UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Economics

The Health Impact of Environmental and Health Policies in Mining Districts: Evidence from Mexico

John Andersson

Supervisor: Ylenia Brilli
Master Degree Project No. 2016:113
Graduate School

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Economics

The Health Impact of Environmental and Health Policies in Mining Districts: Evidence from Mexico

Abstract

In this study I investigate the health impact from an environmental policy implemented in northern Mexico. I use panel data from MxFLS (Mexican Family Life Survey) that is collected on an individual level. The individuals are all over 15 years old and are analysed in two or three different periods (2002, 2005-2006 and 2009-2012). Health is measured in a subjective form where individuals rank their current health status from 1 (*Very Poor*) to 5 (*Very Good*). The impact of the policy is measured by using an interaction term that interacts two dummy variables representing the area where the policy was implemented (city of Torreon) and the post period (2005 to 2012). The results indicate that the policy has been successful on a subjective level. However, these results were not supported by the more objective measures of health, namely self-reported illnesses, hospitalizations and doctor visits.

2016-09-11

Author:
John Andersson

Supervisor:
Ylenia Brilli

Table of Contents

1. Introduction

Environmental pollution is one of the largest causes of diseases and deaths in low- and middle-income countries. An investigation made by Global Alliance on Health and Pollution (GAHP) show that in 2012 approximately 8.4 million people died in the developing world due to pollution in air (both indoor and outdoor), soil and water (Petru, 2015). To understand this number in relative terms, the same year around 3 million people died worldwide from malaria, HIV/AIDS and tuberculosis. Thus the contamination exposure should be considered one of the greatest threats against public health in the developing world. Heavy metal contamination in soil and food from industries caused 10% of the deaths and 8% were accounted by pollution in the local water systems by the form of sewage and industrial waste (GAHP, 2014). In other words, approximately 1.5 million people die annually from soil and water contamination. Environmental and health policies are therefore very important and an urgent matter in mining communities, where soil and water pollution is a common threat for the local inhabitants.

In this study I will investigate the impact of an environmental policy on health status. My area of interest will be Torreon, located in the state of Coahuila in northern Mexico, that for more than a century have suffered from heavy metal exposure. In the middle of the city a smelter is located and as a consequence, Torreon is highly contaminated by lead, cadmium and arsenic. In 1999 the government of Coahuila implemented an environmental policy to improve the poor health status in the region. Some important steps of the new policy included treating children with too high heavy metal levels in their blood, moving the houses of families living dangerously close to the mine and building a new park to improve air quality. The policy also provided the mining industry to decrease the emissions by 25% during one year (1999-2000). But after only one year of decreased emissions, the contamination could continue at the same rate as before the policy was implemented. To my knowledge there are no studies on the subject to provide evidence whether this form of policy is successful or not. I find this investigation highly relevant since it can be used to guide future policy implementations with the objective to improve public health status.

The economical literature on policies against contamination is relatively wide; however, these studies tend to focus only on air pollution in developed countries (Graff Zivin & Nedell, 2013). This is due to scarce supply of data on water pollution and toxins, where the data on air pollution is much more widely available. For this reason there are relatively few investigations focusing on the impact of policy implementations in mining districts, which I find surprising due to the high number of deaths it causes in the developing world. This study will fulfil this gap in the literature. Moreover, previous studies in the same area have focused on children as they are more sensitive to heavy metal pollution and therefore provide more dramatic variation in the blood levels. In this study however, I analyse individuals older than 15 years as I am interested to see how adults' health are affected by a change in pollution exposure. By using data on a micro level from MxFLS (Mexican Family Life Survey) [1], I analyse the change in health, socio-economic and socio-demographic factors over three different periods (2002, 2005-2006 and 2009-2012). The outcome variable in the analysis is self-reported health, i.e. subjective health status. From pooled- and random effects ordered probit models I estimate the change over time in self-reported health but also variables that may have an impact on health status. I estimate the effect of the policy by comparing the outcomes in health status in Torreon during two different periods, namely a pre period (2002) and a post period (2005-2012). I also compare the results from Torreon with its neighbouring city, Gomez Palacio, which is used as control group in this study. Gomez Palacio is located only 9 km from Torreon, and is therefore also affected by the pollution in Torreon (Méndez-Gómez et al., 2008). However, no environmental policy has been implemented in Gomez Palacio as it belongs to a different state[2].

In terms of self-reported health, the results indicate that the policy implementation has been successful as people living in Torreon were more likely to report *Very Good* health status in the post period in comparison with the pre period and Gomez Palacio. However, the findings from the self-reported health variable are not supported by the results when estimating more objective measurements for health status, namely self-reported illnesses, hospitalizations and doctor visits. Actually, the self-reported illnesses increased in the post period in Torreon in comparison with Gomez Palacio. Thus the results are quite confusing regarding this matter and this will be discussed in the conclusion section.

This paper is structured in the following order: Sector 2 presents the literature review that describes results from previous papers on regarding environmental policies and health. Sector 3 is the background section where the issue, region and policy are described. Sector 4 demonstrates the theoretical analysis and sector 5 the data description. The methodology is

---

[1] Rubalcava, L., Teruel, G. (2013). Mexican Family Life Survey, Third Round, Working Paper, www.ennvih-mxfls.org.

[2] Gomez Palacio belongs to the state of Durango and Torreon (where the policy was implemented) the state of Coahuila.

presented in sector 6 followed by the results in sector 7 and heterogeneous results in sector 8. Finally, sector 9 discusses the conclusions and recommendations for future research.

2. Literature review

There are several ways to measure general health status for an individual. Medication costs, mortalities and hospitalizations are some of the most common measurements of health and dominate the academic literature on policy effects. The literature on the impact of policy implementations on self-reported health is, however, much more limited. To my knowledge, there are actually no studies that analyze the possible change in self-reported health from an implementation of new regulations aimed at decreasing pollution. Providing evidence and knowledge in this scarce area will be one of my greater contributions. I believe that self-reported health is a good option to measure general health status when a health policy is relatively new. I mean that regarding more objective measures of health (hospitalizations, mortalities etc.) it might take longer time for the environmental policy to effect the health status while self-reported health could vary more in a short period and is therefore less dependent on time. I also believe that self-reported health could capture other aspects of well being that the more objective measures would not be able to seize. For instance, by using hospitalizations as dependent variable you would capture individuals that have visited the hospital recently, however, it could be the case that these hospitalized individuals overall feel good about their health. Maybe the visit to the hospital was due to an incident that had nothing to do with that person's former health status, for example car accident. Thus the subjective well being for an individual would not be included in the analysis, which would be a great drawback. Therefore more knowledge is needed in this area and this is why I find this paper important for future research.

Mortality is the most common measurement when estimating the impact of a change in emissions on health status. This literature is also dominated by infant mortalities as it allows us to observe the immediate effect of a change in pollution (Currie & Neidell, 2005). In this paper, the authors find that the reductions in carbon monoxide, due to the Clean Air Act Amendments implemented in the U.S., saved approximately 1,000 infant lives in California in the 1990's. Greenstone & Hanna (2013) provide similar results from India where two different environmental policies were implemented in 1995 and 1996. These policies covered several different actions, for instance, installing catalytic converters into certain vehicles, building new roads and restricting industrial pollution. From data between the periods 1987-

2007, they conclude that between the period 1987-1990 there was a decrease in particle pollution by 19% and sulfur dioxide by 69% nationwide, and as a result approximately 86 infant lives were saved per 100,000 births.

The impact on adults is more difficult to investigate as they are less sensitive to contamination and therefore the impact is not as immediate as on children. Chay et al. (2003) provide evidence on this matter when investigating the impact of the Clean Air Act Amendments implemented in 1970 in the U.S. From investigating pollution levels and mortalities between 1969 and 1974, they observe no change in mortalities despite a dramatic decrease in air pollution levels the first year. They discuss however that these findings could be due to the short span of years, which do not allow us to observe changes in the long run. Dechênes et al. (2012) found the opposite results when investigating the budget program on nitrogen oxides implemented in 1999 in the U.S. From observations over a ten-year period (1997-2007) they find that the number of deaths decreased by approximately 2,175 cases annually, where 35-56% of the decline was explained by reductions in cardiovascular/respiratory deaths. Differently from other studies, they did not restrict their analysis to only mortalities but also hospitalizations and medical use. The results demonstrated that the hospitalizations decreased by approximately 1% and the medication costs by around 1.9%. Depending on choice of health variable the results can vary widely and Graff Zivin & Neidell (2013) discuss that it is advantageous to construct an investigation like the one by Dechênes et al. (2012). A wider toolbox of health variables brings greater certainty on the true health impact from a policy implementation. For that reason, in my analysis I also consider measures of health such as self-reported illnesses, hospitalizations and doctor visits, and discuss the relationship between self-reported health and mortality.

