Maryland Population Research Center

The Long-Term Effects of Early Lead Exposure:
Evidence from a Case of Environmental Negligence

Authors:

Tomás Rau
Pontificia Universidad
Católica de Chile

Loreto Reyes
Ministry of Finance,
Chile

Sergio S. Urzúa
University of Maryland

THE LONG-TERM EFFECTS OF EARLY LEAD EXPOSURE:
EVIDENCE FROM A CASE OF ENVIRONMENTAL NEGLIGENCE

Tomás Rau
Loreto Reyes
Sergio S. Urzúa

The Long-term Effects of Early Lead Exposure: Evidence from a case of Environmental Negligence
Tomás Rau, Loreto Reyes, and Sergio S. Urzúa
NBER Working Paper No. 18915
March 2013
JEL No. I15,I18,I25,J17,O1,Q5,Q51,Q53,Q56

ABSTRACT

This paper estimates the effect of early lead exposure on academic achievement and adult earnings. We analyze longitudinal information from individuals attending primary and secondary schools in the city of Arica (in northern Chile). Between 1984 and 1989, Arica received more than 20,000 tons of toxic chemicals containing high concentrations of lead. Initially, the chemical waste was located several kilometers from the city. However, Arica's rapid expansion, which included the construction of housing projects just meters away from the waste deposit, put a large number of families at risk. Our data include information on residential proximity to the polluted area, levels of lead exposure, comprehensive demographic information, nationally representative academic test scores and administrative data on adult earnings.

We document a strong relationship between blood lead levels and student academic performance. We find that an increase of one microgram of lead per deciliter of blood reduces math and language scores by 0.15 and 0.21 standard deviations, respectively. For earnings, we estimate that for each extra microgram of lead, monthly earnings decrease by CLP 11,458 (or USD 22.92). This translates into a reduction of USD 6,000 in lifetime earnings per microgram of lead per deciliter of blood.

"

Tomás Rau
Pontificia Universidad Católica de Chile
Avda. Vicuna Mackenna 4860, Macul
Santiago, Chile
trau@uc.cl

Loreto Reyes
Ministry of Finance, Chile
Teatinos 120, Piso 12
Santiago, Chile
lreyes@hacienda.gov.cl

Sergio S. Urzúa
Department of Economics
University of Maryland
3105 Tydings Hall #3115C
College Park, MD 20742
and NBER
urzua@econ.umd.edu

1 Introduction

In recent decades, lead pollution has received increasing attention, especially in developing countries. This emphasis has led to the development of programs to prevent and communicate the harmful effects of lead on human health, the implementation of public policies aimed at reducing emissions from major pollution sources (fuel, paints, light industries, and mineral processing manipulation, etc.), and more accurate measurement of lead concentration thresholds in soil, air, water, and blood. In 1991 for example, the Centers for Disease Control and Prevention (CDC) defined a threshold of 10 micrograms per deciliter of blood ($10\mu g/dL$), above which negative effects of lead on IQ, impaired hearing, behavioral problems (aggression) and delayed bone growth have been documented.[1] At higher levels, lead is associated with deterioration in the nervous system, severe liver damage, anemia and even death (PTF, 2000; ATDSR, 2007). Furthermore, because brain functions develop in early childhood, young children are especially vulnerable to the effects of lead. Early lead exposure might cause severe and irreversible effects on intellectual development (ATDSR, 2007; Tchernitchin et al., 2005).

In this paper we analyze the effects of blood lead concentration on the academic performance of students in the city of Arica (in northern Chile). As we describe in detail below, this provides a unique opportunity to study the effects of early lead exposure on humans. The presence of lead in Arica is attributed to the importation of 20,000 tons of mineral residues/waste from Sweden, between 1984 and 1989. The waste was imported by the Chilean company PROMEL, initially to be reprocessed to extract gold and silver. However, PROMEL quickly reported this would not be economically profitable and decided not to continue. As a result, the waste was left on the outskirts of the city at a location known as *site F*. According to the company, the toxic waste was removed in 1998, but as documented below, this area still presents high levels of lead today.

In the late 1980s and early 1990s Arica experienced rapid economic growth leading to an increasing demand for social housing projects. Unfortunately, and probably explained by the ignorance surrounding the situation, the regional government decided to construct social housing projects very close to *site F*. In 1995, two housing projects, *Cerro Chuño* and *Los Industriales*, were built only meters away from the site and hundreds of families moved to the area.

---

[1]For further details see CDC (1991).

We follow these precipitating events and evaluate the effect of lead exposure on academic achievement. We exploit the relationship between household distance to the contaminated area (*site F*), and lead blood concentration to estimate the consequences on academic achievement. First, we estimate a reduced-form effect of household proximity to contaminated areas on student academic performance. Then, using an empirical strategy in the spirit of the Two Sample Instrumental Variables (TSIV) (Angrist and Krueger, 1992), we estimate the effect of lead blood levels on academic achievement.

Our empirical strategy utilizes multiple sources of information. From the Chilean College Admission System (*Prueba de Selección Universitaria* or PSU) data set, we gather information on proximity to the polluted area (from home address to *site F*), multiple test scores (math and language) and a comprehensive set of individual and household socioeconomic variables for a large population of individuals living in Arica in 2008. We complement this data with individual's city of residence in 2004, also obtained from administrative records. By combining these sources of information, we identify more than one thousand individuals who lived in Arica between 2004 and 2008, and potentially for many more years. We restrict our analysis to this sample. We also utilize information from a large sample of intravenous tests collected from local residences in 2000. This provides us with individual-level measures of lead blood levels.

Anticipating our main results, we find that moving one kilometer (0.6 miles) away from the polluted site increases math scores by 0.10 standard deviations (controlling for age, gender, parental educational levels, and school characteristics). For language, the estimated effect is 0.13 standard deviations.[2] In terms of lead blood levels, our findings suggest that an increase of one microgram of lead per deciliter of blood reduces math PSU scores by 0.15 standard deviations (after controlling for age, gender, parental educational levels, and school characteristics). The corresponding effect for language scores is a reduction of 0.21 standard deviations.

Finally, we translate our estimated effects into monetary losses. We do this by exploiting the empirical relationship between PSU scores and annual earnings (during adulthood). Our estimates indicate that appropriate compensation would be approximately six thousand dollars per microgram of lead per deciliter of blood. This ignores any non-pecuniary costs associated with blood lead levels

---

[2]In both cases, the thought experiment involves increasing the distance from the place of residence to *site F* from 0.5 kilometers to 1.5 kilometers.

and is therefore likely a lower bound of the true cost. To the best of our knowledge, this is the first estimate of this kind in the literature.

The paper is organized as follows. In Section 2, we explain the origin of the lead pollution in the city of Arica. Section 3 reviews the literature on the effects of lead on health and the association between lead levels in blood and decreased cognitive function and academic performance. In section 4 we describe our data. Section 5 presents our identification strategy and empirical findings. Robustness checks are presented in Section 6. Section 7 concludes.

2   Exposure to Toxic Waste in Arica: A Brief History

The Arrival and Abandonment of Waste Material. PROMEL is a Chilean private company founded in the 1970s. The company operated in the north of Chile and offered secure transportation and processing services for mineral and chemical waste. On March 21st of 1984, PROMEL requested authorization from the Chilean National Border Agency to import mineral waste from Sweden. The request was based on the company's intention to reprocess these minerals to extract gold.[3] The chemical description of the minerals included in the official request indicated the presence of copper, gold, silver, iron, lead, mercury, and a number of other substances. With respect to the level of toxicity, PROMEL described the waste as being "non-toxic metallic mud unless eaten or handled without proper protection." After few months of deliberation, the Chilean government granted permission to import the minerals. As a result, between August 29th, 1984 and July 17th, 1985, Arica received three shipments of 8,096, 6,997 and 4,041 tons of minerals, respectively. The importation of these products continued until January 21st, 1989 (CONGRESS, 2009). By that time, approximately 21,091 tons of toxic waste had entered Chile. The collection point of the waste became known as *site F* (*sitio F*). The area was located northeast of downtown Arica, which at that point, was outside the city limits. PROMEL rented this site from the Bureau of Land Management (*Ministerio de Bienes Nacionales*). The area was initially secured by a concrete perimeter and was marked with warning signs indicating the presence of hazardous waste.

In December 1993, PROMEL requested permission to abandon the mineral waste.[4] The com-

---

[3]The complete process would have been involved cleaning the mineral wastes from arsenic, then converting the resulting material into trioxide and, finally, extracting gold.

[4]The request was directed to the National Border Agency. This agency owned the wastes after PROMEL decided not to pay the taxes on the importation.

pany argued that the high costs of inputs, the constant fluctuation of metal prices in international markets (particularly of cooper and silver), and the exchange rate policy made the originally profitable project nonviable.

The Housing Projects. During the 1980s the city of Arica faced an increasing demand for government subsidized housing. In 1985 the central government launched the construction of the first social housing project, which later became the neighborhood of *Los Industriales I*. It was located only meters away from *site F*. It is unknown why Chilean authorities decided to build at that location, but there was likely a lack of information about the characteristics of the materials contaminating the area.[5] Moreover, even if the local authority (*Municipalidad de Arica*) had wanted to stop or prevent the construction, this would have not been possible since the land was originally outside the city limits (in a non-residential area), and consequently, it was not under its jurisdiction. The neighborhood's location outside of the cities jurisdiction had another consequence on future residents. According to the Chilean legislation, house titles can only be granted by local governments. Therefore, since the housing projects were originally outside the city limits, the new home-owners did not receive official titles for several years. Only after 1996, once the city limits were extended, were house titles granted. As we discuss below, this prevented houses from being sold.

