.7	53,912
Precipitation(mm)	2.16	3.81	0	39.94	53,912
<u>Student's average grade in percent</u>					
All students	66.9	7.74	30	95	17,266
Male students	66.5	7.13	31	94	14,203
Female students	68.17	7.89	33	95	13,579
Students with special needs	59.69	6.46	35	85	4,092
Public school	64.86	7.29	25	95	48,094
Independent school	72.72	8.3	33	95	8,129
<b>Panel B: Ontario</b>					
<u>Pollutants</u>					
CO(ppm)	0.27	0.06	0.2	0.4	996
O3(ppb)	30.82	6.72	13	51	4,288
PM2.5( $\mu\text{g}/\text{m}^3$ )	5.68	3.7	0	18	4,243
AQI	17.57	3.21	11	24	996
<u>Weather</u>					
Temperature( $^{\circ}\text{C}$ )	2.95	4.13	-13.7	12.5	4,215
Precipitation(mm)	4.14	5.65	0	45.3	4,215
<u>Percentage of students who successfully passed the exam</u>					
All students	80.39	13.84	0	100	4,369
Male students	77.39	13.53	0	100	4,206
Female students	84.74	12.46	0	100	4,243

Table 5-1 Summary statistics

As stated in chapter 4, the fixed effects model can control for school fixed effects, municipality fixed effects, grade fixed effects as well as time fixed effects in this study.

<i>Dependent variable: (BC average grade, ON average exam pass rate)</i>									
	BC				ON			BC	ON
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CO	1.195*** (0.449)				5.365 (3.922)			-1.304 (1.296)	1.325 (4.614)
O3		-0.023*** (0.008)				-0.007 (0.024)		0.010 (0.024)	-0.039 (0.087)
PM2.5			0.013 (0.016)				0.069 (0.052)	0.036 (0.056)	0.348 (0.212)
PM10				0.029* (0.015)				0.028 (0.029)	
Observations	9,454	15,355	12,678	5,653	993	4,141	4,106	2,975	993
R <sup>2</sup>	0.870	0.851	0.863	0.866	0.935	0.905	0.907	0.874	0.935
Adjusted R <sup>2</sup>	0.832	0.810	0.820	0.797	0.914	0.889	0.891	0.807	0.914
Residual Std. Error	3.243 (df = 7296)	3.336 (df = 12042)	3.356 (df = 9643)	3.360 (df = 3742)	4.651 (df = 755)	4.662 (df = 3551)	4.613 (df = 3524)	3.259 (df = 1934)	4.648 (df = 753)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01								

Table 5-2 Regressing school test result on pollutants in BC and Ontario

First, we regress the test results on each pollutant (without controlling for the others), controlling for weather effects and fixed effects. In table 5-2, column (1) to (7) show that both in BC and Ontario, the estimated coefficients for O3 are negative, while the estimated coefficients for CO, PM2.5 and PM10 are positive. The reason may be that O3 is the most responsible pollutant for AQI (chapter 4). Secondly, all the pollutants are included in the fixed-effects model: column (8) and (9) show that CO has negative impact in BC but positive

impact in Ontario; whereas, O3 has positive impact in BC but negative impact in Ontario; PM2.5 remains positive in the two provinces. However, all the coefficients are not statistically significant. The coefficients in the two provinces are quite different or even opposite. The reason could be that the independent variables might be correlated with each other and led to multi-collinearity, which generates erratical change in response to small changes in the data or the model.

Consequently, the correlation between the independent variables are examined by Pearson test and the result are shown in Table 5-3:

	BC			ON	
	CO	O3	PM2.5	CO	O3
O3	-0.227 (p-value<2.2e-16)	1	/	-0.177 (p-value=1.9e-08)	1
PM2.5	0.164 (p-value<2.2e-16)	-0.413 (p-value<2.2e-16)	1	0.213 (p-value=1.1e-11)	-0.4 (p-value<2.2e-16)
PM10	0.047 (p-value=3.1e-08)	-0.105 (p-value<2.2e-16)	0.421 (p-value<2.2e-16)	/	/