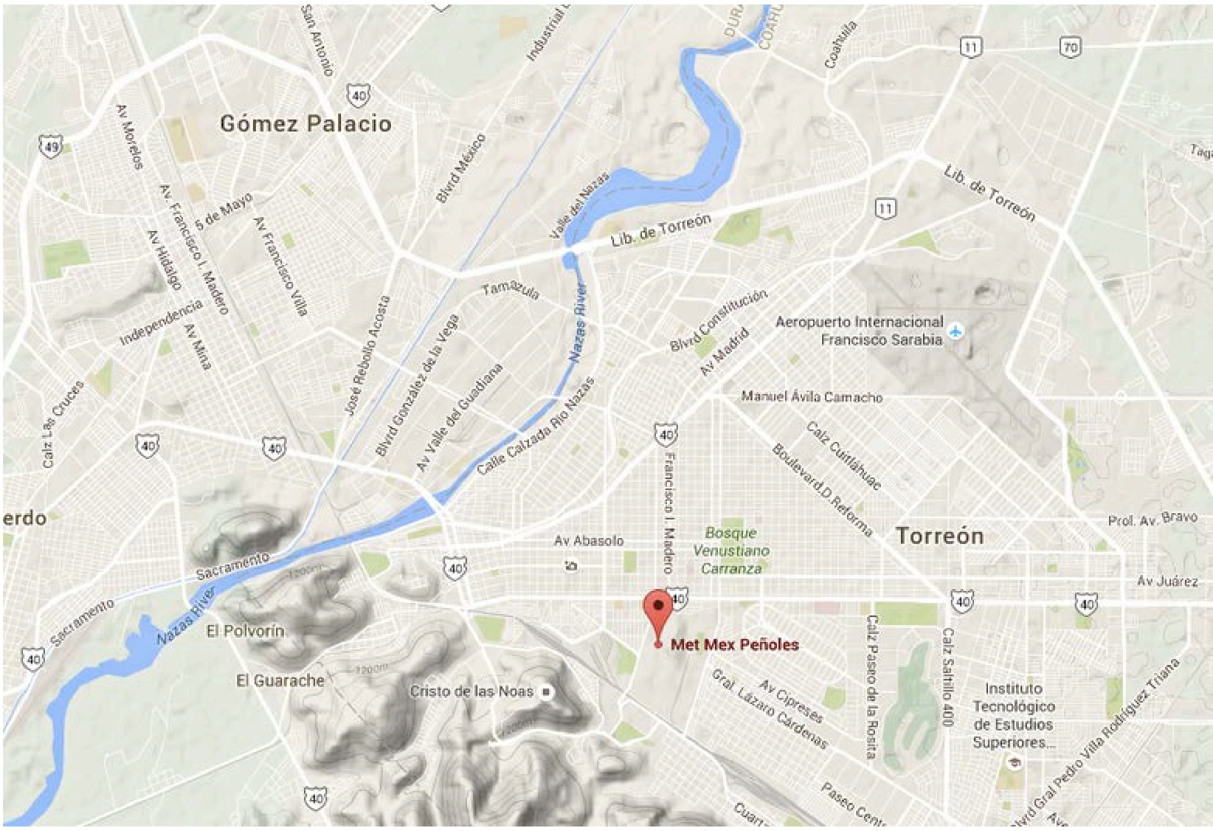
3. Background

In the last 20 years there has been a vast explosion in Mexico concerning investments in the mining industry due to a mining law that was implemented in 1992. This law permits foreign investors to own up to 100% of the Mexican mines and as a result, more and more companies have entered the Mexican mining industry. Thus the production of different metals has increased vastly, for instance, between 2001-2012 the government accepted concessions of mining exploitation in Mexico representing 61.8 million of acres (Calderón Mendoza & Jiménez, 2015). As a consequence, the contamination (caused by the mining industry) around the nation has increased which could be severely dangerous for the human health. Different heavy metals have been observed at hazardously high levels in children's hair, urine and blood. In this study I investigate the health conditions in Torreon, Coahuila. Torreon is one of the most traditional mining cities in Mexico with more than 100 years of mining extraction.

3.1 Lead, cadmium and arsenic in Torreon

Peñoles Met-Mex is a Mexican mining company that have been extracting minerals in the city of Torreon, in northern Mexico, for decades. From 1901 to 2000, Peñoles Met Mex was producing without limits, contributing to a layer of black dust in the area. This was not considered to be a public problem up-until the 1970's when the Mexican government sold the land nearby the hill to the people. Thus the inhabitants started living in the dust and in the 1990's the concerns were growing about the pollution and its effects on human health (The Economist, 2011). In figure 1 we observe a map of the district, where the Peñoles smelter is located in the southern part of Torreon, surrounded by houses. Méndez-Gómez et al. (2008) show that not only the children in Torreon are negatively affected by the pollution of Peñoles Met-Mex, but also the children in Gomez Palacio. Considering the short distance, where only the river "Nazas" separates the two cities, these results are not very surprising.

Figure 1 – Map of the Laguna district



3.2 Policy implementation

After a study made by University of Juárez in the state of Durango, in 1999 the authorities implemented some new regulations for Peñoles Met-Mex. Thus twelve new directives were made to decrease the pollution and improve the environmental and health situation in the region. These directives forced Peñoles Met-Mex to invest large amounts in health care for contaminated children, cleaning the city, moving families living close to the mine, building parks to clean the air, among other rules (Valdéz Perezgasga & Cabrera Morelos, 1999). One of the most important actions in this program was to decrease all kinds of emissions. Big warehouses, a total of 45,000 m^2, were constructed for the purpose of storing material that will be re-used and raw goods instead of letting it stand out in the open as these polluted materials easily could be carried with the wind into the city. Moreover, installations of electrostatic precipitators and water cleaning plats were made. An environmental management system called ISO 14000 was also implemented to increase the workers knowledge of contamination prevention. A second action was to take samples from all over Torreon to gain knowledge of which areas were the most affected by heavy metal pollution. In the most affected areas, cleaning programs were applied and a total of 360,000 m^2 were paved with cement to cover the metal exposure. Moreover, in 2000 more than hundreds of families were moved from the most polluted area, called colonia Luís Echeverría, to cleaner neighbourhoods. A total of 448 houses were destroyed in this area and in 2001 the park "El Centenerio" was built to recover the air quality in Torreon. 9,600 trees were planted in a zone of 18 hectares and cactuses were planted in an area of 5,000 m^2. Community health centres were also installed in the city where the citizens of Torreon could help from doctors with toxicology as expertise. Furthermore, a campaign was made to give the citizens guidance in hygiene and in aspect regarding food (Chaparro, 2007).

According to Recio-Vega et al. (2012), the large amount of money invested has had a positive effect on children's health. By analyzing treated children, i.e. the most intoxicated children who received medical treatment as part of the program, they find evidence that the blood lead levels of the treated children of this program had decreased by 69.75% between 1999 and 2010. Another interesting result from this paper was that the biggest decline in blood lead levels was observed in the first three years, i.e. from 1999 to 2002. Nevertheless, Garcia-Vargas et al. (2014) discuss that the blood lead levels in children and teenagers are still too high. They find results showing that over 86% of teenagers in lead hot spots exceeded the reference level settled by Centers for Disease Control and Prevention (5mg/dl). The major part of these hot spots, i.e. where the highest levels of lead were observed, were detected in the western and northwestern part of the city. Rosado et al. (2007) agrees with Garcia-Vargas that the exposure of heavy metals is still too high. They find that the levels of arsenic in the drinking water were moderately high throughout the city between 2006 and 2007 where 52% of the children investigated had concentrations of arsenic (in their urine) above 50 μg/L and 10% were above 100 μg/L.

I have chosen to investigate this area due to three different reasons. First, the environmental policy in Torreon is quite different from other regulations since the emissions only decreased over a one-year period. Thus it seems like the aim is to create a win-win situation where Peñoles Met-Mex can continue to pollute at the same pace as previously but compensate the inhabitants by cleaning the city and treating the polluted children. This makes the study much more interesting as there is a gap in previous literature on policies that are not principally aimed at decreasing the emissions. Second, except for Torreon I have only heard of a few policy implementations in Mexico related to mining, namely in the state of Zacatecas, however these were implemented in 2011 and therefore are “out of reach” for my dataset (as it ends in the period 2009-2012). Third, that I can use Gomez Palacio as a control group. The investigation of Méndez-Gómez et al. (2008) provides evidence that Gomez Palacio also is heavily affected by the pollution in Torreon and is thereby possibly a good control group. I believe it would be very hard to find better control groups in other cities or communities more suited for this study.