The construction of *Los Industriales I* (1985) was followed by *Los Industriales II* and *III* (1990), *Los Industriales IV* (1991) and *Cerro Chuño I-VII* (between 1993 and 1996). Table 1 presents the exact date of completion for each project. Figures 1 and 2 present satellite photographs of the area before and after the construction of the housing projects. They clearly illustrate the proximity of the neighborhoods to *site F*. Thus, hundreds of families were exposed, many for more than a decade, to the toxic waste.

Government Responses, Evaluation of Consequences and Compensations. In December of 1996, after multiple complains from the local residents, the Health Service of Arica began a series of actions aimed at removing the toxic waste from *site F*. In 1997 the department of chemistry at the University of Tarapaca (a local university) carried out the first scientific study of the levels of pollution in Arica. The results documented the presence of high concentrations of toxic minerals

---

[5]It is worth noting that the original decision was made by the central government, specifically by the Chilean Department of Housing (*Ministerio de Vivienda y Urbanismo*).

in the city.[6] These were confirmed by a second study conducted in the same year, prepared by the National Service of Geology and Mining (*Servicio Nacional de Minería y Geología*, SERNA-GEOMIN). The second study included a detailed analysis of the concentration of material in *site F*'s stockpiles. The analysis concluded that the material was approximately 10% of arsenic, 4.4% lead, 2.1% zinc, 1.5% copper, and also contained other heavy metals such as mercury, antimony, bismuth, and tin.

Given this evidence, the Chilean government declared the situation to be an "environmental health emergency." The re-location of the waste began in January of 1998, under the supervision of the Arica Department of Environmental Protection, and with the approval of the local government.[7]

Between July 14th and September 8th of 2000, the Chilean Ministry of Health carried out a third study, "Lead in *Cerro Chuño* and *Los Industriales*" (CONAMA, 2009). In this study, blood samples were taken from about 5000 residents. The tests showed high concentrations for 579 of the residents. As a result of these studies, residents of *Cerro Chuño* and *Los Industriales* initiated legal actions against PROMEL and the Arica Health Service, demanding financial compensation. In 2007, the Supreme Court upheld the claims, and decided on a compensation of 8 million pesos (approximately USD 16,000) per claim (Suprema, 2007).

Between August of 2008 and October of 2009, the National Environmental Commission (*Comisión Nacional de Medio Ambiente*, CONAMA) carried out a new environmental study specifically designed to identify sources of pollution in the city.[8] The results demonstrate the presence of three different sources of pollution: *site F*, the Arica-La Paz Line Railroad and a set of 26 points whose lead levels exceeded the international standard of 400 milligrams of lead per kilogram of soil analysis (400mg/Kg).[9] *Site F* was identified as the oldest source of lead contamination in Arica. The study also showed that both the rail line and the 26 points were associated with contemporary

---

[6]The study was motivated by the increasing complains of the residents of *Cerro Chuño* and *Los Industriales*, and it had the support of the Peace and Justice Service of Chile (*Servicio de Paz y Justicia de Chile*, SERPAJ).

[7]Some of the actions outlined in this plan were materialized through the re-location of the toxic wastes to *Quebrada Encantada*, an area located northeast, almost straight ahead, to *site F*. Figures 1 and 2 show the location *site F*, *Cerro Chuño*, *Los Industriales*, and *Quebrada Encantada*. Notice that the distance between houses located at the eastern edge of *Cerro Chuño* and the new point of collection of toxic waste is approximately 600 meters (or 0.375 miles).

[8]In order to establish the chemical composition of soil and subsoil of the city of Arica, the study followed a sampling scheme that considered the division of the city in 40 quadrants. For each quadrant, an average of 11 random samples were taken in 3 levels (surface, 15 cm and 30 cm below surface).

[9]The Environmental Protection Agency (EPA) established a risk value of 400mg/Kg lead to sites where children are active (Soil Hazard Standard for Children Play's Area, SHSCPA).

sources of contamination (Agriquem, 2009). Below we examine the relative importance of each of these sources of pollution in our sample.

Finally, and using as an input the report by CONAMA, the Department of Health established a multidisciplinary committee to investigate the effects of metal residue contamination in Arica and to develop policies leading to the relocation of the affected families. This committee focused on three highly contaminated areas: the harbor area (one of the 26 locations originally identified by CONAMA), *Maestranza* (site associated with the Arica-La Paz railroad system) and *site F*. They also considered two airstreams contributing to the dispersion of pollutants: south-southwest (predominant) and northeast. They found lead (and arsenic) not only in the air, but also Arica's drinking water. Among the main policy initiatives resulting from this analysis was the relocation of approximately 1,880 housing sites adjacent to *site F* and the development of special educational programs for the students attending schools in the affected areas.[10]

3 Literature Review

The toxicological effects of lead on human health have been widely documented in the literature. Depending on the dose and length of exposure, lead is associated with neurological, cardiovascular, kidney and liver damage, and can even cause death in very high concentrations (ATDSR, 2007). Although these effects date back to ancient times,[11] only in the last few decades has chronic exposure to lead raised greater concerns worldwide. These concerns are reflected in recent decision to amend the international standards of concentrations of lead in gasoline, paints, soil and air of urban areas, all of which are considered major sources of lead exposure (Mielke and Reagan, 1998; Mielke, 2002).

Notwithstanding the sustained reduction of lead levels in the population worldwide, the effects of lead on the deterioration of cognitive and behavioral functions, especially in children, is the main focus of research in a number of developed countries (Koller et al., 2004). Relative to adults, children are more susceptible to the toxicological effects of lead for (at least) three reasons. First, small children are more likely to put contaminated items (or any items) in their mouths. This increases the

---

[10] For details see *The Intervention in Areas with presence of polymetallic in Arica: A Master Plan* (CONAMA, 2009). During 2010 and 2011, the Ministry of Housing granted 898 housing subsidies and approved the construction of two housing projects to relocate 230 families of *Cerro Chuño* and *Los Industriales* (GORE, 2011).

[11] Possibly explaining even the fall of the Roman Empire (Lewis, 1985)

likelihood of ingesting lead from the environment via powder or through wall drawings (Lanphear et al., 2002; Tchernitchin et al., 2005; ATDSR, 2007). Second, children's digestive systems also have a higher degree of lead absorption (Ziegler et al., 1978; Tchernitchin et al., 2005; ATDSR, 2007). Thus, for a given amount of exposure young children are more likely to absorb harmful quantities. Finally, children's developing nervous systems increase their vulnerability to lead exposure because some functions of hormonal receptors, hormones, and neurotransmitters are altered, preventing the generation of beneficial substances for the organism (Lidsky and Schneider, 2003; Tchernitchin et al., 2005).

The effects of prolonged exposure to lead, especially when occurring at a young age, are irreversible. They can cause delay in bone growth, decreased intelligence and memory, and behavioral changes. Needleman et al. (1979) and Banks et al. (1997) document a negative relationship between blood lead levels and child IQ. Canfield et al. (2003) supports this relationship and also states that the deterioration of cognitive function of children could be manifested even under minimal exposure (below the threshold of $10\mu g/dL$ set by CDC). Using longitudinal data, Bellinger et al. (1992) find that low-level lead exposure at age 24 months is associated with a decrease of 5.8 points in the WISC-R Full Scale IQ test[12], 8.9 points in K-TEA Battery Composite[13] and significant reductions in intellectual performance at age 10 (considering an increase of $10\mu g/dL$). Wasserman et al. (2000) show that, depending on whether the exposure to lead occurs in the pre- or post-natal period, the decline in IQ can vary within the range of 1.07 to 2.82 points. Lanphear et al. (2000) document that even after controlling for gender, parental education and children's health status, there is strong evidence supporting the inverse relationship between levels of lead in blood and four scores associated with cognitive functioning tests. They estimate that for every extra micrograms per deciliter of blood of lead concentration, math and language scores decrease by 0.7 and 1 points, respectively.

---

[12]The Wechsler Intelligence Scale for Children-Revised (WISC-R) is a widely used individually administered intelligence test for children (Wechsler, 1974).

[13]The Kaufman Test of Educational Achievement (Kaufman and Kaufman, 1997).

Table A.1 in the appendix presents the effects of lead on IQ documented in the literature. [14],[15],[16]

The literature also examines the effects of lead exposure on violent behavior and mortality. Wolpaw (2007), using a reduction of the lead content in gasoline introduced by the United States Clean Air Act of 1970, studies the relationship between lead exposure and crime. Her results indicate a significant relationship between lead exposure during childhood and violent crime behavior later in life. The estimates show that the elasticity of violent crime and childhood lead exposure is approximately 0.8. This suggests that the reduction in lead exposure between 1970 and 1980 accounts for more than the 50% of the decrease in crime during 1990s (1992-2002). Clay et al. (2006) analyze the effect of lead water pipes on child mortality. Their results suggest that approximately 13 infants per 1,000 children died in 1900 as a consequence of lead pipes. This is staggering when compared to the 5.7 deaths per 1,000 children from *all causes* in 2000. Using variations in water composition of different types of water delivery system across cities, Clay et al. (2010) extend their previous analysis to study the effect of water-born lead exposure on child mortality and worker productivity between 1900 and 1920. Exploiting the chemical relationship between water lead content in pipes and its level of acidity (measured by the water's pH), the authors find that an increase in pH from 6.7 to 7.5 was responsible for a decrease of 12.43% in infant mortality between 1900 and 1920 in cities with lead pipes. They also find that manufacturing workers living in cities with lead pipes have a lower salary and suggest this was caused by the lower productivity resulting from lead exposure.