Table 5-3 Pearson correlation test for pollutants in BC and Ontario

Table 5-3 shows that all the p-values both in BC and Ontario are less than 0.001, so the correlations within the independent variables are extremely significant. CO and particulate matters (PM10 and PM2.5) are positively correlated, this could be because that some CO emission sources also produce large amount of particulate matters, such as forest and grassland fires, fossil fuel combustion, residential wood heating.<sup>48</sup> The positive correlation between PM10 and PM2.5 is due to particulate matter emission source always produces these two pollutants simultaneously. Furthermore, because the formation of ozone needs

sunlight and hot weather, so commonly it has highest concentration in summer and lowest concentration or almost nonexistent in the winter, while CO and particulate matter commonly have highest level in winter since some activities that produces CO and particulate matters happen more often or only happen in winter, such as fossil fuel combustion, residential wood heating. This explains the negative correlation between ozone and the other three pollutants.

Since the independent variables are correlated, I create an integrated index to reflect the synergetic impacts of the four pollutants.

## **5.2 Air Quality Index**

Based on the regression results and correlation results in section 5.1, and to further explore the potential relationship between ambient air pollution and school test results, the AQI is calculated following the function (4-2) and (4-3).

The following figures show the variation of AQI on test days over years. To overview the distribution of the samples better, the points has been “jittered” in figure 5-1 and figure 5-2.

The darker the color is, the more samples there are.

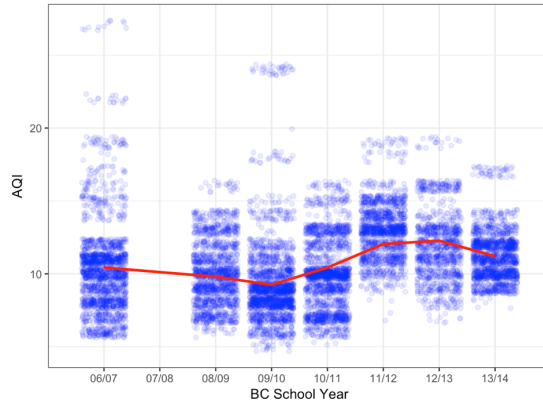


Figure 5-1 AQI on exam days over years in BC

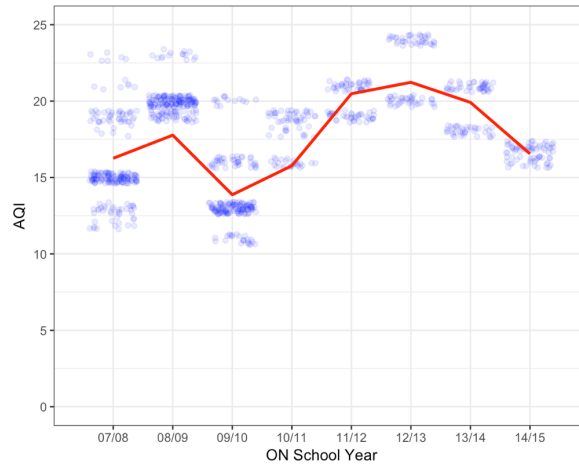


Figure 5-2 AQI on exam days over years in Ontario

Figure 5-1 shows the AQI on BC exam days varied from 5 to 27 with annual average fluctuating from 9 to 13 over years. Figure 5-2 shows the AQI on OSSLT exam days varied from 11 to 24 with annual average fluctuating from 14 to 22 over years. The samples in figure 5-2 are relatively few is because that: a. OSSLT does not have as many schools and subjects as BC provincial test; b. there is lots of missing data for CO in Ontario which led to very few AQI. And this might be a problem when doing regression and will be discussed in the following part.