4. Theoretical framework

The health production function is widely used throughout the health economics literature. It demonstrates how different inputs, both medical and non-medical, can affect the health outcome. It provides understanding of how different inputs interact and how they may impact overall health status (Wibowo & Tisdell, 1992). Graff Zivin & Neidell (2013) construct such a model where health (H) depends on pollution exposure (P), medical expenditures (M) and avoidance behaviour (A):

$$H = H(P, M, A)$$

where avoidance behaviour represents activities committed before being exposed to pollution to decrease its damage while medical spending is a post pollution action where the damage has already been done on health capital. Both the avoidance behaviour and medical expenditures have a positive impact on health while pollution exposure obviously is negatively related to health status:

$$\frac{\partial H}{\partial P} < 0, \qquad \frac{\partial H}{\partial M} > 0, \qquad \frac{\partial H}{\partial A} > 0$$

The policy implemented in Torreon contains several different steps discussed in the background and these actions can be applicable in the function above. For example, the decrease in the pollution exposure, as a result of the cleaner streets and the newly constructed park, should have a positive impact on health. Note that we do not know for sure that the pollution exposure has decreased, however, theoretically this action should have a positive impact on the inhabitants Torreon. Moreover, moving families living dangerously close to the mine is an example of avoidance behaviour since these families no longer will be exposed to metal contamination. In this case though, the damage has already been done, nevertheless the emigration from the heavily exposed area ought to have a positive impact on health status. Finally, regarding medical expenditures the policy does not have a direct impact on my sample as only children were treated for metal poisoning.

5 Data description

5.1 Mexican Family Life Survey

In this investigation I use data from one single source, the Mexican Family Life Survey (MxFLS), constructed by Iberoamerican University (UIA), Center for Economic Research and Teaching (CIDE) and Duke University. The MxFLS is collected from the entire country and contains longitudinal data at an individual level over three different periods of time, 2002, 2005-2006 and 2009-2012. However, due to a low sample size in the second period (2005-2006), I merge the observations from this period with the third period. Thus a new variable is constructed, “Post” that represents the observations between 2005-2012. In this study I only investigate the outcomes from adults (15 years and older), where I also drop individuals from my sample that for some reason moved during the survey process, as these individuals would not contribute to this study. The individuals that were not born and raised in the cities where they currently live were also dropped to get rid of the possible issue of endogenous migration.

From the handpicked selection of individuals, the dataset contains a total of 297 adults, 203 from Torreon and 94 from Gomez Palacio. The observations from these individuals are weighed against the entire population thanks to expansion weights offered by MxFLS. These weights are collected from the last period (2009-2012) on an individual level. Moreover the weights are constructed in a longitudinal structure, which means that the sample is expanded by the weights to the original population in the first period (2002). The sample weights are needed to control for unequal probabilities of selection, to make the sample conform to a known population distribution with respect to key variables (such as age, sex etc.) and finally to reduce bias due to non-responses (Yansaneh, 2003). To be more precise, the weight is the inverse of the probability of being included in the sample based on the sample design (IDRE, 2016).

The MxFLS contains detailed information concerning health, socio-demographic and socio-economic characteristics. Moreover, the individuals are divided into different states, municipalities and even localities. In 2002 the complete dataset contained observations from 19,764 individuals. However, by taking advantage of the categorical division of the individuals, I construct my dataset on observations from the Laguna district, i.e. the localities of Torreon and Gomez Palacio. Torreon and Gomez Palacio are situated in two different states, Coahuila and Durango, and therefore the regulations vary between the two cities. Despite the fact that Gomez Palacio is affected by the pollution from Peñoles Met-Mex due to the short distance between the two cities (9 km), no environmental policy has been implemented there. Therefore Gomez Palacio will be the control group in this analysis.

5.2 Variables

I analyse the possible change in health status by using self-reported health where the individuals answer the question "*Currently, could you say your health is (...)?*" where the options to choose from are: 1. *Very poor*, 2. *Poor*, 3. *Regular*, 4. *Good* and 5. *Very good*. The self-reported health fluctuates less than illnesses and therefore is a good measure of the health capital level. However, this could also be a disadvantage since people sometimes underestimate (or overestimate) their health status and therefore the self-reported health does not represent the true health of some individuals. These issues will be further discussed in the robustness section below.

In the majority of the previous studies in the Laguna district, discussed in the background section, the investigations have not constructed econometric regressions controlling for different socio-economic factors. Nevertheless, these are important aspects for the analysis to get a better understanding on the elements that may have an impact on health. To control for these socio-economic effects, I include variables for education and employment. Regarding education I include the variable, "higher education level", which refers to studies above the mandatory level (secondary school). Employment is based on the job situation, i.e. if the individual is currently working or not. To investigate the impact of socio-demographic factors I include variables for age, gender and if individual is married. To understand how other health related features might have an impact on general health status, I include the variables controlling for health insurance, smoking and exercising.

In table 1 we observe the summary of the dependent and independent variables. In the table I have also included "illness", "hospitalized" and "doctor visits" despite the fact that they are not used in my model. These variables are used in the robustness section to compare the results from self-reported health with more objective measurements of health status. Illness refers to if individual has reported any illness, by answering *Yes* on the question "*In the last 4 weeks, have you suffered of (…)?*"[3]. Doctor visits refer to short visits at hospitals, clinics, health centre etc. or if a doctor has visited this person at his/her home.

Table 1
Summary variables

Variables	Description
Self-reported health	Subjective health of individual ranked as: 1=Very poor, 2=Poor, 3=Regular, 4=Good & 5=Very good
Insuranced	Individual has health inurance: 1=yes & 0=no
Exercises	Individual exercises: 1=yes & 0=no
Smoking	Individual smokes: 1=yes & 0=no
Age	Age of individual (continous)
Married	Individual is married: 1=yes & 0=no
Female	Individual is female: 1=yes & 0=no
Torreon	Individual lives in Torreon: 1=yes & 0=no
Working	Individual is currently working: 1=yes & 0=no
Higher level education	Individual has a higher education level than mandatory (secondary school): 1=yes & 0=no
Post	Dummy variable representing observations in the post period: 1=Post & 0=Pre
Illness	Self-reported illness by individual (fever, breathing problems, chest pain etc.) last 4 weeks: 1=yes & 0=no
Hospitalized	Individual has been hospitalized last year: 1=yes & 0=no
Doctor visits	Individual has visited hospital, clinic etc. last 4 weeks

5.3 Descriptive statistics

In table 2 we observe the summary statistics of the independent variables in the Laguna district. We can see that only 39.8% of the adults have an education level that is higher than the mandatory (secondary school). Moreover, a big proportion of the individuals in the sample have reported that they have suffered from an illness in the last 4 weeks. However, few of them have sought help in hospitals or clinics, despite the fact that the majority of them have health insurance. Table 3 show that the majority of the people in the Laguna district

[3] The illnesses included in the survey are: breathing issues, chest pain, body ache, headache, cough, diarrhea (at least three times per day), irritated or itching skin, irritated eyes, fever, flu, strong stomach pain, nausea/vomit, swollen joints, teeth pain, fever, throat pain, respiratory/urinary/digestive problems, blood pressure and finally other possible illnesses.

would rank their health as *Good*. Matter of fact, approximately 89% would rank their health as *Regular* or *Good*. Noteworthy is also that the proportion of people estimating their health as *Very good* is relatively big, namely 8.34%.

Table 2
Descriptive statistics

Variables	Number of observations	Mean	Min	Max
Age	708	37.028	15	72
Higher level education	740	0.398	0	1
Female	740	0.582	0	1
Working	740	0.590	0	1
Married	734	0.589	0	1
Torreon	740	0.620	0	1
Post	740	0.590	0	1
Insuranced	740	0.755	0	1
Exercises	740	0.264	0	1
Smoking	740	0.198	0	1
Illness	740	0.574	0	1
Hospitalized	740	0.078	0	1
Doctor visits	740	0.131	0	1

Table 3
Proportion of Rank

Self-reported health	Proportion
Very good	8.34%
Good	55.9%
Regular	33.18%
Bad	2.48%
Very bad	0.1%

In table 4 I demonstrate the average values for my independent variables for both Torreon and Gomez Palacio in the pre period. I also compare the values in the two cities by constructing t-tests. As we can see, there are only two independent variables that are significantly and statistically different between the two cities, namely the "Age" and "Married".