Tchernitchin et al. (2005) review the main sources of human exposure to lead in Chile. The study identifies gasoline and lead-based paint as the most common sources of lead exposure. It also identifies the cities of Arica and Antofagasta as *examples* of cities with high levels of lead contamination. The study also reviews two previous small-scale studies examining the concentrations of lead in the blood of children exposed to pollution in Arica. The first study analyzes the results of 32

---

[14]In adults, the effects of exposure to lead may be partially reversible after removal of the source of contamination, decreasing levels of blood lead and after application of chelation treatments (Tchernitchin et al., 2005).

[15]There is broad consensus on the effects of lead at high concentrations. However, the effects at low exposures are still subject of debate (Bellinger et al., 1992; Lanphear et al., 2000). Kaufman (2001) review a series of articles studying the relationship between low levels of blood lead and decreased IQ. The main criticisms are that there is no adequate control for variables that could influence a child's IQ (household socioeconomic status, parenting skills, childhood diseases, parental IQ, among others), and that the evidence usually examines only the most extreme cases omitting intermediate cases. According to Kaufman (2001), due to these considerations, the link between low-level lead concentrations and IQ must be assessed with caution.

[16]For a deeper discussion of the methodological issues in this literature, see Koller et al. (2004).

blood samples collected in 1998 from children with residences in *Cerro Chuño* and *Los Industriales* (located within 200 meters of *site F*). Nine samples from children living in *Villa Santa María* were initially used as control group (this village is located approximately 1.7 kilometers from *site F*). However, the results showed high levels of lead in all three neighborhoods, potentially as a result of the small sample size in the control areas. A second study extended the original sample. In this second study, children living in *Pampa Nueva* (a village located 4 kilometers from *site F*) were chosen as a control group. In addition to the blood sample tests, the study included hair samples and data on tests of motor responses to visual stimulus. The results confirmed the high levels of lead found in 1998 and also documented poor motor responses to visual stimulus even among children with levels of lead below $10\mu g/dL$.[17]

4   The data

In order to estimate the effect of early lead exposure on academic achievement we use several sources of information.

Blood samples. We use data from the study "Lead in *Cerro Chuño* and Los *Industriales*" (LCCI), carried out by the Chilean Ministry of Health in 2000 to get blood lead level concentrations. This study included the collection of 4,990 intravenous blood tests from local residents: 3,240 children (aged 0 to 15 years) and 1,750 adults (aged 16 or older). The results show that 4,411 people had concentrations of lead of 10 micrograms per deciliter of blood or less, the maximum level of lead allowed by the U.S. Centers for Disease Control and Prevention (CDC).[18] For the remaining 579 cases the study found levels exceeding the maximum standards. Of these, 538 were re-examined by the Institute of Public Health through the *leadcare* system.[19] This new examination found that a total of 407 samples had concentrations of lead below the CDC standard, while 131

---

[17]Coria et al. (2009) study the long-term consequences of lead poisoning on children from another unfortunate experience of lead exposure that occurred in 1996 in the city of Ñuble (south of Chile). The authors combined historical records of lead concentration for 77 people 10-25 years old, scores from intelligence tests, and the results from physical examinations taken place ten years after exposure. The findings showed that 21% of the sample had subnormal levels of IQ, and the lowest among individuals whose lead exposure occurred prior to the age of six. In particular, the psychometric results show lower performance on tests of arithmetic reasoning, which revealed that the lead was compromising frontal lobe functions.

[18]In the absence of substantial changes in lead exposure blood lead levels in adults are stable and for their environmental monitoring a single blood lead concentration is an excellent biological indicator (Delves et al., 1984).

[19]The final report indicated that the remaining 41 individuals were contacted multiple times (including attempts via home visits), but they did not attended the examination.

had levels at or above the standard (120 were children between ages 0 to 15 years). Importantly, the data contains information on each individual's street address. This allows us to link lead levels to distance to *site F*. Figure 3 presents the distributions of lead levels in blood for the two examinations.[20]

Test scores. We use data from two national standardized tests: the SIMCE (*Sistema de Medición de la Calidad de la Educación*) and the Chilean college admission tests or PSU (*Prueba de Selección Universitaria*).

The SIMCE (language and math) tests are aimed at measuring the quality of teaching and the degree of student learning at various educational levels in Chile. The tests are taken annually by students enrolled in 4th, 8th, and 10th grade.[21] The system started in 1988 so we are able to identify schools near *site F* at the time the chemical waste was brought to the country. As discussed below, this allows us to analyze test scores before and after students were exposed to lead pollution from *site F* (fourth graders in 1988 were born in 1978 or 1979). Unfortunately, from 1988 to 1997 the SIMCE data is available at the school level only. This is not the case for the data from the college admission test system (PSU). Therefore, we perform our main analysis using individual level data from the 2008 college admission system.

The 2008 PSU data contain information on individual characteristics (age, gender), individual-level math and language test scores, family background variables (mother's and father's education, family income, mother or father absence), high school-level information (private, voucher or public), and residential address. A large proportion of the sample has just graduated from high school (18 year old on average). For this group we gather information from the 2004 SIMCE test scores when they were enrolled in 8th grade. This allows us to observe multiple measures of academic achievement, and also to identify each individual's place of residence (municipality) in 2004 and 2008.[22] We restrict our main analysis to the sample of individuals who reported living in Arica in both years.

Descriptive statistics for this sample are presented in Table 2. We use the data from those not

---

[20]Table A.2 in appendix presents the descriptive statistics of this sample.

[21]Since 2006, SIMCE tests 4th graders and alternates between 8th and 10th graders every other year.

[22]From the 2004 SIMCE data we extract the exact location of individual's high school. We use this information to analyze the impact of pollution on test scores as a function of the distance from *site F* to high schools. Unfortunately, the information on individual's place of residence is only available in 2008 PSU.

living in Arica in either 2004 or 2008 to perform robustness checks.[23]

Labor market outcomes. Lastly, we use administrative records from the Chilean Unemployment Insurance (UI) system to analyze the impact of pollution on adult earnings. The UI data contain information on labor income and individual labor market histories in the formal sector (approximately 80% of the employment in the country) on a monthly basis since November 2002. We merge individual-level data on earnings with PSU test scores. As described below, we use the estimated relationship between earnings and PSU scores to compute the appropriate monetary compensation for the irreversible health damage and loss of productivity associated with early lead exposure. Given that a sizable fraction of individuals who took the PSU in 2008 do not show up in the UI system - many are still enrolled in tertiary institutions- we use information from the 2004 PSU and the 2010-2011 UI data. Hence, most individuals enrolled at a higher education institutions in 2004 were working in 2011, and the information about their wages is reported in the UI data. We choose the 2004 PSU because the 2004 version is the first year of the PSU. This ensures the comparability of our results.

Table 3 presents summary statistics for the sample of individuals who took the PSU in 2004 and for whom we have data on monthly earnings between 2010 and 2011.

5 Empirical Analysis

5.1 Identifying Sources of Pollution

Over the last 20 years the population of Arica has been exposed to multiple sources of pollution. In order to identify the main source of pollution in the city, we first evaluate the importance of all of the sources of pollution originally identified by CONAMA: *site F*, Arica-La Paz line railroad and a set of 26 points whose lead levels exceeded the international standard of 400 milligrams of lead per kilogram of soil (400mg/Kg) (see CONAMA, 2009, for further details). Since current levels of lead at each location are not available, we establish their relative importance by analyzing the relationship between the lead levels in blood from the sample of residents of *Cerro Chuño* and *Los*

---

[23]The sample of students in both 2004 SIMCE and 2008 PSU datasets are not exactly the same. We do not observe PSU scores for those who took the SIMCE in 2004 and did not take the PSU in 2008. Therefore, we do not observe the PSU scores nor the place of residence in 2008 for those who retained a course during high school, dropped out of high school or for those high school graduates who opted to not take the PSU in 2008. The 2004 SIMCE data contains 272,198 observations whereas the 2008 PSU contains information for 240,682 high school graduates.

Industriales and the distance between their places of residence and the different locations.[24] As previously explained, the original sample of blood tests was collected in 2000 from 4,990 residents of *Cerro Chuño* and *Los Industriales* (Servicio de Salud de Arica, 2001). We focus our analysis on the 579 individuals whose blood lead levels exceeded the risk threshold defined by the CDC ($10\mu g/dL$). As previously explained, for them we have data from two different independent blood lead level measures.

Using this data, we evaluate which of the three sources of lead exposure is most highly correlated with the blood lead levels observed among individuals living in *Cerro Chuño* and *Los Industriales* by 2000. We estimate linear regression models of individual's blood lead levels on (ln) distance from home to each of the three sources of contamination.[25] We include as additional controls in these regressions gender and age (at time of testing). Finally, in order to reduce the measurement error we use the average blood lead level of the two samples.[26]

Thus, for each of the sources of pollution we estimate a regression model of the form:

$$Pb_i = \alpha_0 + \alpha_1 \ln D_{home,i} + \alpha_2 Gender_i + \alpha_3 \text{Age}_i + \eta_i, \tag{1}$$

where Pb_i is the blood lead level associated with individual i (measured in micrograms per deciliter of blood, $\mu g/dL$), $D_{home,i}$ is the distance between individual's i place of residence and the location of the respective contaminated area in Arica.