The result of regression between school test result and AQI is as follows:

<i>Dependent variable:</i>		
	BC average grade (1)	ON average exam pass rate (2)
AQI	-0.065* (0.034)	-0.171 (0.142)
Observations	2,975	993
R <sup>2</sup>	0.874	0.935
R <sup>2</sup> (non-fixed effects)	0.004	0.003
Residual Std. Error	3.257 (df = 1937)	4.652 (df = 755)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

Table 5-4 Regressing school test result on AQI in BC and Ontario

Column (1) of table 5-4 indicates that there is a negative association between AQI and air pollution: with 1 unit increase in AQI, the school test result decreases 0.065 units. And this estimated coefficient is significant. This is comparable with Victor Lavy's study result (2014), which is that with PM<sub>2.5</sub> increasing by 1 unit, the student's test score decreases 0.046 points; with CO increasing by 1 unit, the student's test score decreases by 0.085 points.<sup>19</sup> Moreover, the BC's regression result can also be interpreted as: with one standard deviation increase in AQI, the average grade decreases 0.23 percentage points<sup>+++</sup>.

As column (2) of table 5-4 shown, we also found that AQI has a negative impact on OSSLT test result, although it is not significant. This is might because that the Ontario does not have enough AQI data, and OSSLT does not have as many samples (schools and subjects) as BC provincial test. Moreover, the size of the estimated coefficients in Ontario is larger than

<sup>+++</sup> the estimated coefficient (-0.065) × the standard deviation of AQI (3.5) = -0.23.

BC. The reason for this difference may be that in BC dataset, the test result is recorded as averaged grade, while in Ontario the OSSLT result is recorded as exam pass rate.

### 5.3 Logarithmic Transformation

<i>Dependent variable:</i>		
	Log (BC average grade)	Log (ON average exam pass rate)
	(1)	(2)
Log(AQI)	-0.012** (0.006)	-0.004 (0.046)
Observations	2,975	996
R <sup>2</sup>	0.875	0.916
Adjusted R <sup>2</sup>	0.809	0.890
Residual Std. Error	0.051 (df = 1937)	0.105 (df = 760)
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01

Table 5-5 Log-Log model in BC and Ontario

Column (1) of table 5-5 shows that in BC, there is a negative relationship between Log (BC average grade) and Log(AQI), and the coefficient is significant. According to the Log-Log transformation mathematical theory (section 4.3), the regression result indicates that with 1% increase in the integrated AQI, the school average grade decreases by 0.012%.

Similarly, column (2) of table 5-5 shows that the coefficient of Ontario is also negative, meaning that with 1% increase of the AQI, the OSSLT result decreases by 0.004%. However, the coefficient of Ontario is not significant, as stated in the above, it may be because that we do not have enough observations in Ontario school data, and lots of missing data in air pollution data.

## 5.4 Non-linear relationship Exploration

In addition to a logarithmic transformation, to test if there is a non-linear relationship between AQI and school test results, the AQI is cut into 5 bins in BC and 4 bins in Ontario since the range of AQI in Ontario is narrower than BC. A different coefficient is fitted for each bin. The results are as follows:

	<i>Dependent variable:</i>	
	BC average grade (1)	ON average exam pass rate (2)
AQI (9,13)	-0.277 (0.187)	
AQI (13,18)	-0.213 (0.292)	-0.964 (0.644)
AQI greater than 18	-1.384** (0.661)	-1.173 (0.895)
Observations	2,975	993
R <sup>2</sup>	0.875	0.935
Adjusted R <sup>2</sup>	0.807	0.914
Residual Std. Error	3.257 (df = 1935)	4.653 (df = 754)
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01