Table 4
Average Values in Pre Period

Variables	Torreon Pre-period	Gomez Palacio Pre-period	t-test
Insuranced	0.705	0.648	-0.749
Exercises	0.411	0.314	-1.271
Smoking	0.201	0.21	0.393
Age	34.943	32.01	-2.006**
Married	0.629	0.519	-1.870*
Female	0.585	0.562	-0.519
Working	0.513	0.552	0.714
Higher level education	0.335	0.314	-0.523

6. Methodology

6.1 Pooled Ordered Probit Model

In this study, I analyse the impact of the policy implementation on self-reported health, which is ordinal. I therefore construct an ordered probit model, where I estimate the unobserved self-reported health through the observed values in the dataset. The ordered probit model is a standard measurement when dealing with a dependent variable with ordinal nature, and seems to dominate the literature on self-reported health status. The model is constructed as a maximum likelihood estimation of the different ordinal results, i.e. probability that an individual reports *Very Good* health status based on the values observed from the data. The maximum likelihood model is preferred over an OLS since the dependent variable is ordinal and therefore non-linear. Non-linear variables present different error structure and their variances are not constant (Min, 2013). An OLS model would therefore provide biased and inconsistent results when estimating self-reported health due to its nature.

As the dependent variable is ordinal, I must take into consideration that there are unobserved factors that impact the ranking of each individual. Thus I estimate a latent model where y_{it}^{c*} is the unobserved variable, so called latent variable of the true health:

$$y_{it}^{c*} = \beta_0 + \beta_1 Torreon + \beta_2 Post + \beta_3 (Torreon * Post) + \beta_4' x_{it}^{c} + \varepsilon_{it}^{c}$$

and x_{it} represents the explanatory variables: “Insurance”, “Exercise”, “Smoking”, “Age”, “Age squared”, “Married”, “Female”, “Working” and “Higher level education”. ε_{it} is the error term, i represent individual and finally c the city. “Torreon” is dummy variable representing the individuals in the sample living in Torreon, i.e. 0 if the person is from Gomez Palacio and 1 if he/she is from Torreon. "Post” is also a dummy that is 0 if the individual answered the survey in the pre period and 1 if post period. Together, these two variables construct the main variable of interest in this study, namely the interaction term. This term represents the policy implementation and therefore a positive β_3 would indicate that the environmental policy has been implemented successfully. This would indicate that people in Torreon are more likely to report *Very good* self-reported health in the post period than in the pre period in comparison with Gomez Palacio. Gomez Palacio, as a control group, work as a reference point indicating the self-reported health levels that would be observed in Torreon if it were not for the policy implementation. The interaction term is a difference-in-differences specification and is defined as the difference in average outcome in the treatment group before and after the treatment minus the difference in average outcome in the control group before and after the treatment (Albouy, 2004):

$$\beta_3 = \bar{y}_1^T - \bar{y}_0^T - (\bar{y}_1^{GP} - \bar{y}_0^{GP})$$

In other words, the difference estimator estimates the outcome of the actual $\bar{y}_1^T$ with the outcome that would occur without the policy implementation. Therefore it is very important to use a control group that is representative for the treatment group; however, this will be further discussed in the assumption section below.

As self-reported health is ranked from 1 to 5, where 1 is very poor and 5 very good, the estimation of our unobserved dependent variable will be determined by the observations from this variable. The estimation will be structured as following,

$$y_{it} = 1\ if\ y^* \geq \mu_1$$

$$y_{it} = 2\ if\ \mu_1 < y^* \leq \mu_2$$

$$\dots$$

$$y_{it} = 5\ if\ \mu_4 < y^* \leq \mu_5$$

where μ are unknown threshold parameters to be estimated with β. Moreover, since we use a probit structured model, we assume for the error terms to be normally distributed and therefore normalize the term by setting its mean at 0 and variance at 1 (Greene, 2012). Based on these assumptions, we estimate the probabilities that an individual crosses a threshold and change the value of self-reported health, where the probabilities are structured in the following form:

$$Prob(y_{it} = 1|x) = \Phi(\mu_1 - x'_{it}\beta)$$

$$Prob(y_{it} = 2|x) = \Phi(\mu_2 - x'_{it}\beta) - \Phi(\mu_1 - x'_{it}\beta)$$

$$\dots$$

$$Prob(y_{it} = 5|x) = 1 - \Phi(\mu_4 - x'_{it}\beta)$$

where the Φ represent the probability density function of a standard normal distribution. These probabilities help us estimate the likelihood function that is,

$$L(\beta', \mu_1, \mu_2, \mu_3, \mu_4) = Prob(y_{it} = 1|x) * Prob(y_{it} = 2|x) * \dots * Prob(y_{it} = 5|x)$$

As demonstrated in the function, the probabilities are dependent on the threshold variables and the parameters. Moreover the parameters are obtained through maximizing the log of this likelihood function.

6.2 Random Effects Ordered Probit Model

As the name indicates, the pooled ordered probit pools together the data on the different individuals and therefore ignores individual differences, so called random effects. It is however important to have the random effects in mind when analysing longitudinal data, otherwise serial correlation could occur in the error term. Thus I construct a random effects ordered probit model:

$$y_i^{c*} = \beta_0 + \beta_1 Torreon + \beta_2 Post + \beta_3 (Torreon * Post) + \beta_4' x_{it}^c + \gamma_i + \varepsilon_i^c$$

γ_i is the unobserved individual differences (random effects). The random effects estimator allows me to estimate the effects of variables that are individually time-invariant, such as my Torreon variable that is crucial for this study. This is a big advantage with the random effects model, as the fixed effect specification does not allow for time-invariant variables as they are absorbed by the intercept (Greene, 2012). Both the random effects and error term are assumed to be zero on average with constant variance.

6.3 Assumptions & Endogeneity issues

There are some assumptions we have to consider when using a random effects model, for instance that the random effects are uncorrelated between individuals and time $\{cov(\gamma_i, \gamma_j) = 0 \; i \neq j\}$ (Carter Hill et al., 2011). Moreover the random effects are assumed to be uncorrelated with the error term $\{ cov(\gamma_i, \varepsilon_i) = 0 \}$ and explanatory variables $\{ cov(\gamma_i, x_i) = 0 \}$. If these assumptions are not fulfilled, the investigation suffers from endogeneity issues. A risk of endogeneity in this study could be from omitted variable bias as, for instance, I do not include an income variable in my model. According to Carter Hill et al. (2011) this could lead to biased estimators as variables related to income, for instance education, might be overestimated. For instance, it could be the case that the citizens of Torreon are richer due to migration as highly educated people from other parts of Mexico move there to work, as Torreon is very industrialized. In this case, Gomez Palacio would not be a good control group for Torreon. It could also be the case that the levels of pollution exposure vary by much in the pre period between the two cities. If Torreon were to be more polluted, since it is located closer to the smelter, then Gomez Palacio would not be representative for Torreon. However, I assume that the level of heavy metal exposure is similar between the two cities and moreover I assume that the emission of metal particles from Peñoles Met-Mex has not decreased. I cannot provide evidence that the emissions have not changed due to lack of data. The lack of data is one of the greatest limitations in this study, especially considering the parallel trend assumption (Albouy, 2004). The parallel trend assumption relates to the reliability of the control group and should be considered when using difference-in-differences specifications, i.e. interaction term in my case. The difference-in-differences method assumes that the treatment- and control group are similar regarding certain variables of interest. The issue however is that there often are differences between the two groups but it is hard to test how great these differences are. One possible test is to analyse the trend concerning certain variables of interest over a short time period before the policy was implemented (Meyer, 1995). However, as mentioned, in this study no such data is available for neither Torreon nor Gomez Palacio. Thus there is a great risk that the estimator of the interaction will be biased.

As discussed, the lack of data causes many problems in this study. A final issue, which I believe also is the greatest limitation in this paper, is the lack of data from before the policy implementation (as I only have data from 2002 when the policy was implemented in 1999). It could be the case that health changed remarkably during the three years (1999 to 2002) that my study would not be able to observe. Thus my results would not show the true impact of the environmental policy on health status. However, we should bear in mind that decreasing the levels of heavy metal exposure is a long-term process. Moreover, in this study I only investigate the outcomes from adults (15 years and older), which is an advantage since it usually takes some time to observe changes in adults in health status from decreased exposure.