Table 4 presents the estimated coefficients for our measures of exposure to lead and the three

---

[24]We follow (Sinnott, 1984) and generate the distance between two locations A and B as follows. Let A and B be two different locations, with associated latitudes and longitudes $(\text{lat}(A), \text{lon}(A))$ and $(\text{lat}(B), \text{lon}(B))$, respectively. We compute the distance between A and B ($d(A, B)$) from:

$$\tau(A, B) = \left(\frac{\sin(\text{lat}(B) - \text{lat}(A))}{2}\right)^2 + \cos(\text{lat}(A)) \times \cos(\text{lat}(B)) \times \left(\sin\frac{(\text{lon}(B) - \text{lon}(A))}{2}\right)^2$$

$$d(A, B) = 6367 \times (2 \times \arctan(\sqrt{\tau(A, B)}, \sqrt{1 - \tau(A, B)}))$$

where $d(A, B)$ is expressed in kilometers, $\arctan(x, y)$ is the arctangent of x/y, and 6,367 is the radius of the earth expressed in kilometers.

[25]Distances were obtained as follows. We first define 33 centroids of pollution covering the full extent of the railway. Similarly, we define 26 points of contamination in soil. The distance from household to railway is defined as the minimum distance between household and the 33 centroids. Likewise, distance from household to leaded soil is calculated as the minimum distance of each household to any of the 26 centroids. In the case of *site F*, we consider *Chapiquiña 3066* as reference address. This corresponds to the place rented by PROMEL to carry out the tasks of extracting gold and silver from the chemical wastes. GoogleEarth did not recognize some of the addresses in the data (mostly rural), and the sample size was reduced to 390 tests.

[26]Our results are robust to the use of either sample (ISP or SERMUS). Results are available from the authors upon request.

measures of distances.[27] The results presented in column (1) indicate a significant correlation between concentration of lead in blood and distance to *site F*. Specifically, we estimate that increasing the distance in one kilometer (0.6 miles), from 0.16 to 1.16 kilometers to *site F*, would imply a reduction of 1.26 micrograms of lead per deciliter of blood (or 0.413 standard deviation).[28] Thus, the evidence confirms the association between distance to *site F* and blood lead levels. As columns (2) and (3) show, the coefficients associated with distances to the railway and *leaded soil* sites have the opposite sign from expected and they are not significant.

The results are conclusive in pointing to *site F* as the main source of lead pollution in the city of Arica. Consequently, we focus our analysis on this source.

5.2 Distance to *Site F* and Academic Achievement

We first analyze school-level data for different years from SIMCE data sets. Figure 4 presents the (standardized) school-level average test scores from 4th graders enrolled in schools located less than 1.5 kilometers from *site F*, in schools located between 1.5 and 3 kilometers from *site F*, and in schools located 3 or more kilometers away from *site F*. We present the evolution of the test scores from 1988 to 2010.

The results show a clear pattern. At the time the chemical waste arrived to the area (1988), schools near *site F* had a better performance than those located between 1.5 and 3 kilometers from the site. Just four years after the arrival of the material, average test scores for those schools start reducing dramatically, potentially because of the prolonged exposure to lead. Table 5 confirms these findings. Here we show raw school-level averages on math and language PSU tests as a function of each school's distance to *site F* for the period 1988-2010. As in the previous figure, the evidence does not suggest any relationship between distance from the school to *site F* and scores between 1988 and 1992. However, we find that after 1992 schools located closer to *site F* (i.e at less than 1 kilometer) have lower math and language test scores than schools located further away

---

[27]To avoid influential observations, we restrict our sample to those individuals whose blood lead levels were lower than the 99th percentile using Cook's distance (Cook, 1979).

[28]The non-linear association between Pb_i and $D_{home,i}$ implies that the magnitudes of the effects depend on the original distance to *site F*. For example, increasing the distance from 0.66 to 1.66 kilometers to *site F* would imply a reduction of 0.587 micrograms of lead per deciliter of blood (or 0.193 standard deviation). Likewise, increasing the distance from 1.3 to 2.3 kilometers would imply a reduction of 0.363 micrograms of lead per deciliter of blood (or 0.119 standard deviation).

from the site.[29,30] However, these strong patterns are obtained using aggregate data and without controlling for confounding factors. We utilize individual level data from PSU to deal with these potential concerns.[31]

Let PSU_i denote individual's i PSU test score and $D_{home,i}$ the distance between student's place of residence and *site F*. We estimate the following model:

$$PSU_i = \gamma_0 + \gamma_1 \ln D_{home,i} + \gamma_2 Age + \gamma_3 Gender + \gamma X_i + \gamma_5 Z_{school,i} + \eta_i, \tag{2}$$

where X_i represents a set of individual-specific exogenous controls (family background, and variables controlling for the direction of wind), $Z_{school,i}$ is a vector of variables describing the characteristics of the high school from which the individual graduated from, and η_i is an idiosyncratic component.[32] Table 6 presents the results from estimating equation (2) for math and reading PSU test scores under different specifications.[33] The evidence shows that, even after controlling for family background, high school characteristics, and geographic factors, there is a strong relationship between PSU scores and distance from student's place of residence to *site F*. Our preferred specifications (columns (5) and (6), imply that increasing the distance one kilometer (starting at 100 meters from the site) would increase language and math test scores in approximately 29.55 (0.31 standard deviations) and 21.06 (0.22 standard deviations) points, respectively. The nonlinear model implies decreasing effects as we move away from the pollution site. These magnitudes are robust to different specifications.[34] The larger effects for language when compared to math suggest that lead exposure might have a higher negative impact on brain functions associated with language skills than arithmetic and math skills.[35]

---

[29]We use the school-level data on test scores ($T_j(t)$) and distance from school to *site F* ($Dist_j$), to estimate the model $T_j(t) = \alpha(t) + \beta \ln Dist_j + \nu_j(t)$, where t denotes year. We obtain $\hat{\beta}$=0.19 (t=2.21). The results from the alternative model $T_j(t) = \alpha(t) + \beta_0 \ln Dist_j + \beta_1 \ln Dist_j \times D_{t>1996} + \nu_j(t)$, where $D_{t>1996}$ is a dummy variable that takes a value of one if t is larger than 1996, delivers $\hat{beta}_0$=-0.008 (t=-0.06) and $\hat{beta}_1$=0.36 (t=2.04). These results confirm the patterns observed in Figure 4 and Table 5.

[30]Since 1999 SIMCE test scores changed their scales. This explain the difference in levels between 1988-1996 and 1999-2010.

[31]Furthermore, since the LCCI data does not contain information on the school attended, we could not implement our empirical strategy using only school-level data.

[32]Currie and Walker (2011) use a similar reduced form specification to study the link between health outcomes and distance to a source of pollution.

[33]In order to restrict our analysis to individuals living in Arica, we restrict the PSU sample to those individuals reporting a home street address at most four kilometers away from the *site F*. This is consistent with the maximum distance from *site F* to the edge of the city.

[34]These results are available upon request from the authors.

[35]An obvious concern regarding the previous regression model is that our measure of distance $D_{home,i}$ might

5.3 Blood levels and academic achievement

We now assess the impact of blood lead levels on individual test scores. Our identification strategy is similar to the Two Sample Instrumental Variables (TSIV) approach developed by Angrist and Krueger (1992). Specifically, consider the following system of equations:

$$PSU_i = \beta_0 + \beta_1 Pb_i + \beta_2 X_i + \beta_3 Z_{school,i} + \varepsilon_i, \tag{3}$$

$$Pb_i = \alpha_0 + \alpha_1 \ln D_{home,i} + \alpha_2 Age_i + \alpha_3 Gender_i + \eta_i. \tag{4}$$

We interpret equation (3) as the structural equation describing the effect of blood lead levels, Pb_i, on individual's test scores, PSU_i, whereas equation (4) represents our model for blood lead levels as a function of distance to the source of pollution. As discussed above, we can use the LCCI data to estimate equation (4). These results were reported in Table 4. However, we do not have blood lead levels (Pb_i) and test scores (PSU_i) for the same sample of individuals, so equation (3) cannot be estimated directly. This is where TSIV strategy is useful. By substituting expression (4) into (3) we can derive the reduce form equation (2), linking PSU scores to distance to *site F*. In this context, it is straightforward to show that γ_1 in equation (2) is equal to ($\beta_1 \times \alpha_1$). Thus, and given that the model is exactly identified, we can estimate β_1 from equations (2) and (4) as:

$$\hat{\beta}_{1,TSIV} = \frac{\hat{\gamma}_1}{\hat{\alpha}_1}.$$

By combining the results from equations (2) (reduced form model) and (4) (first stage), we can obtain the (structural) effect of lead exposure on academic achievement.[36] We obtain standard errors using the Delta Method.[37]

---

be endogenously determined. In principle, families could have decided to live far from the *site F* depending on their characteristics, leading to biased estimates for the effect of distance on test scores. In order to investigate the endogeneity of $D_{home,i}$, we empirically investigate the association between our distance variable and family background variables. Table A.3 presents the results. We find that parental education is positively correlated with distance from home to *site F*, so our results suggest that more educated parents have avoided living close to *site F*. In the estimation of equation 2 we account for this by including parental education in the set of controls.