Table 5-6 Non-linear relationship in BC and Ontario

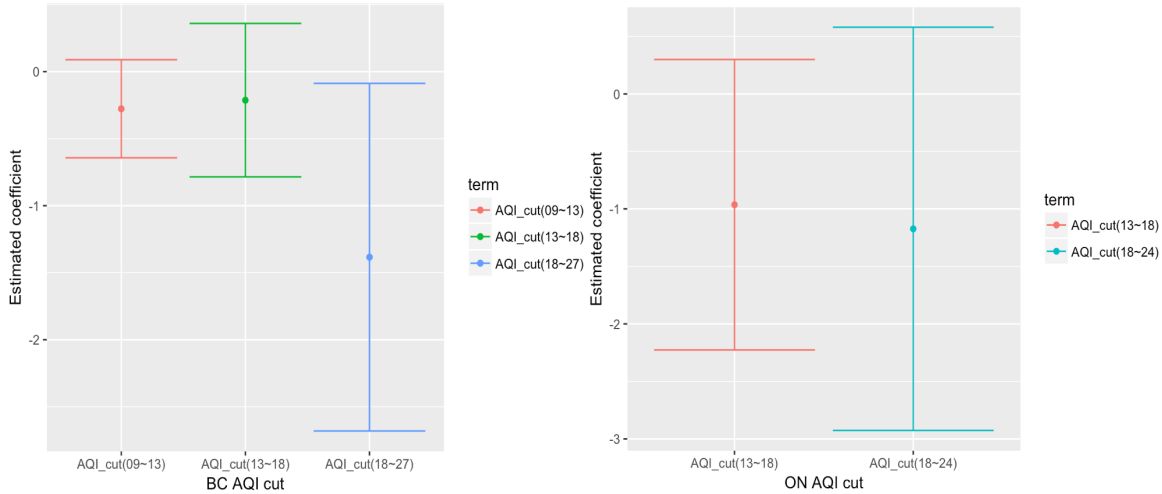


Figure 5-3 Non-linear relationship in BC and Ontario

Table 5-6 and figure 5-3 shows that to avoid “dummy variable trap”###, the (0~9) bin is excluded from the regression. It indicates that the coefficients keep falling with the increase of AQI, indicating that air pollution’s potential negative impact on student’s test results gets worse under poorer air quality.

### 5.5 Robustness Check

To check the robustness in empirical studies, a commonly-used method is adding or removing regressors to examine how the core regression coefficient responses. If the coefficient maintains robust and plausible, it will be considered as an evidence of structural validity. In our study, the core regression is between school test results and AQI, and one and both the weather variables are excluded from the model to check the robustness.<sup>21, 42</sup> In

### When multiple dummy variables are used as regressors, if all n dummy variables are included in the regression along with a constant, the regressors will be perfect multicollinearity, which is called dummy variable trap. So to avoid this, one of the dummy variable is usually excluded from the multiple regression.

addition, we remove the schools with distance to its closest NAPS station larger than 10km, and check how the coefficient changes. The result is as follows:

<i>Dependent variable: (BC Average grade, ON Average exam pass rate)</i>									
	Baseline	No weather control	Distance to NAPS < 50km	Distance to NAPS < 10km	Time linear trend	Inverse distance weight	Fixed effects (a)	Fixed effects (b)	Fixed effects (c)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>Panel A: British Columbia</b>									
AQI	-0.065*	-0.071**	-0.061*	-0.055	-0.085**	-0.091***	-0.031	-0.065*	-0.060
	(0.034)	(0.033)	(0.035)	(0.036)	(0.035)	(0.021)	(0.043)	(0.036)	(0.043)
Observations	2,975	3,136	2,838	2,627	2,975	16,587	2,975	2,975	2,975
R <sup>2</sup>	0.874	0.873	0.872	0.869	0.887	0.852	0.637	0.874	0.911
Residual Std. Error	3.257 (df = 1937)	3.256 (df = 2086)	3.230 (df = 1852)	3.225 (df = 1753)	3.175 (df = 1838)	3.355 (df = 13013)	4.617 (df = 2785)	3.384 (df = 1794)	3.359 (df = 1293)
<b>Panel B: Ontario</b>									
AQI	-0.171	-0.084	-0.171	-0.234	-0.014	-0.131			
	(0.142)	(0.128)	(0.142)	(0.157)	(0.144)	(0.151)			
Observations	993	996	990	942	993	993			
R <sup>2</sup>	0.935	0.934	0.935	0.938	0.959	0.935			
Residual Std. Error	4.652 (df = 755)	4.662 (df = 760)	4.652 (df = 755)	4.639 (df = 716)	4.317 (df = 551)	4.655 (df = 755)			
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01								

Table 5-7 Robustness check of AQI regression in BC and Ontario<sup>§§§§</sup>

In Panel A of table 5-7, column (1) and (2) show that in BC, the coefficient remains similar and significant when excluding weather variable from our model. Therefore, the fixed-