7. Results

7.1 Ordered Probit Models

The aim of this analysis is to investigate the impact of different socio-economic and demographic factors but mainly how the self-reported health changes over time. In both table 5 and 6, column 1 and 2 show the results from the pooled ordered probit while column 3 and 4 the random effects ordered probit. Moreover, Column 2 and 4 contain results from the clustered standard errors. The clustered standard errors allow for the error terms to be correlated within each family. If we do not relax the assumption of uncorrelated standard errors between the family members over time, the standard errors could become too small, which would falsely overstate the reliability of the estimators (Hill et al., 2011). In other words, the clustered standard errors should be greater than the unclustered standard errors. Another advantage with the clustered standard errors is that it solves the issue of autocorrelation between the errors over time for the same individual. This is an important aspect when using panel data as the observations in the post period could be dependent on observations from the pre period. Finally, the only difference between these two tables is the interaction term, which I included in table 6.

Table 5 shows that individuals with a higher education level are more likely to report higher self-reported health than the low educated. Moreover, the older you get the less likely you are to report *Very good* health status. Males are more likely to report higher levels than females and those with health insurance are also more likely to report higher health status than the individuals without. Finally, an interesting finding with respect for this analysis is that people in Torreon are more likely to report higher self-reported health than the inhabitants in Gomez Palacio.

Table 5
Determinants of Self-Reported Health

Variables	Coefficients			
	Pooled Ordered Probit		Random Effects Ordered Probit	
	(1)	(2)	(3)	(4)
Torreon	0.277***	0.306**	0.277***	0.306**
	(0.0985)	(0.127)	(0.104)	(0.123)
Post	0.112	0.130	0.112	0.130
	(0.110)	(0.113)	(0.0955)	(0.112)
Age	-0.0900***	-0.102***	-0.0900***	-0.102***
	(0.0245)	(0.0296)	(0.0244)	(0.0300)
Age^2	0.000753***	0.000848**	0.000753**	0.000848**
	(0.000288)	(0.000348)	(0.000296)	(0.000359)
Higher level education	0.448***	0.466***	0.448***	0.466***
	(0.113)	(0.141)	(0.122)	(0.142)
Female	-0.210*	-0.242	-0.210	-0.242
	(0.119)	(0.150)	(0.129)	(0.152)
Working	0.122	0.132	0.122	0.132
	(0.134)	(0.152)	(0.118)	(0.139)
Married	0.0793	0.0538	0.0793	0.0538
	(0.113)	(0.137)	(0.119)	(0.131)
Insuranced	0.479***	0.521***	0.479***	0.521***
	(0.121)	(0.131)	(0.116)	(0.139)
Exercises	0.110	0.0720	0.110	0.0720
	(0.127)	(0.149)	(0.134)	(0.157)
Smoking	-0.134	-0.138	-0.134	-0.138
	(0.117)	(0.137)	(0.110)	(0.129)
Observations	682	682	682	682
Clustered Standard Errors	No	Yes	No	Yes

*** p<0.01, ** p<0.05, * p<0.1

The issue in table 5 is that we do not know if the likelihood for the inhabitants in Torreon to report higher self-reported health has changed over time. For a successful policy implementation we would expect that people in Torreon are more likely to report higher health status in the post period than in the pre period compared with the citizens of Gomez Palacio. To investigate this matter, I construct an interaction term containing the two dummy variables, Torreon and Post. In table 6 we observe that this new interaction term, representing the policy effect, is positive and statistically significant. It seems as if the environmental policy implementation has had a positive impact on people's subjective health status. Additional notes from the results are that the Torreon variable in this table is statistically insignificant while the post variable is statistically significant. However, these results are most possibly due to high multicollinearity where the interaction term correlates with its components.

Table 6
Policy Effect on Self-Reported Health

Variables	Coefficients			
	Pooled Ordered Probit		Random Effects Ordered Probit	
	(1)	(2)	(3)	(4)
Torreon	-0.109	-0.118	-0.109	-0.118
	(0.144)	(0.169)	(0.140)	(0.163)
Post	-0.287*	-0.315**	-0.287**	-0.315**
	(0.149)	(0.160)	(0.135)	(0.155)
Torreon*Post	0.648***	0.727***	0.648***	0.727***
	(0.199)	(0.221)	(0.178)	(0.213)
Age	-0.0896***	-0.102***	-0.0896***	-0.102***
	(0.0243)	(0.0297)	(0.0242)	(0.0301)
Age^2	0.000743***	0.000840**	0.000743**	0.000840**
	(0.000286)	(0.000349)	(0.000293)	(0.000359)
Higher level education	0.451***	0.472***	0.451***	0.472***
	(0.113)	(0.141)	(0.121)	(0.141)
Female	-0.214*	-0.247	-0.214	-0.247
	(0.120)	(0.153)	(0.131)	(0.155)
Working	0.101	0.105	0.101	0.105
	(0.133)	(0.150)	(0.115)	(0.136)
Married	0.0893	0.0691	0.0893	0.0691
	(0.112)	(0.138)	(0.120)	(0.133)
Insuranced	0.501***	0.554***	0.501***	0.554***
	(0.119)	(0.128)	(0.111)	(0.134)
Exercises	0.152	0.128	0.152	0.128
	(0.133)	(0.161)	(0.139)	(0.167)
Smoking	-0.125	-0.119	-0.125	-0.119
	(0.119)	(0.142)	(0.114)	(0.135)
Observations	682	682	682	682
Clustered Standard Errors	No	Yes	No	Yes

*** p<0.01, ** p<0.05, * p<0.1

In table 7 below I also report the marginal effects of the predicted probabilities keeping all variables at their means. The different outcomes of the self-reported health variable are categorized in the columns and the interaction term (with its components) are listed in the rows. The marginal effects for the interaction term are captured from the model represented in table 6, whereas the marginal effects of “Torreon” and “Post” are taken from model specification in table 5. The reason why the two components are calculated from the model without the interaction is to observe the true effect of these variables as the interaction term could cause misleading estimations due to multicollinarity.

As we can see, “Post” is statistically insignificant for all possible outcomes, whereas “Torreon” is statistically significant in all cases except for those who have reported *Very bad* health status. However, as demonstrated in table 4, this is probably due to the small proportion of people reporting *Very bad* health status. For *Very good* and *Good* the terms are positive, while the proportion of people in Torreon that have reported *Regular* and *Bad* health status decreases. More specifically, keeping all variables at their means, the predicted probabilities of reporting *Very good* or *Good* health status increases by 3.3 percentage points respectively 6.7 percentage points. This means that the inhabitants of Torreon are 3.3 percentage points more probable of reporting *Very good* health status than the inhabitants of Gomez Palacio. Moreover, the predicted probabilities of reporting *Regular* or *Bad* health status decreases by 9 percentage points respectively 0.9 percentage points. Regarding the interaction term we observe the same pattern indicating that the predicted probabilities of reporting *Very good* or *Good health* status increases by 7.4 percentage points respectively 15.9 percentage points. At the same time the predicted probabilities of reporting *Regular* or *Bad* health status decreases by 21.2 percentage points respectively 2.1 percentage points.

Table 7
Marginal effects at Means

	Very good	Good	Regular	Bad	Very bad
Torreon	0.033**	0.067***	-0.09***	-0.009**	-0.0003
Post	0.013	0.027	-0.037	-0.004	0.0001
Torreon*Post	0.074***	0.159***	-0.212***	-0.021**	-0.0006

*** p<0.01, ** p<0.05, * p<0.1

From now on I only use the pooled ordered probit for my regressions, as it is my preferred model for the upcoming probit models in the robustness section. The reason for this selection is because the pooled probit model is simpler to use but still show almost identic results as the random effects model. I therefore assure that relaxing the assumption of individual differences will not be an issue for further analysis.

7.2 Robustness checks

7.2.1 Merging *Good* and *Very good* Self-Reported Health Status

To analyse if the results in previous regressions are robust, I now estimate the change in self-reported health by using a pooled probit model. I do this by constructing a new dummy variable that represents the individuals that have reported either *Good* or *Very good* health status. In other words, instead of having five different possible outcomes for self-reported health, I now only have two. These two outcomes could be seen as “good health” or “bad health”.

In table 8 I report the results from this new specification of the self-reported health variable. As we can see, the results are very similar with the exception that “Torreon” is statistically significant at a 10% level in column 2 representing the model with clustered standard errors. Moreover, in table 9 I report the marginal effects with the same specification as in table 7, but now for my new model. In this case, the two components are both statistically insignificant while the interaction term is statistically significant at a 1% level.

Thus the marginal effects does not seem to show a robust results as both “Torreon” and the interaction term were statistically significant in table 7.