[36] Our non-linear specifications are motivated by the estimation of non- and semi-parametric models of the relationship between PSU test scores and distance from place of residence to *site F*. Figures A.1 and A.2 in the appendix present the results for math and language scores, respectively. The results suggest that the natural logarithm of the distance variable provides a good description of the data. Fortunately, this specification also allows us to directly implement the TSIV strategy.

[37] Assuming that $cov(\hat{\gamma}_1, \hat{\alpha}_1) = 0$, the variance of $\hat{\beta}_{TSIV}$ is giving by the following expression: $Var(\hat{\beta}_{TSIV}) = Var(\hat{\gamma}_1)/\hat{\alpha}_1^2 + (\hat{\gamma}_1^2/\hat{\alpha}_1^4) \times Var(\hat{\alpha}_1)$.

Table 7 presents the estimated effect of lead blood level on math and language test scores. The estimated $\hat{\beta}_{1,TSIV}$ implies a reduction of 13.81 (or 0.15 standard deviations) points for one microgram of lead per deciliter of blood for math and 19.37 (or 0.21 standard deviations) points for one microgram of lead per deciliter of blood for language. Overall our results suggest negative and large effects of proximity to the pollution site on test scores.

Although we find significant results for $\hat{\gamma}_1$ and $\hat{\alpha}_1$, as expected, we lose power when computing the ratio. Nevertheless, we obtain significant results (at the 10% percent level) in the case of language test scores. This finding is in line with the results for the effect of lead on verbal IQ as reported in Needleman et al. (1979).[38]

Caveats. A natural threat to our identification strategy is the possibility of endogenous migration. Once alerted to the presence of the chemical waste, families could have decided to move away from the vicinity of *site F* and even to out-migrate from the city. Other related threat is endogenous sorting of educated parents that may decide to change their kids from schools near *site F* to others away from it. In what follows, we demonstrate that these might not be serious concerns.

Table 8 presents the proportion of parents with college education with children enrolled in schools in Arica. The numbers are obtained from the SIMCE records. We present these numbers by school distance to *site F* for 1989, 1999 and 2008. The proportion of parents with college education for schools located 1.5 kilometer or less from *site F* is very stable in almost 20 years, suggesting that endogenous migration as a function of parental education did not occur during the period. On the other hand, we observe a sharp increase in the fraction of college educated parents with kids enrolled in schools located between 1.5 and 3 kilometers from the site. However, most of the increase can be explained by a single school. [39] For schools located more than 3 kilometers away from *site F* there is a less pronounced increase which agrees with the national trend during

---

[38]Also, we perform linear regression models to study the effect of the school's proximity to *site F* on the academic achievement of the students of Arica using SIMCE scores between 2003 and 2006. As we previously described, in SIMCE we only have the distance from schools to *site F*. Tables A.4 and A.5 in the appendix present descriptive statistics of SIMCE 2003-2006 scores and the results from the linear regression models, respectively. The finding confirm our main results. Even after controlling for gender, parental education, school characteristics, we obtain positive effects from (natural logarithm) distance between the school and *site F* on school-level math and language test scores.

[39]Colegio Arica College explains most of the increase in parental education between 1999 and 2008. If we exclude this school from this calculation, the distribution of parents with college education with children enrolled in schools between 1.5 and 3 kilometers from *site F* increases by 1.3 percentage point (instead of 10) during the period.

the period. These results suggest endogenous migrations might not be a serious threat to our identification strategy.

Figure 5 presents further evidence along these lines. It shows the trends of the distributions of enrollment by school distance to *site F* for fourth graders between 1988 y 2008 in Arica. Again, we observe that the proportions are stable over time. Figure A.3 in the appendix presents similar trends but for the distributions of elementary schools. The results confirm the stability of enrollment.

These results are consistent with the literature documenting little regional migration in Chile. Soto and Torche (2004) show that on average the regional migration rate for the 1965-2000 period was 0.6% per year, which is especially low considering the existence of factors that could promote greater regional movements (homogenous population across regions and high rates of urbanization). Also, data from the CASEN 2006 survey[40] shows that 86.17% of respondents aged 17 or younger who lived in Arica reported that they were born in Arica. Likewise, among the same individuals, 96% reported they were living in Arica in 2002. On the other hand, the literature shows that residential mobility is largely explained by schooling and employment decisions (Leslie and Richardson, 1961). For Chile, Ruiz-Tagle and Sehnbruch (2009) estimate that the average job spell for a formal worker is about 5 years. Therefore, if we believe that residential mobility is related to employment opportunities, on average, families should live in the same place for at least 5 years, providing extra support for our approach.

Institutional features of the social housing system also support our empirical strategy. Specifically, the 18th article of the Supreme Decree on "The Regulation on Social Housings" (MINVU, 1990), establishes that houses in *Cerro Chuño* and *Los Industriales* could not be sold before 5 years after the inscription date of the house in the registry of deeds (*Conservador de Bienes Raíces*). As mentioned in Section 2, since the housing projects were originally outside the city limits, the new home-owners did not receive official titles for several years. In fact, only after 1996, once the city limits were extended, were house titles granted.[41] Therefore, we can assume that the families and their children were exposed to lead for a long period of time, although the possibility remains that they could have rented their houses and moved to other places in order to avoid lead contamination.

---

[40]CASEN (*Socioeconomic Characterization Survey*) 2006 corresponds to a cross-sectional household survey representative of the population in 2006.

[41]The law establishes that if the owners want to sell the house before the mandatory deadline, they have to request authorization from the Housing Service (*Servicio de Vivienda y Urbanismo*, SERVIU). This rule increases the probability that those families lived for at least 5 years near the toxic waste.

Therefore, even though we cannot rule it out, endogenous migration does not seem to pose a serious threat to our identification strategy. Nevertheless, we control for it by analyzing residents of Arica in 2004 and 2008. In this context, our analysis suggests that four years of lead exposure is enough time to observe its long-term effects on academic performance, and potentially, they can be extrapolated to different setups.

5.4 Blood lead levels and earnings: Estimating compensations

The literature documenting the monetary consequences of lead exposure at the individual level is scarce. Landrigan et al. (2002) quantify the IQ loss in children exposed to lead and link this loss to lower revenues for future generations.[42] We extend this literature by documenting the effects of early lead exposure on future earnings at a more disaggregate level.

After obtaining the effect of blood lead level on PSU scores for students in Arica, we can compute the effect on earnings by linking labor market outcomes and PSU scores for a different but similar sample. This is a significant contribution to the public policies associated with determining compensation for the serious (and irreversible) damage to the health of residents of *Cerro Chuño* and *Los Industriales*. It also provides a comparison to the actual monetary compensations paid to a fraction of the population over the past few years. In 2007 the Supreme Court ruled that there should be a compensation payment of CLP 8 million (USD 16,000) per claimant made by the Arica Health Service to 365 residents of *Cerro Chuño* and *Los Industriales* (Suprema, 2007).

We use information on monthly earnings from the Unemployment Insurance Database (UI) and PSU scores to estimate the following Mincer-type regression model:

$$\ln w_i \;\;=\;\; \beta_0 + \beta_1 PSU_{i,math} + \beta_2 PSU_{i,lang} + \beta_3 Q_i + \epsilon_i, \tag{5}$$

where w_i corresponds to the average monthly wage obtained by individual i between 2010 and

---

[42]Landrigan et al. (2002) estimate the annual costs of the incidence, prevalence, mortality and pediatric assistance produced by environmental pollution on the health of American children. Their results show that the annual costs because of children lead poisoning in the U.S. correspond to USD 43.4 billion, a figure higher than the estimated annual cost of asthma (USD 2 billion), cancer (USD 0.3 billion) and mental disability (USD 9.2 billion). Muennig and Bao (2009), extends the estimated costs of lead exposure in children in New Jersey, including special education costs, direct medical costs associated with treatment and hospitalization, social costs for juvenile and adult crime and dropout graduate rates. The results indicate that if blood lead levels in a single cohort of children between 0 and 6 years were reduced to less than 1 microgram per deciliter of blood, the contribution to the society of New Jersey would be at least USD 27 billion.

2011 and $PSU_{i,j}$ denotes his 2004 PSU test scores (j={$math, language$}).[43] We also include a vector of observable characteristics, Q_i, which include gender, age, age squared, parent education, high-school characteristics and regional dummies. Table 9 presents our results. We find positive and significant effects of math and language scores on monthly earnings. Our estimates suggest that each additional point in math test score increases earnings approximately by 0.17 percent. In the case of language scores, the effect is 0.01 percent.

In order to provide a better understanding of the magnitudes of these effects, we compute the partial derivative of monthly earnings (in levels) with respect to PSU scores.[44] Our results, presented in Table 10, indicate that one extra point on the PSU increases monthly earnings by 745 and 60 Chilean pesos (or USD 1.49 and 0.12) for math and language PSU scores, respectively.