§§§§ The first row is the regression result without dropping any weather controlled variable. The second row is the adjusted regression result with temperature-controlled variable dropped. The third row is the adjusted regression result with precipitation-controlled variable dropped. The fourth row is the adjusted regression result with both temperature and precipitation controlled variables dropped.

effects model used in the study is robust in BC. Column (3) and (4) show that in BC, when only including schools with distance to the closest pollution station less than 50km and 10km respectively, the coefficient remains similar, although not significant for the latter one. To account for slow-moving changes that occur at the school level – for example, changes in income or demographics, a time linear trend interacted with schools is included in column (5), and the coefficient remains negative and similar. To weight the pollution observations by the distances between schools and pollution stations, the inverse distance weighting is conducted in column (6), and the coefficient remain negative and similar. In the baseline scenario, the fixed effects of subject under each school, and year fixed effects are controlled. To test whether the results are sensitive to different fixed effects, I changed which fixed effects are included in the model: school, year and subject fixed effects are controlled in column (7); school-subject (combination of school and subject) and year fixed effects are controlled in column (8); school-subject (combination of school and subject) and school-year (combination of school and year) fixed effects are controlled in column (9). It shows that the coefficient remains negative and similar across all the models, although coefficient in column (7) is a little smaller.

In Panel B of table 5-7, the same robustness check is conducted in Ontario, except for including different fixed effects, which is because that the Ontario school data only have one subject and that school, year fixed effects are already controlled in the baseline scenario. Column (1) to (6) show that in Ontario, the coefficient remains negative but with a larger

interval, and not significant across all the models. The reason could be that there is bias resulted from not having enough observations in Ontario school dataset and lots of missing data in air pollution dataset.

## 5.6 Heterogeneity

To examine the heterogeneity of ambient air pollution’s impact on student’s test result, we examine the different sub-group’s response to ambient air pollution in this step.

<i>Dependent variable: (BC average grade, ON average exam pass rate)</i>								
	BC						ON	
	Baseline	Male	Female	Students with special needs	Public school	Independent school	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AQI	-0.065* (0.034)	-0.075** (0.038)	-0.131*** (0.037)	-0.221** (0.100)	-0.051 (0.037)	-0.115 (0.093)	-0.380** (0.177)	0.068 (0.144)
Observations	2,975	2,523	2,442	763	2,380	595	946	956
R <sup>2</sup>	0.874	0.845	0.883	0.740	0.864	0.847	0.878	0.915
Adjusted R <sup>2</sup>	0.807	0.762	0.819	0.565	0.796	0.739	0.840	0.889
Residual Std. Error	3.257 (df = 1937)	3.291 (df = 1649)	3.182 (df = 1578)	4.223 (df = 456)	3.170 (df = 1579)	3.609 (df = 349)	5.625 (df = 723)	4.577 (df = 730)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01							

Table 5-8 Heterogeneity of sub-groups in BC and Ontario

In table 5-8, column (2) and (3) indicate that air pollution’s impact on female students is 74.7% larger than male students in BC. Column (1) and (4) indicate that the impact on students with special needs is 3.4 times of the average level for other students in BC.

Column (5) and (6) indicate that air pollution's impact on independent schools is about 2 times of public schools in BC, although the two coefficients are not significant.\*\*\*\* Column (7) and (8) show that in Ontario, the coefficient for female is positive but it is not significant; the coefficient for male is negative and significant. This is different from BC, and the reason may be that the school data is average grade in BC, while it is exam pass rate in Ontario.

---

\*\*\*\* The ratios are calculated as: coefficient of female/coefficient of male - 1 =  $(-0.131)/(-0.075) - 1 = 0.747$ ;  
coefficient of students with special needs/coefficient of baseline =  $(-0.221)/(-0.065) = 3.4$ ;  
coefficient of independent schools/coefficient of public schools =  $(-0.115)/(-0.051) = 2$ .

## CHAPTER 6 CONCLUSIONS

Our study is to research whether there is a detectable association between air pollution and student's educational performance, as well as if the association can be taken as a causal impact. BC provincial test results of grade 10, 11 and 12 from 2006/2007 to 2013/2014 and Ontario secondary school literacy test from 2007/2008 to 2014/2015, along with corresponding air pollution and weather data were used to examine this question.