Table 8
Policy Effect on Self-Reported Health

Variables	Coefficients	
	Pooled Ordered Probit	
	(1)	(2)
Torreon	-0.291	-0.291*
	(0.187)	(0.176)
Post	-0.368*	-0.368**
	(0.197)	(0.160)
Torreon*Post	0.657***	0.657***
	(0.246)	(0.198)
Age	-0.0939***	-0.0939***
	(0.0288)	(0.0285)
Age^2	0.000727**	0.000727**
	(0.000346)	(0.000341)
Higher level education	0.468***	0.468***
	(0.133)	(0.163)
Female	-0.373**	-0.373**
	(0.146)	(0.180)
Working	0.104	0.104
	(0.145)	(0.128)
Married	0.00948	0.00948
	(0.135)	(0.140)
Insuranced	0.608***	0.608***
	(0.133)	(0.130)
Exercises	0.115	0.115
	(0.151)	(0.150)
Smoking	-0.182	-0.182
	(0.161)	(0.152)
Observations	682	682
Clustered Standard Errors	No	Yes

*** p<0.01, ** p<0.05, * p<0.1

Table 9
Marginal effects at Means

	Very good/Good
Torreon	0.041
Post	0.016
Torreon*Post	0.235***

*** p<0.01, ** p<0.05, * p<0.1

7.2.2 Self-Reported Health Without Extreme Ages

We observed in both table 5 and 6 that education has a strong significant impact on self-reported health status. As the sample does not contain an income variable, “Higher level education” is the variable that is the most representative for socio-economic status. The individuals, who study longer, receive higher salary than the less educated, and as a consequence they are more likely to report *Very Good* health status. This could cause a problem in the previous two tables as the individuals who are students in the pre period and workers in the post period, could go from being poor to rich in only a couple of years. This would obviously impact the results by making the interaction term more significant than it actually is. For the old, however, the change in socio-economic status would be the contrary. In other words, they would go from having a stable income from employment to a much smaller income due to retirement. This could on the other hand have a negative impact on self-reported health. For this reason, in table 10 I drop individuals from my sample that were younger than 22 or older than 60 in the pre period. The limit was drawn at 22 since the majority of the Mexican students graduate from university at this age. Moreover, the upper limit was drawn at the age of 60 since this is the age when the Mexican seniors tend to retire. Table 10 presents the results from this sample, and as we can see, the outcome is basically the same as in table 6. Thus it seems as these two socio-economic changes in life do not have an impact on self-reported health and the model is robust in this sense.

Table 10
Policy Effect on Self-Reported Health

Variables	Coefficients	
	Pooled Ordered Probit	
	(1)	(2)
Torreon	-0.135	-0.150
	(0.166)	(0.191)
Post	-0.286*	-0.311*
	(0.163)	(0.178)
Torreon*Post	0.668***	0.746***
	(0.216)	(0.243)
Age	-0.0932***	-0.106***
	(0.0330)	(0.0366)
Age^2	0.000755**	0.000861**
	(0.000373)	(0.000416)
Higher level education	0.419***	0.445***
	(0.120)	(0.145)
Female	-0.244*	-0.255
	(0.128)	(0.156)
Working	0.107	0.128
	(0.142)	(0.159)
Married	0.124	0.133
	(0.116)	(0.142)
Insuranced	0.526***	0.591***
	(0.129)	(0.135)
Exercises	0.144	0.117
	(0.146)	(0.180)
Smoking	-0.157	-0.170
	(0.126)	(0.147)
Observations	614	614
Clustered Standard Errors	No	Yes

*** p<0.01, ** p<0.05, * p<0.1

7.2.3 Self-Reported Illnesses, Hospitalizations and Doctor Visits

There has been a wide discussion on the relevance of using self-reported health as a measure of health status. Some investigators have found a significant difference in the results between self-reported health with other health related factors, see for instance Lambrinos (1981), Anderson & Burkhauser (1985) and Bazzoli (1985). For this reason I find necessary to also analyse the possible effect of the policy implementation on other measurements for health status; namely illnesses, hospitalizations and doctor visits. Thus I construct three new models where I drop self-reported health as exploratory variable and exchange it with these variables that are more objective estimates for health status. As these variables are binary, I now construct pooled probit regressions instead of ordered probit models. However, before estimating these new variables I wish to analyse the correlation between self-reported health, illnesses, hospitalizations and doctor visits. These results are presented in table 8 below and as we can see, the self-reported health variable is negatively correlated with the other three more objective measures. This means that illnesses, hospitalizations and doctor visits affects self-reported health negatively. These “harder” measures of health, on the other hand, are positively correlated with each other. Logically, illnesses and doctor visits have the strongest correlation as people in Mexico normally prescribe medicines at clinics when being ill.

Table 11
Correlation of the Dependent Variables

	Self-reported health	Self-reported illnesses	Hospitalized	Doctor visits
Self-reported health	1			
Self-reported illnesses	-0.128	1		
Hospitalized	-0.066	0.058	1	
Doctor visits	-0.153	0.205	0.069	1

In table 12 I now estimate the three new variables for health status where column 1 and 2 show the results from the unclustered and clustered standard errors for illnesses, column 3 and 4 for hospitalizations, finally column 5 and 6 for doctor visits. Interestingly, we observe that the interaction term is positive and statistically significant in column 1 and 2, meaning that self-reported illnesses in Torreon has increased over time. I.e. these results do not support the previous findings from table 6 where people in Torreon were more likely to report higher self-reported health in the post period. These are quite surprising results as illnesses ought to affect self-reported health as discussed previously regarding table 8. The interaction terms are unfortunately statistically insignificant when estimating hospitalizations and doctor visits in column 3-6.

Table 12
Policy Effect on Self-Reported Illnesses, Hospitalizations & Doctor Visits

Variables	Coefficients					
	Pooled Probit					
	Illness		Hospitalized		Doctor visits	
	(1)	(2)	(3)	(4)	(5)	(6)
Torreon	-0.634***	-0.634***	0.143	0.143	-0.251	-0.251
	(0.189)	(0.210)	(0.250)	(0.250)	(0.225)	(0.235)
Post	-0.356*	-0.356	-0.0819	-0.0819	-0.153	-0.153
	(0.204)	(0.234)	(0.273)	(0.277)	(0.229)	(0.226)
Torreon*Post	0.518**	0.518*	-0.228	-0.228	-0.00384	-0.00384
	(0.244)	(0.277)	(0.339)	(0.349)	(0.301)	(0.296)
Age	0.0167	0.0167	0.0703*	0.0703*	0.0760**	0.0760*
	(0.0277)	(0.0260)	(0.0397)	(0.0423)	(0.0361)	(0.0402)
Age^2	-0.000189	-0.000189	-0.000868*	-0.000868*	-0.000780*	-0.000780
	(0.000333)	(0.000315)	(0.000475)	(0.000507)	(0.000435)	(0.000483)
Higher level education	0.00775	0.00775	-0.0163	-0.0163	0.0758	0.0758
	(0.124)	(0.118)	(0.193)	(0.177)	(0.166)	(0.149)
Female	0.175	0.175	0.0515	0.0515	0.758***	0.758***
	(0.134)	(0.128)	(0.209)	(0.190)	(0.184)	(0.199)
Working	-0.259*	-0.259*	-0.335*	-0.335*	0.0273	0.0273
	(0.134)	(0.145)	(0.181)	(0.186)	(0.167)	(0.198)
Married	-0.000986	-0.000986	-0.432**	-0.432**	0.154	0.154
	(0.130)	(0.134)	(0.186)	(0.178)	(0.164)	(0.163)
Insuranced	0.0266	0.0266	0.401**	0.401**	0.0311	0.0311
	(0.131)	(0.118)	(0.187)	(0.179)	(0.168)	(0.153)
Exercises	-0.0983	-0.0983	-0.179	-0.179	0.147	0.147
	(0.137)	(0.135)	(0.183)	(0.176)	(0.167)	(0.141)
Smoking	0.149	0.149	0.306	0.306	0.189	0.189
	(0.153)	(0.143)	(0.229)	(0.209)	(0.186)	(0.160)
Observations	682	682	682	682	682	682
Clustered Standard Errors	No	Yes	No	Yes	No	Yes