In order to compute the individual-level monetary impact of lead exposure, we combine our previous results with the marginal effect of lead on PSU test scores. Formally, we compute:

$$\frac{\partial W}{\partial Pb} = \frac{\partial W}{\partial PSU_{lang}} \times \frac{\partial PSU_{lang}}{\partial Pb} + \frac{\partial W}{\partial PSU_{math}} \times \frac{\partial PSU_{math}}{\partial Pb}. \tag{6}$$

Table 11 presents these partial derivatives obtained using the results in Table 7 ($\partial PSU/\partial Pb$) and Table 10 ($\partial W/\partial PSU$). We estimate that for each extra microgram of lead, monthly earnings decrease by 11,458 Chilean pesos (or USD 22.92). This represents a 2.9% decrease in average monthly wages (of the average monthly earnings for our sample: 395,447 pesos per month).[45]

We extend this analysis to study the effects of proximity to a source of pollution on individual earnings. In particular, we analyze the effect of the distance from home to *site F* on average monthly earnings between 2009 and 2010 using the following expression:

$$\frac{\partial W}{\partial D_{home}} = \frac{\frac{\partial W}{\partial PSU_{lang}} \times \frac{\partial PSU_{lang}}{\partial \ln(D_{home})}}{D} + \frac{\frac{\partial W}{\partial PSU_{math}} \times \frac{\partial PSU_{math}}{\partial \ln(D_{home})}}{D}. \tag{7}$$

---

[43]As it was explained in Section 4, we use data for the first version of the PSU (2004) linked with UI wage data because a sizable fraction of students who took the PSU in 2008 are still enrolled in tertiary institutions in 2010-2011 and, thus, not participating in the labor market.

[44]The marginal effect of an increase on PSU score on the monthly earnings is given by $\hat{\beta}_1 \times \exp(\hat{\beta}_0 + \hat{\beta}_1 PS\bar{U}_{math} + \hat{\beta}_2 PS\bar{U}_{lang} + \hat{\beta}_3 \bar{Q} + \hat{\sigma}^2/2)$, where we evaluate the regressors at the respective sample means.

[45]We work with a nationwide sample to estimate the parameters for the Mincer equation. For consistency purposes, we estimate the effects of lead on earnings using the whole sample. We performed the analysis using a subsample including only Arica finding very similar results. Indeed, the average monthly wage of the subsample of Arica is CLP 381,938, very close to the national wage.

Using the estimated coefficients from equations (2) and (5) (presented in Tables 6 and 9, respectively), we obtain the effect of an increase of one kilometer in the distance from home to *site F* for different values of the initial distance. The estimated coefficients of this relationship are presented in Table 12. Specifically, the first row shows that for each kilometer increase in distance from a base of 0.1 km, individual monthly earnings increase by US$ 145.75. This corresponds to a 18.4% increase in monthly earnings. The impact decreases for individuals located far away from *site F*.

Finally, we use our results to estimate the losses in terms of lifetime earnings (present discounted value) due to lead exposure. Using the results presented in Table 11 we simulate the stream of earnings of an individual affected by lead exposure and compare it with the simulated profile of a representative individual not exposed to the sources of pollution.[46] We find a difference of CLP$ 3 million (or US$ 6,000 dollars) for each microgram of lead per deciliter of blood. At an average of 10 $\mu g/dL$ in the sample, the mean compensation should have been about 3.8 times the amount dictated by the Supreme Court in 2005.

6 Robustness Checks

Although we are controlling for a comprehensive set of family background variables in the reduced form equation, the potential correlation between the residential location decision and unobserved variables could compromise the validity of distance as an instrument for blood lead levels.

We attempt to tackle this issue by implementing the following strategy. First, we isolate the impact of family background, individual and school characteristics on PSU scores using a large sample of individuals living in areas not affected by lead pollution. More precisely, from a sample of individuals residing in three provinces not affected by lead pollution (*Parinacota*, *Iquique*, and *Tamarugal*), we estimate the following version of our reduced form model:

$$PSU_i = \gamma_0 + \gamma_2 Age_i + \gamma_3 Gender_i + \gamma_4 X_i + \gamma_5 Z_{school,i} + \eta_i. \tag{8}$$

The estimated parameters are consistent estimates of the association between the exogenous variables and PSU scores. Obviously, we do not include $\gamma_1 \ln D_{home,i}$ since the regression is estimated

---

[46]The predicted lifetime profile of annual earnings is generated from Mincer-type regressions estimated using the UI data. We assumed a discount rate of 4%.

with data from provinces excluding Arica, and thus, far away from *site F*.

Second, using the estimated parameters from equation (8) and the sample of individuals exposed to pollution in Arica, we construct a predicted value for the PSU, $P\hat{S}U_{-D}$, as follows:

$$P\hat{S}U_{-D} = \hat{\gamma}_0 + \hat{\gamma}_2 Age + \hat{\gamma}_3 Gender + \hat{\gamma}_4 X_i + \hat{\gamma}_5 Z_{school,i}.$$

Thus, $P\hat{S}U_{-D} = P\hat{S}U_{-D}(Age, Gender, X, Z)$ captures the effect of family background, individual and school characteristics on PSU. Since, it is constructed from estimated parameters obtained from a large sample of individuals not exposed to pollution, we avoid the consequences of the potential biases due to the endogeneity of distance to the *site F*.

Finally, we estimate the reduced form model of PSU on $P\hat{S}U_{-D}$ and distance from home to *site F* for the sample of students exposed to lead pollution:

$$PSU = \gamma_0 + \gamma_1 \ln D_{home} + \gamma_2 P\hat{S}U_{-D} + \varepsilon. \tag{9}$$

Note that if we follow a Constrained Least Squares (CLS) approach, by fixing γ_2, we avoid any residual correlation between the family background and unobservables determining the PSU test scores. Additionally, note that by restricting $\gamma_2 = 1$, under the null hypothesis of no correlation between family background and distance from home to *site F*, the coefficient γ_1 should be comparable to the one obtained in Table 6.

Table 13 presents the results of this exercise. Our findings are consistent with those of Table Table 6. For math scores, we find a point estimate that is very similar to the effect of distance on PSU score reported in Table 6 (but with lower statistical power). Similarly, for language test scores we find comparable results to those presented in Table 7 (significant at 1% level).

In Table 14 we perform a sensitivity analysis constraining γ_2 at different values. In practice, we vary γ_2 from 0.8 to 1.2. Thus, by fixing γ_2 and estimating by CLS, we avoid any residual correlation between family background and unobservables determining PSU score. Interestingly, even when the family background parameter is set to 1.2, i.e. when the role of family background is "amplified", the estimated effects of distance to *site F* are consistent with the results presented in Table 6.

We also analyze the robustness of our results analyzing the association between distance to *site*

F and test scores for those individuals who took PSU 2008 in Arica but SIMCE 2004 in another region and for those who took SIMCE 2004 in Arica but PSU 2008 in another place. In the first case we obtain positive but non-significant effects, whereas for the second group we obtain ambiguous signs and non-significant effects. These results are presented in Tables A.7 and A.6 in the appendix.

7   Conclusions

In this paper we estimate the long term effects of early lead exposure on academic performance and labor market outcomes. We exploit a massive and unfortunate lead intoxication incident that occurred in the city of Arica in northern Chile. Between 1984 and 1989 the Chilean company PROMEL imported 20,000 tons of chemical waste to extract gold and silver. Years later, PROMEL decided not to continue with the project and abandoned the waste in an area known as *site F*. In the early 1990s, the neighborhoods of *Cerro Chuño* and *Los Industriales* were built a few meters from *site F*. The residents of these neighborhoods became exposed to abnormally high levels of chemical waste.

We find a strong empirical relationship between blood lead levels and student academic performance. Specifically, we find that an increase of one microgram of lead per deciliter of blood reduces math and language scores by 0.15 and 0.21 standard deviations, respectively. These results persist even after controlling for individual and school characteristics. We also show that endogenous migration is not a serious threat to our identification strategy.

Using administrative data on labor market outcomes, we also investigate the impact of lead exposure on future earnings. The results suggest that an increase of one microgram of lead per deciliter of blood results in an average reduction of 2.9% in annual earnings during adulthood. Comparing outcomes by residential proximity to *site F*, we find that a one kilometer increase in the distance between home and Site F increases individual monthly earnings by USD 145.75. This corresponds to a 18.4% increase in monthly earnings.

Finally, we calculate the effect of lead exposure on the net present value of earnings. Our results suggest a decrease of approximately USD 6,000 per microgram of lead per deciliter of blood. This evidence should inform policy makers when determining appropriate compensation for individuals exposed to events such as the one observed in Arica.

References

Agriquem (2009, October). Análisis químico de suelos de la ciudad de Arica.

Angrist, J. and A. Krueger (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association 87*, 328–336.

ATDSR (2007, August). Toxicological profile for lead. *Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services*.

Banks, E., L. Ferretti, and D. Schucard (1997). Effects of low level lead exposure on cognitive function in children: a review of behavioral, neuropsychological and biological evidence. *Neurotoxicology 18*(1), 237–281.

Bellinger, D., K. Stiles, and H. Needleman (1992). Low-level lead exposure, intelligence and academic achievement: A long term follow-up study. *Pedriatrics 90*, 855–861.

Canfield, R., C. Henderson, D. Cory-Slechta, C. Cox, T. Jusko, and B. Lanphear (2003, April). Intellectual impairment in children with blood lead concentrations below $10 \leq \mu g$ per deciliter. *The New England Journal of Medicine 348*(16), 1517–1526.

CDC (1991). Preventing lead poisoning in young children. Report, Atlanta, GA: US Department of Health and Human Services.