We did not find significant association between student test results and each pollutant independently either in British Columbia or Ontario. This is because that the relationship between school result and pollutants is more complicated, and that the pollutants are significantly correlated with each other and resulted in multicollinearity.

Consequently, an Air Quality Index is calculated to integrate the four pollutants together. In BC, it is found that there is a negative impact of air pollution on student test results, which is an increase of 1% in AQI leads to 0.012% decrease in student's test result. Or, one standard deviation increase in AQI, leads to a 0.23 percentage points decrease in the average grade.

In Ontario, we do not find significant association between OSSLT results and AQI, and this might be because that we do not have enough observations in Ontario school data and there is lots of missing data in the air pollution dataset. However, we find that the association between ambient air pollution and OSSLT results is also negative, which helps to support BC's results. Furthermore, we find that ambient air pollution's impact on students with special needs experience 3.4 times of the average level of other students in BC.

It should be noted that our research has some limits. First, in BC, for each subject exam, one student could take more than once time to participate and the highest score will be recorded, and our data is the school-level aggregated results, so we could not match one unique date to each subject's exam each year. Instead, we used a ratio of weight based on the number of students who participated on each exam date. In Ontario, the sample size of school data is not big enough and there is lots of missing data in the air pollution dataset. Second, the air pollutants' concentrations matched for the exams are the monitoring data of the closest air pollution stations from schools. Therefore, different schools could be matched with the same air pollution data (if the closest air pollution station is the same for them), nevertheless the air pollution levels for them may be different. For example, if one school is close to a busy roadway while the other school is not, the actual air pollution of the two schools could be different even though they share the same closest air pollution station. Third, there is a phenomenon that poorer schools usually locate in the more polluted region, if so, then our research result could be overestimated; however, if wealthier schools locate in the more polluted region (for example, downtown region with higher population density), parent's higher education background and affordable for tutors usually lead to higher scores, and if so, our research result could be underestimated. Fourth, it is possible that there are remaining unobserved factors correlated with both school test results and air pollution, and this can lead to a bias in our study.

In spite of the above limits, our research provides some evidence indicating ambient air pollution has a negative impact on educational outcomes. And the air pollution's impact might be underestimated by only focusing on human health effects.

## BIBLIOGRAPHY

1. Auchincloss AH, Diez Roux AV, Dvornch JT, Brown PL, Barr RG, Daviglius ML, Goff DC, Kaufman JD, O'Neill MS (2008). "Associations between recent exposure to ambient fine particulate matter and blood pressure in the Multi-ethnic Study of Atherosclerosis (MESA)." *Environ Health Perspect.* 116(4):486-91.
2. BC Ministry of Education, Handbook of Procedures 2016-2017.
3. Ben-Porath, Y. (1973) "Labor force participation rates and the supply of labor", *Journal of Political Economy*, 81: 697-704.
4. Brauer, Michael; Reynolds, Conor ; Hystad, Perry. 2013. Traffic-related air pollution and health in Canada. *CMAJ : Canadian Medical Association journal*, Vol.185(18), pp.1557-8
5. Canadian Medical Association. (2008). "No Breathing Room National Illness Costs of Air Pollution", [http://www.cma.ca/multimedia/CMA/Content/Images/Inside\\_cma/Office\\_Public\\_Health/ICAP/CMA\\_ICAP\\_sum\\_e.pdf](http://www.cma.ca/multimedia/CMA/Content/Images/Inside_cma/Office_Public_Health/ICAP/CMA_ICAP_sum_e.pdf)
6. Charles G Plopper and Michelle V Fanucchi. Do urban environmental pollutants exacerbate childhood lung diseases? *Environmental health perspectives*, 108(6):A252, 2000.
7. Curran, Jason, author. "Exposure to traffic-related air pollution and perinatal health" National Collaborating Centre for Environmental Health, issuing body, publisher: Canadian Electronic Library (Firm), distributor. 2014
8. Desvousges, W.H., F.R. Johnson, M.C. Ruby, and A.R. Gable, 1997, "Valuing Stated Preferences for Health Benefits of Improved Air Quality: Results of a Pilot Study", Working Paper T-9702, Triangle Economic Research.
9. Dockery DW, and Pope, CA 3rd (1994). "Acute respiratory effects of particulate air pollution." *Ann Rev Pub Hlth* 15:107-13.
10. Education Quality and Accountability Office 2011, <http://www.eqao.com/en/assessments/OSSLT/assessment-docs/information-bulletin-osslt.pdf>
11. Environment and Climate Change Canada, 2010a, Human Health Costs, [www.ec.gc.ca/Air/default.asp?lang=En&n=085A22B0-1](http://www.ec.gc.ca/Air/default.asp?lang=En&n=085A22B0-1)
12. Graff Zivin, Joshua S. and Matthew J. Neidell. 2011. "The Impact of Pollution on Worker Productivity." 2012. *American Economic Review*. Vol. 102. No. 7. pp. 3652-73.
13. Ghio, A. J., Kim, C., and Devlin, R. B. (2000). "Concentrated ambient air particles induce mild pulmonary inflammation in healthy human volunteers." *Am. J. Respir. Crit. Care Med.* 162: 981-988.
14. Ham, John C., Zweig, Jacqueline S. and Edward Avol. (2011) "Pollution, Test Scores and the Distribution of Academic Achievement: Evidence from California Schools 2002-2008." IZA Working Paper.
15. Hanna, Rema and Paulina Oliva (2011). "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City." 2015. *Journal of Public Economics*. Vol. 122. pp. 68-79.
16. Hsiao Cheng, 2003. "Analysis of Panel Data". Book, page 2.