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

There are several studies showing a strong relationship between self-reported health and mortality, see for example Benyamini & Idler (1997) and Benyamini & Idler (1999). Reindl Benjamins et al. (2004) extend the literature by claiming that this relationship depends on the cause of death. Deaths due to diabetes, infectious and respiratory diseases are strongly correlated to self-reported health. Such diseases on the other hand can be provided by arsenic and lead exposure. For this reason I will focus more on illnesses in this section, but specifically on illnesses related to arsenic, lead and cadmium exposure. Diseases that could be caused by arsenic pollution are high blood pressure and itching skin (WHO, 2012). Moreover, lead could provide stomach problems and headache (NIOSH, 2013). Finally vomit, diarrhea, lung problems and body ache could be a result from cadmium pollution (ATSDR, 2012). Thus in my new illness variable I now only include the diseases: cough, breathing issues, stomach pain, nausea/vomit, diarrhea, itching skin, headache, fever, body ache, chest pain, respiratory/urinary/digestive problems and blood pressure. Moreover, I drop the individuals from previous regression that did not report that they had suffered from any illness. This is a necessary action since I wish to analyse if the statistically significant interaction term in previous table is mainly due to illnesses related to heavy metal exposure. In other words, I do this to isolate the effect of illnesses related to metal pollution from other illnesses included in the previous illness variable. Table 13 show the results from this regression and as we can see, the interaction term is statistically insignificant in this case. Thus I cannot conclude that the positive interaction term in previous regression is due to an increase in self-reported illnesses related to heavy metal exposure. However, I should note that 84.5% of the individuals in this sample reported that they had suffered from a metal related illness. Thus the insignificant results are not very surprising. Nevertheless, it provides some evidence that most likely the increase in self-reported illnesses, presented in table 12 above, are due to illnesses that are not caused by heavy metal pollution (for instance, flu, throat pain and stomach pain). These illnesses are not as severe as illnesses related to heavy metal pollution, which might be reason why we do not observe the same pattern for self-reported health as for illnesses. Namely because these illnesses does not have a strong impact on general health status as the chronic ones caused by metal exposure and therefore should not have a strong impact on self-reported health. Another interesting fact however is that the post variable is statistically significant in the clustered pooled model. This means that individuals who have suffered from an illness during the last month are more likely to report that they have suffered from an illness related to metal exposure in the post period.

Table 13
Policy Effect on Disorders Related to Metal Pollution

Variables	Coefficients	
	Pooled Probit	
	Metal illness	
	(1)	(2)
Torreon	0.0746	0.0746
	(0.269)	(0.251)
Post	0.0896	0.0896
	(0.283)	(0.228)
Torreon*Post	0.391	0.391
	(0.380)	(0.351)
Age	-0.0110	-0.0110
	(0.0416)	(0.0427)
Age^2	5.88e-05	5.88e-05
	(0.000492)	(0.000493)
Higher level education	0.477**	0.477**
	(0.190)	(0.196)
Female	0.581***	0.581***
	(0.210)	(0.213)
Working	-0.331	-0.331
	(0.203)	(0.224)
Married	0.298	0.298
	(0.215)	(0.249)
Insuranced	-0.152	-0.152
	(0.231)	(0.198)
Exercises	0.420*	0.420*
	(0.229)	(0.220)
Smoking	0.180	0.180
	(0.227)	(0.218)
Observations	396	396
Clustered Standard Errors	No	Yes

*** p<0.01, ** p<0.05, * p<0.1

8. Heterogenous effects

In this section I will construct similar regressions as in table 6, however, now I will divide the sample depending on three different attributes; namely gender, education level and age. The reason for this is to get a clearer view of how my variables of interest may impact an individual's health status depending on these attributes.

Table 14

Self-Reported Health (Gender, Education & Age)

Variables	Coefficients					
	Pooled Probit					
	Male	Female	Higher education	Lower education	Age≤32	Age>32
	(1)	(2)	(3)	(4)	(5)	(6)
Torreon	-0.595***	0.0742	-0.182	-0.0321	-0.373*	0.0375
	(0.223)	(0.189)	(0.231)	(0.172)	(0.207)	(0.197)
Post	-0.118	-0.392*	-0.148	-0.303*	-0.442**	-0.181
	(0.213)	(0.206)	(0.225)	(0.178)	(0.212)	(0.178)
Torreon*Post	0.683**	0.677***	0.688**	0.532**	0.728**	0.590**
	(0.273)	(0.254)	(0.303)	(0.241)	(0.291)	(0.235)
Age	-0.0582	-0.0939***	-0.131***	-0.0729***	-0.148	-0.146**
	(0.0429)	(0.0264)	(0.0475)	(0.0246)	(0.166)	(0.0605)
Age^2	0.000565	0.000728**	0.00121**	0.000528*	0.00225	0.00130**
	(0.000491)	(0.000326)	(0.000570)	(0.000294)	(0.00332)	(0.000600)
Higher level education	0.474**	0.413***	-	-	0.603***	0.383**
	(0.192)	(0.157)			(0.152)	(0.174)
Female	-	-	0.0108	-0.370**	0.161	-0.509***
			(0.221)	(0.154)	(0.166)	(0.175)
Working	0.145	0.0766	0.363	-0.0387	0.194	0.0243
	(0.251)	(0.126)	(0.228)	(0.128)	(0.162)	(0.164)
Married	-0.218	0.140	0.274	-0.00828	0.167	-0.0614
	(0.227)	(0.148)	(0.248)	(0.133)	(0.194)	(0.160)
Insuranced	0.594***	0.482***	0.177	0.633***	0.424***	0.617***
	(0.197)	(0.141)	(0.196)	(0.136)	(0.148)	(0.162)
Exercises	0.331*	0.181	0.558***	-0.165	0.101	0.275*
	(0.171)	(0.192)	(0.164)	(0.192)	(0.222)	(0.152)
Smoking	-0.176	-0.220	-0.225	-0.104	-0.189	-0.0563
	(0.190)	(0.183)	(0.185)	(0.156)	(0.165)	(0.162)
Observations	283	399	239	443	207	412
Clustered Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

8.1 Gender

It is still common in developing countries that men and women fulfil different roles in the society, where men work and women are housewives. Mexico is not an exception, and therefore I find it relevant to compare the probability of reporting *Very Good* health status between men and women as their roles in the Mexican society could determine their health status. I analyse 122 men (column 1) and 175 women (column 2), and find that there is a difference between male and female regarding the effect of different independent variables. For example, a man from Torreon is significantly less probable of reporting *Very Good* health than if he were from Gomez Palacio, while for women this is not the case. On the other hand, women in the Laguna district are less probable of reporting *Very Good* health status in the post period in comparison with the pre period. This might have to do with age though as the age-variable has a stronger impact on women. Interestingly, the interaction term is statistically significant for both male and female that indicates that independently on gender, people in

Torreon feel healthier in the post period in comparison with Gomez Palacio and the pre period.

8.2 Education

Education is widely seen as the main component to improve health status (Grossman and Kaestner, 1997). It is believed that more educated people are more efficient producers of health as they are better at maximizing their production functions of health. For example, they are aware of the consequences that cigarettes provide and therefore they avoid smoking or they understand better the importance of an appropriate diet. Moreover, level of education is often causal to socio-economic factors, such as income. People with higher income on the other hand can spend more on medications, healthy food and water, but also housing in areas that are not exposed to pollution. As this study does not have an income variable, the education variable becomes even more important and to get deeper knowledge on the impacts of education on health. For this reason I, as in previous section, divide the sample depending if the individuals have higher education level or not to analyse if there is a change between the results of the two groups.

In table 11 we observe in column 3 and 4 we observe that, independently on education level, people in Torreon feel healthier in the post period in comparison with Gomez Palacio and the pre period. Moreover, women that have not continued studying after secondary school are less likely to report *Very Good* health status than men. However, the highly educated women demonstrate no such patterns. Also the low educated individuals with insurance are more likely to report *Very Good* health status, but this is not the case for the highly educated. Finally, highly educated people that exercises are more likely to report *Very Good* health status than those who do not exercise. The low educated on the other hand show no such pattern, which might be due to their financial situation as exercising usually is an activity for the high- and middle class.

8.3 Age

To get a clearer impression of the impact of age on health status, I divide the sample into two groups depending on age, namely those who are 32 or under and those over 32. The limit was drawn at 32 since it is the median age for my sample in the pre period. Previously we observed that age had a stronger negative impact on women, in column 5 and 6 we observe that these results mainly accounts for women above 32. However the fact that individuals at 32 or below are less likely to report *Very Good* health status in the post period than in the pre period is quite surprising. Concerning the interaction term, as in previous cases, people in Torreon feel healthier in the post period in comparison with Gomez Palacio and the pre period. Interestingly, however, is the fact that the interaction term is much larger for young people than the old.