Clay, K., W. Troesken, and M. R. Haines (2006, October). Lead Pipes and Child Mortality. *National Bureau of Economic Research, NBER, Working Paper 12603*.

Clay, K., W. Troesken, and M. R. Haines (2010). Lead, Mortality and Productivity. *National Bureau of Economic Research, NBER, Working Paper 16480*.

CONAMA (2009). Programa Maestro de Intervención en Zonas con presencia de Polimetales en Arica. Report, Comisión Nacional de Medio Ambiente.

CONGRESS (2009). Informe sobre acopios de residuos tóxicos y peligrosos internados ilegalmente por la empresa Promel en sectores poblacionales de la comuna de Arica. Report, Chilean Congress.

Cook, R. D. (1979, March). Influential observations in linear regression. *Journal of the American Statistical Association 74*(365), 169–174.

Coria, C., A. Cabello, E. Tassara, E. López, H. Rosales, M. Pérez, C. Zavala, P. Muñoz, G. Orellana, M. I. Inostroza, L. Contreras, and L. Kristen (2009). Long term consequences among children exposed to lead poisoning. *Revista Médica de Chile 137*, 1037–1044.

Currie, J. and R. Walker (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics 3*.

Delves, H. T., J. C. Sherlock, and M. Quinn (1984, November-December). Temporal stability of blood lead concentrations in adults exposed only to environmental lead. *Human and Experimental Toxicology 3*(4), 279–288.

GORE (2011). Cuenta Pública 2011. Report, Gobierno Regional de Arica y Parinacota.

Kaufman, A. (2001). Do low levels of lead produce IQ loss in children? A careful examination of the literature. *Archives of Clinical Neuropsychology 16*(4), 303–341.

Kaufman, A. and N. Kaufman (1997). K-TEA: Kaufman Test of Educational Achievement: Comprehensive form Manual. *American Guidance Service*.

Koller, K., T. Brown, A. Spurgeon, and L. Levy (2004, June). Recent developments in low-level lead exposure and intellectual impairment in children. *Environmental Health Perspectives 112*(9), 987–994.

Landrigan, P. J., C. B. Schecter, J. M. Lipton, M. C. Fahs, and J. Schwartz (2002, July). Environmental pollutants and disease in american children: estimates of morbidity, mortality and cost for lead poisoning, asthma, cancer and developmental disabilities. *Environmental Health Perspectives 110*(7), 721–728.

Lanphear, B., K. Dietrich, P. Auinger, and C. Cox (2000, November-December). Cognitive deficits associated with blood lead concentrations $10\mu g/dL$ in US children and adolescents. *Public Health Reports 115*(6), 521–529.

Lanphear, B., R. Hornung, M. Ho, C. Howard, S. Eberly, and K. Knauf (2002, January). Environmental lead exposure during early childhood. *The Journals of Pediatrics 140*(1), 40–47.

Leslie, G. R. and A. H. Richardson (1961, Diciembre). Life-cycle, career pattern and the decision to move. *American Sociological Review 26*(6), 894–902.

Lewis, J. (1985, May). Lead poisoning: A historical perspective. *EPA Journal*.

Lidsky, T. and J. S. Schneider (2003, January). Lead neurotoxicity in children: basic mechanisms and clinical correlates. *Brain 126*(1), 5–19.

Mielke, H. (2002, April). Research ethics in pediatric environmental health: Lessons from lead. *Neurotoxicology and Teratology 24*, 467–469.

Mielke, H. and P. Reagan (1998, February). Soils is an important pathway of human lead exposure. *Environmental Health Perspectives 106*(1), 217–229.

MINVU (1990). Decreto Supremo 140 que Reglamenta Programa de Viviendas Sociales. Technical report, División Jurídica, Ministerio de Vivienda y Urbanismo.

Muennig, P. and P. Bao (2009, December). The social cost of childhood lead exposure in New Jersey. *New Jersey Department of Public Advocate*.

Needleman, H., C. Gunnoe, A. Leviton, R. Reed, H. Peresie, C. Maher, and P. Barret (1979, March). Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *The New England Journal of Medicine 300*(13), 689–695.

PTF (2000, February). Eliminating childhood lead poisoning: A federal strategy targeting lead paint hazards. *President's Task Force on Environmental Health Risks and Safety Risks to Children*.

Ruiz-Tagle, J. and K. Sehnbruch (2009). Elaboración de un Indicador de la Calidad del Empleo. Report, United Nations Procurement Division (UNPD).

SERPAJ (1996). Diagnóstico medioambiental población Los Industriales y Cerro Chuño. Report, Servicio de Paz y Justicia.

Servicio de Salud de Arica (2001). Listado de personas con niveles de plomo en sangre superiores a $10\mu g/dL$ con contramuestras confirmadas por el I.S.P. efectuadas el año 2000 a las poblaciones

Cerro Chuño y Los Industriales. *Servicio de Salud de Arica, Subdirección Médica, Unidad de Epidemiología.*

Sinnott, R. W. (1984). Virtues of the Haversine. *Sky and Telescope 98*(2), 159.

Soto, R. and A. Torche (2004, Diciembre). Spatial inequality, migration and economic growth in Chile. *Cuadernos de economía 41*(7), 401–404.

Suprema, C. (2007). Fallo Corte Suprema Rol 3174-2005, Resolución 12947. Technical report, Corte Suprema.

Tchernitchin, A., L. Nina, L. Molina, G. Molina, N. Tchernitchin, C. Acevedo, and P. Alonso (2005). Human exposure to lead in chile. *Reviews of Environmental Contamination and Toxicology 185*(6), 93–139.

Wasserman, G., X. Liu, D. Popovac, P. Factor-Litvak, J. Kline, C. Waternaux, N. Lolacono, and J. Graziano (2000, November-December). The Yugoslavia prospective lead study: contributions of prenatal and postnatal lead exposure to early exposure to early intelligence. *Neurotoxicology and Teratology 22*(6), 811–818.

Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children Revised. *New York: Psychological Corporation.*

Wolpaw, J. (2007, May). Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime. *The B.E. Journal of Economic Analysis and Policy, Berkeley Electronic Press 7*(1).

Ziegler, E., B. Edwards, R. Jensen, K. Mahaffey, and S. Fomon (1978, January). Absorption and retention of lead by infants. *Pediatric Research 12*(1), 29–34.

Table 1: Construction, completion, and reception dates of Housing Projects in Arica

Housing Project	Construction date	Number of houses built	Completion date	Building authorization date	Reception date	Number of houses receptioned
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Los Industriales I	1985	536	March, 1989	N.A.	N.A.	N.A.
Los Industriales II	N.A.	316	29th April, 1991	21th November, 1994	N.A.	N.A.
Los Industriales III	1990	279	29th April, 1991	11th June, 1997	16th September, 1997	279
Los Industriales IV	1991	405	17th December, 1992	21th November, 1994	N.A.	N.A.
Cerro Chuño I	1993	149	N.A.	4th January, 2001	9th September, 2001	118
Cerro Chuño II	1994	60	N.A.	4th January, 2001	9th September, 2001	N.A.
Cerro Chuño III-a	1994	203	N.A.	4th January, 2001	24th September, 2001	105
Cerro Chuño III-b	1995	38	N.A.	4th January, 2001	11th September, 2001	19
Cerro Chuño IV	1994	185	N.A.	4th January, 2001	9th October, 2001	44
Cerro Chuño V	1995	110	N.A.	4th January, 2001	4th October, 2001	47
Cerro Chuño VII	1996	134	N.A.	4th January, 2001	9th October, 2001	36

Notes: The information provided in this table was obtained from documents of Chilean Congress and from a report elaborated by Peace and Justice Service (SERPAJ, 1996). Columns (2), (3), and (4) show the construction dates, number of houses built, and houses completion dates. Since the neighborhoods were built in areas destined for industrial activities, the Housing Ministry (*Ministerio de Vivienda y Urbanismo*) did not request the building authorization of the neighborhoods to the Municipality. In turn, the Municipality of Arica could not authorize the construction of housings projects due to the fact that areas were out of its jurisdiction. As soon as the areas were incorporated into the regulatory plane of the city, the Housing Ministry requested the building authorizations to the Municipality in the dates in column (5) and the property titles domain were delivered to the residents. The final reception dates and the number of houses receptioned by the Municipality are presented in columns (6) and (7), respectively.

Table 2: Summary Statistics. PSU 2008

Variable	Mean	Std. Dev.	Min.	Max.
Distance from *site F* to home	2.321	0.938	0.222	4.049
Home located to the east of *site F*	0.153	0.360	0	1
Age	18.591	0.535	18	21
Male	0.446	0.497	0	1
Mother absence	0.059	0.236	0	1
Father absence	0.275	0.446	0	1
Mother with less than 8 years of education	0.165	0.371	0	1
Mother with 8-11 years of education	0.150	0.357	0	1
Mother with 12 years of education	0.237	0.426	0	1
Mother with more than 12 years of education	0.386	0.487	0	1
Father with less than 8 years of education	0.121	0.327	0	1
Father with 8-11 years of education	0.108	0.311	0	1
Father with 12 years of education	0.182	0.386	0	1
Father with more than 12 years of education	0.312	0.463	0	1
Private school	0.0481	0.214	0	1
Voucher school	0.459	0.498	0	1
Public school	0.492	0.500	0	1
Math PSU 2008 score	502.805	92.341	203	800
Language PSU 2008 score	483.632	93.900	204	747
N. of Obs.		1,076		

Source: PSU 2008. Notes: Distances are expressed in kilometers. In order to consider the winds currents and the dispersion of pollutants around the city, we include the variable home located to the east of *site F*. It is equal to 1 if the student's home is located to the east of the *site F* (main contaminated area) and 0 in otherwise. Parent's absence is equal to 1 if the biological or stepfather/stepmother parent does not live with the student. Parental education was tabulated only if parents were present.