17. Jeannie Allen , 2002. "Chemistry in the Sunlight"  
[https://earthobservatory.nasa.gov/Features/ChemistrySunlight/chemistry\\_sunlight3.php](https://earthobservatory.nasa.gov/Features/ChemistrySunlight/chemistry_sunlight3.php)
18. Kathleen Mortimer, Romain Neugebauer, Frederick Lurmann, Siana Alcorn, John Balmes, and Ira Tager. Air pollution and pulmonary function in asthmatic children: effects of prenatal and lifetime exposures. *Epidemiology*, 19(4):550-557, 2008
19. Lavy Victor, Avraham Ebenstein and Sefi Roth (2014) "The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation" NBER Working Paper No. 20648
20. Lavy Victor, Avraham Ebenstein and Sefi Roth (2014) "The Long Run Human Capital and Economic Consequences of High-stakes Examinations" NBER Working Paper No. 20647
21. Lu, Xun and Halbert White (2014). "Robustness checks and robustness tests in applied economics." *Journal of Econometrics*, 178(1).
22. McKittrick, Ross, 1965- Fraser Institute (Vancouver, B.C.) 2006. "Air pollution policy in Canada improving on success".
23. Neidell, Matthew. 2004. "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *Journal of Health Economics*, 23: 1209–1236.
24. OECD. (2014). "The Cost of Air Pollution - Health Impacts of Road Transport". Published on May 21, 2014.
25. OECD, "DEFINING AND MEASURING PRODUCTIVITY".  
<https://www.oecd.org/std/productivity-stats/40526851.pdf>
26. Oliva, P., and R. Hanna. 2011. "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City." NBER Working Paper No. 17302.
27. Ontario Medical Association, 2005, "The illness cost from air pollution: 2005-2026".
28. Ontario Medical Association, The Illness Costs of Air Pollution in Ontario. 2005.  
<https://www.oma.org/Resources/Documents/e2005HealthAndEconomicDamageEstimates.pdf>
29. Ostro, Bart (1983). "The Effects of Air Pollution on Work Lost and Morbidity." *Journal of Environmental Economics and Management*, 10(4): 371-382.
30. Ozkaynak, H., Xue, J., Spengler, J., Wallace, L., Pellizzari, E., & Jenkins, P. (1996). "Personal exposure to airborne particles and metals: results from the Particle TEAM study in Riverside, California." *Journal of Exposure Analysis and Environmental Epidemiology*, 6(1), 57-78
31. Pandey M.D. and Jatin S. Nathwani, 2003, "Canada Wide Standards for Particulate Matter and Ozone: Cost-Benefit Analysis Using a Life Quality Index" in *Risk Analysis*, Vol. 23, No. 1, 2003.
32. Piantadosi, C. A. 2002. "Carbon Monoxide Poisoning." *The New England Journal of Medicine*, 347(14):1054-1055.
33. Pope, CA 3rd (2000). "Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk?" *Environ Health Perspect.* 108(4):713-23.
34. Prashant Bharadwaj, Matthew Gibson, Joshua Graff Zivin, Christopher A. Neilson (2017) "Gray Matters: Fetal Pollution Exposure and Human Capital Formation" *Journal of the Association of Environmental and Resource Economists*, Vol. 4, No. 2.
35. Real-time World Air Quality Index, <https://waqi.info/>