9. Conclusions & Discussion

The purpose of this study was to investigate if the environmental policy implemented in Torreon has had a positive impact on subjective health status. The results from this paper are very interesting considering the fact that very few (if any) studies have been made on this subject, namely the policy impact on self-reported health in a mining district. Moreover, the different structure of the policy makes the analysis even more unique as the emissions from the smelter did not decrease during the process. In the results section we could see that the interaction term (Torreon*Post), representing the environmental policy implementation, had a statistically significant impact on self-reported health status. In other words, people overall feel healthier in Torreon in the post period than in the pre period, when using Gomez Palacio as control group. I also provided evidence that the interaction term was not affected by great socio-economic changes (graduation or retirement) in the individuals' lives. However, interestingly these results are not supported by the analysis on hard measures of health. The interaction term, which was positive and statistically significant for self-reported health, show no evidence that this was due to a decrease in self-reported illnesses, hospitalizations or doctor visits. Rather the opposite since the self-reported illnesses increased in Torreon from the pre period to the post period. I believe that there are four possible reasons for these results. First, as we saw in the robustness section, the increase in the illness variable seem to be due to diseases that are not caused by metal pollution, and therefore are not that severe towards general health status. For instance, throat ache should not have as severe impact on self-reported health status as breathing issues. Second, it could be the case that self-reported health is affected by psychological factors. In other words, people in Torreon feel healthier simply due to the knowledge of the policy implementation in the city. Third, another reason might be that hard measures of health (illnesses, hospitalizations and doctor visits) need more time to improve from a positive environmental change. As discussed in the data description section, subjective health could be more flexible over a short period of time than hard measures. Fourth, as I only include individuals in my sample that are born and raised in the city they currently live in, it could be the case that objective health cannot change for the better after growing up in a polluted area like the Laguna district.

I have three limitations worth mentioning that are important to have in mind when analysing the results. Firstly, the insecurity that Gomez Palacio is a good control group for Torreon, as I cannot test the parallel trend assumption due to lack of data. Secondly, the small sample size could be a reason for the low number of statistical significant variables. Thirdly, I do not have any observations from the time period before the implementation was made. As mentioned in the methodology section, this is probably the greatest limitation of this study. Therefore I cannot conclude that for example illnesses have not decreased in Torreon, as it could be the case that they decreased remarkably the time period before my observations (1999-2002).

For decision-makers, my advice regarding policy implementations is to construct programs aimed at decreasing emissions in a more direct form, as I believe that such a policy guarantees improvements in health status. From this study we have learned that the health outcomes concerning policy implementations (as the one in Torreon) are more uncertain, i.e. by keeping production at its normal pace and compensate by cleaning the cities, building parks etc. However, I recommend future investigators to do more researches on this subject. If it were the case that such a win-win policy is successful, this would completely change the decision-making process of public policies. For future research I would also recommend for investigators to use data on level of contamination. In other words, level of heavy metals in soil and water and not on emissions, as the purpose is to analyse the effect from a policy similar to the one in Torreon. Finally, it would be interesting to see the results from a similar investigation, but with bigger sample size. As I discussed previously, the small sample size could be a reason for the few variables that are statistically significant.

References

Albouy, D. (2004). Program Evaluation and the Difference in Difference Estimator. Section Notes, Economics 131.

Anderson, K.H., Burkhauser, R.V. (1985). The Retirement-Health Nexus: A New Measure of an Old Puzzle. Journal of Human Resources 20, 315-330.

ATSDR, Agency for Toxic Substances & Disease Registry (2012). Public Health Statement for Cadmium. 1.5 How Can Cadmium Affect My Health?

Bazzoli, G.J. (1985). The Early Retirement Decision: New Empirical Evidence on the Influence of Health. Journal of Human Resources 20, 214-234.

Benyamini, Y., Idler, E.L. (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. Journal of Health and Social Behaviour, 38(1): 21-37.

Benyamini, Y., Idler, E.L. (1999). Community Studies Reporting Association Between Self-Rated Health and Mortality. Research on Aging, 21(3): 392-401.

Calderón Mendoza, R., Jiménez, A. (2015). Mining 2015 Mexico. Latin Lawyer: The Business Law Resource for Latin America. Question 1 and 3.

Carter Hill, R, Griffiths, W.E., Lim, G.C. (2011). The Principles of Econometrics. (4th Edition) John Wiley & Sons, Inc. The Difference-in-difference Estimator (pp. 286-287).

Chaparro, E.A., (2007). Buenas Prácticas en la Industria Minera: El Caso de Grupo Peñoles en México. División de Recursos Naturales e Infraestructura, 29-41.

Chay, K., Dobkin, C., Greenstone, M. (2003). The Clean Air Act of 1970 and Adult Mortality. Journal of Risk and Uncertainty. Vol. 27, Iss. 3:279-300.

Currie, J., Neidell, M. (2005). Air Pollution and Infant Health: What Can We Learn from California’s Recent Experience? Quarterly Journal of Economics, 120(3): 1003-1030.

Dechênes, O., Greenstone, M., Shapiro, J.S. (2012). Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program and Ozone Reductions. The National Bureau of Economic Research, Working paper no. 18267.

GAHP (2014). Pollution – The Silent Killer of Millions in Poor Countries. Global Alliance on Health and Pollution.

Garcia-Vargas, G.G., Rothenberg, S.J., Silbergeld, E.K., Weaver, V., Zamoiski, R., Resnick, C., Rubio-Andrade, M., Parsons, P.J., Steuerwald, A.J., Navas-Acién, A., Guallar, E. (2014). Spatial Clustering of Toxic Trace Elements in Adolescents Around the Torreon, Mexico Lead-Zink Smelter. Journal of Exposure Science & Environmental Epidemiology, 24, 634-642.

Graff Zivin, J., Neidell, M. (2013). Environment, Health and Human Capital. Journal of Economic Literature, vol. 51, issue 3:689-730.

Greene, W.H. (2012). Discrete Choices and Event Counts. Econometric Analysis, 7th Edition. U.S.

Greenstone, M., Hanna, R. (2013). Environmental Regulations, Air and Water Pollution, and Infant Mortality in India. American Economic Review, vol. 104, iss. 10: 3038-3072.

Grossman, M. & Kaestner, R. (1997). The Effects of Education on Health. The Social Benefits of Education, Behrman, J. & Stacy, N. Michigan.

IDRE (2016). Applied Survey Data Analysis in Stata 9. Statistical Computing Seminars. Institute for Digital Research and Education. University of California.

Lambrinos, J. (1981). Health: A Source of Bias in Labor Supply Models. Review 63, 81-98.

Méndez-Gómez, J., García-Vargas, G.G., López-Carillo, L., Calderón-Aranda, E.S., Gómez, A., Vera, E., Valverde, M., Cebrián, M.E., Rojas, E. (2008). Genotoxic Effects of Environmental Exposure to Arsenic and Lead on Children in Region Lagunera, Mexico. New York Academy of Sciences, 1140:358-367.

Meyer, B. D. (1995). Natural and Quasi-Experiments in Economics. Journal of Business & Economic Statistics, Vol. 13, No. 2.

Min, H. (2013). Ordered Logit Regression Modeling of the Self-reported Health in Hawaii, With Comparisons to the OLS Model. Journal of Modern Applied Statistical Methods, Vol. 12, Iss.2, Article 23.

NIOSH, The National Institute for Occupational Safety and Health (2013). Centers for Disease Control and Prevention. Health Problems Caused by Lead.

Petru, A. (2015). Report: Pollution is the Leading Cause of Death in the Developing World. Triple Pundit, January 28th.

Recio-Vega, R., Valdez-Abrego, C., Adame-Lopez, B., Gurrola-Mendez, A. (2012). Surveillance of elevated blood lead levels in children in Torreon, Coahuila, Mexico, 1998-2010. International Journal of Hygiene and Environmental Health, vol. 215, 507-513.

Reindl Benjamins, M., Hummer, R.A., Eberstein, I.W., Nam, C.B. (2004). Self-Reported Health and Adult Mortality Risk: An Analysis of Cause-Specific Mortality. Social Science & Medicine, 59(6): 1297-1306.

Rosado, J.L., Ronquillo, D., Kordas, K., Rojas, O., Alatorre, J., Lopez, P., Garcia-Vargas, G.G., del Carmen Caamaño, M., Cebrián, M.E., Stoltzfus, R.J. (2007). Arsenic

Exposure and Cognitive Performance in Mexican Schoolchildren. Environmental Health Perspectives, 115(9), 1371-1375.

The Economist (2011, 3 Sept). Black Sand in the Desert. The Economist Newspaper Limited.

Valdéz Perezgasga, F., Cabrera Morelos, V.M. (1999). La Contaminacion por Metales Pesados en Torreon, Coahuila, Mexico. Texas Center for Policy Studies, 27-31.

WHO, World Health Organization (2012). Arsenic, Health Effects. Fact Sheet No. 372.

Wibowo, D., Tisdell, C. (1992). Health Production Function for Preventive Health Programs. Proceedings of the Fourteenth Australian Conference of Health Economics, p.106-133.

Yansaneh, I., S. (2003). Construction and Use of Sampling Weights. Handbook on Designing of Household Sample Surveys.