36. Regulations Amending the Vessel Pollution and Dangerous Chemicals Regulations  
<http://gazette.gc.ca/rp-pr/p1/2012/2012-07-21/html/reg2-eng.html>
37. Sam Brockmeyer, Amedeo D'Angiulli. (2016) "How air pollution alters brain development: the role of neuroinflammation"  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5017593/>
38. Sawyer, D., S. Stiebert and C. Welburn, 2007, Evaluation of Total Cost of Air Pollution Due to Transportation in Canada, Marbek Resource Consultants Ltd and RWDI Inc.  
[http://publications.gc.ca/collections/collection\\_2008/tc/T22-148-2007E.pdf](http://publications.gc.ca/collections/collection_2008/tc/T22-148-2007E.pdf)
39. Seaton, A., MacNee, W., Donaldson, K., and Godden, D. (1995). "Particulate air pollution and acute health effects." *Lancet* (8943): 176-178.
40. Stock J, Watson MW. Introduction to Econometrics. New York: Prentice Hall; 2003.
41. Teng Li, Haoming Liu and Alberto Salvo (2015) "Severe Air Pollution and Labor Productivity" IZA DP No. 8916
42. Tom Chang, Joshua S. Graff Zivin, Tal Gross, Matthew J. Neidell (2014) "Particulate Pollution and the Productivity of Pear Packers" 2016. *American Economic Journal: Economic Policy*. Vol.8, No.3. pp. 141-69.
43. Thatcher, T. L., & Layton, D. W. (1995). "Deposition, resuspension, and penetration of particles within a residence." *Atmospheric Environment*, 29(13), 1487-1497.
44. United States Environmental Protection Agency, "Health Effects of Ozone Pollution".  
<https://www.epa.gov/ozone-pollution/health-effects-ozone-pollution>
45. Vette, A. F., Rea, A. W., Lawless, P. A., Rodes, C. E., Evans, G., Highsmith, V. R., & Sheldon, L. (2001). "Characterization of indoor-outdoor aerosol concentration relationships during the Fresno PM exposure studies." *Aerosol Science & Technology*, 34(1), 118-126.
46. World Health Organization, 2016, [www.who.int/mediacentre/news/releases/2016/air-pollution-rising/en/](http://www.who.int/mediacentre/news/releases/2016/air-pollution-rising/en/)
47. William H. Greene, *Econometric Analysis*, Pearson Prentice Hall, 7th edition, 2011. ISBN-13: 978-0131395381 ISBN-10: 0131395386
48. Environment and Climate Change Canada, Pollution Sources,  
<https://www.ec.gc.ca/air/default.asp?lang=En&n=F963E49C-1>
49. Environment and Climate Change Canada, National Air Pollution Surveillance Program (NAPS). <https://www.ec.gc.ca/rnspa-naps/>
50. Environment and Climate Change Canada, weather historical data.  
[http://climate.weather.gc.ca/historical\\_data/search\\_historic\\_data\\_e.html](http://climate.weather.gc.ca/historical_data/search_historic_data_e.html)
51. Does Classroom Temp. Effect Student Performance?  
<https://sites.psu.edu/siowfa16/2016/10/05/does-classroom-temp-effect-student-performance/>
52. Does bad weather affect student performance in school? <http://theconversation.com/does-bad-weather-affect-student-performance-in-school-75461>