Table 3: Summary Statistics. PSU 2004 and Wages 2009-2010

Variable	Mean	Std. Dev.	Min.	Max.
Average monthly wage 2010-2011 (Chilean pesos)	395,447	282,095	10,044	2,136,248
Language PSU 2004 score	480.356	119.291	190	840
Math PSU 2004 score	482.956	107.917	112	840
Age	26.096	3.229	20	64
Male	0.501	0.500	0	1
Mother with less than 8 years of education	0.166	0.372	0	1
Mother with 8-11 years of education	0.287	0.452	0	1
Mother with 12 years of education	0.302	0.459	0	1
Mother with more than 12 years of education	0.243	0.429	0	1
Father with less than 8 years of education	0.222	0.416	0	1
Father with 8-11 years of education	0.231	0.421	0	1
Father with 12 years of education	0.271	0.444	0	1
Father with more than 12 years of education	0.273	0.445	0	1
Public school	0.466	0.498	0	1
Region	9.459	3.867	1	13
N. of Obs.		65,497		

Source: PSU 2004 and Unemployment Insurance Database. Notes: Parent's absence is equal to 1 if the biological or stepfather/stepmother parent does not live with the student. Parent's education was tabulated only if parents were present. Region corresponds to a fixed effect to consider the place where students took the test in 2004.

Table 4: The Effect of Distances to Sources of Pollution on Blood Lead Levels (BLLs)

Variable	(1)	(2)	(3)
(ln) distance from *site F* to home	-0.636**		
	(0.318)		
(ln) distance from railway to home		0.114	
		(0.337)	
(ln) distance from leaded soil to home			0.159
			(0.148)
Constant	9.973***	10.242***	10.501***
	(0.324)	(0.287)	(0.341)
Age and Male	Yes	Yes	Yes
Observations	377	377	378
R^2	0.0428	0.0369	0.0377

Notes: All distances are expressed in kilometers. BLLs are measured in micrograms per deciliter of blood (μg/dL). As we described in the text, we use the average of two independent blood tests taken to residents of *Cerro Chuño* and *Los Industriales* during 2000. One of these samples was taken by Municipal Health Service (SERMUS) and another, which is the confirmation of these findings, were performed by the Public Health Institute (ISP). We present the estimated coefficients of regressions of home distance to the nearest centroid of pollution of each of the three sources of lead contamination using the average of these two measures of blood lead levels. In this case we control for influential observations using Cook's distance (Cook, 1979). All the numbers in the table are robust to this procedure. Standard errors in parentheses. *** $p<0.01$,** $p<0.05$, * $p<0.1$.

Table 5: School-level average math and language SIMCE test scores: Schools in Arica, 1988-2010

Math

Distance from school to *Site F*	1988 (1)	1989 (2)	1990 (3)	1991 (4)	1992 (5)	1994 (6)	1996 (7)	1999 (8)	2002 (9)	2005 (10)	2006 (11)	2008 (12)	2010 (13)
Less than 1.5 kilometer	38.88	61.22	67.92	64.88	71.52	69.00	69.94	252.5	239.49	235.51	236.79	225.2	240.65
	(3.87)	(11.26)	(7.70)	(7.40)	(7.29)	(7.05)	(7.57)	(27.08)	(17.38)	(26.92)	(27.35)	(28.49)	(16.89)
Between 1.5 and 2.5 kilometers	36.41	61.96	64.5	58.2	73.25	73.67	70.45	244.22	249.33	246.17	256.91	249.54	265.71
	(12.32)	(7.69)	(7.51)	(9.15)	(5.29)	(6.92)	(6.59)	(16.48)	(13.42)	(26.74)	(27.19)	(37.21)	(34.11)
Between 2.5 and 3.5 kilometers	41.38	59.24	69.17	57.72	73.6	75.10	75.64	255.42	248.05	257.96	242.81	244.87	255.12
	(11.16)	(10.63)	(9.01)	(10.98)	(8.11)	(7.29)	(4.57)	(10.60)	(14.66)	(20.06)	(19.59)	(23.53)	(27.05)
Between 3.5 and 4.5 kilometers	33	58.74	65.24	55.66	71.07	74.10	73.61	255.92	251.12	254.5	251.52	250.93	254.28
	(13.94)	(12.54)	(8.97)	(11.10)	(6.48)	(7.79)	(6.28)	(22.37)	(20.85)	(15.75)	(15.25)	(21.52)	(24.27)
More than 4.5 kilometers	32.03	59.57	64.74	48.7	72.98	71.90	75.16	247.25	252.74	247.37	259.37	242.06	259.91
	(11.70)	(14.29)	(12.00)	(16.91)	(9.69)	(7.59)	(6.26)	(34.84)	(22.73)	(37.11)	(37.30)	(34.65)	(29.91)

Language

Distance from school to *Site F*	1988 (1)	1989 (2)	1990 (3)	1991 (4)	1992 (5)	1994 (6)	1996 (7)	1999 (8)	2002 (9)	2005 (10)	2006 (11)	2008 (12)	2010 (13)
Less than 1.5 kilometer	43.42	60.43	67.73	61.73	71.36	65.65	70.17	249.17	239.49	242.02	239.97	244.54	259.8
	(4.56)	(6.82)	(6.77)	(3.07)	(5.81)	(4.89)	(6.94)	(25.96)	(21.36)	(25.34)	(26.38)	(24.94)	(17.58)
Between 1.5 and 2.5 kilometers	42.53	60.93	66.48	55.77	73.06	69.30	70.98	243.67	248.89	249.41	254.89	264.66	278.44
	(11.75)	(6.98)	(6.83)	(4.14)	(6.12)	(6.00)	(6.87)	(18.68)	(13.56)	(23.16)	(22.21)	(27.21)	(27.26)
Between 2.5 and 3.5 kilometers	46.99	64.08	69.43	60.6	72.61	70.99	75.05	252.25	252.21	256.04	244.23	260.26	265.68
	(10.68)	(8.67)	(6.28)	(8.72)	(7.35)	(6.08)	(4.22)	(11.99)	(16.77)	(20.37)	(18.33)	(19.61)	(14.79)
Between 3.5 and 4.5 kilometers	42.49	60.46	66.42	56.93	71.77	70.97	74.91	257.58	252.08	261.42	253.98	261.42	270.57
	(18.47)	(8.59)	(9.78)	(6.75)	(7.09)	(6.34)	(6.11)	(27.68)	(23.01)	(15.22)	(13.98)	(19.72)	(19.28)
More than 4.5 kilometers	39.53	61.2	65.57	53.98	71.95	66.34	76.27	245.25	251.94	233.54	259.31	270.56	275.29
	(16.63)	(13.65)	(10.63)	(15.61)	(11.36)	(7.63)	(5.87)	(39.90)	(29.61)	(31.96)	(32.55)	(26.19)	(25.66)

Source: SIMCE 1988-2010. Notes: We present the weighted average SIMCE scores for math and language. Distances are expressed in kilometers. From 1999, SIMCE changed its measurement scale. This explain the difference of scores between 1988-1992 and 1999-2010. Standard deviations in parentheses.

Table 6: Estimates from Reduced Form Model: The effect of distance on math and language PSU scores

Variable	Math (1)	Language (2)	Math (3)	Language (4)	Math (5)	Language (6)
(ln) distance from *site F* to home	13.901**	17.430***	10.005*	13.234***	8.783**	12.325***
	(6.204)	(4.922)	(5.097)	(3.801)	(3.951)	(4.074)
Home located to the east of *site F*	-7.028	-5.306	-3.731	0.848	-4.665	0.153
	(13.108)	(12.077)	(12.304)	(10.627)	(10.764)	(9.983)
Constant	675.479***	614.243***	642.450***	572.473	635.424***	567.244***
	(89.074)	(72.226)	(87.259)	(76.082)	(85.916)	(80.579)
Age and Gender	Yes	Yes	Yes	Yes	Yes	Yes
Mother and Father absence	No	No	Yes	Yes	Yes	Yes
Parental education	No	No	Yes	Yes	Yes	Yes
Type of school	No	No	No	No	Yes	Yes
R^2	0.0396	0.0269	0.0939	0.0789	0.0939	0.0871
Observations			1,076			

Notes: Distances are expressed in kilometers. In order to consider the winds currents and the dispersion of pollutants around the city, we include the variable home located to the east of *site F*. It is equal to 1 if the student's home is located to the east of *site F* (main contaminated area), 0 otherwise. Parent's absence is equal to 1 if the biological or stepfather/stepmother parent does not live with the student. Parent's education was tabulated only if parents were present. Clustered standard errors (at the school level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: The effect of lead on math and language PSU test scores

Stage	Math (1)	Language (2)
First Stage	-0.636**	-0.636**
	(0.318)	(0.318)
Reduced From	8.783**	12.325***
	(3.951)	(4.074)
TSIV	-13.810	-19.379*
	(9.288)	(11.615)
Observations	1,076	1,076

Notes: The causal effect of lead on PSU scores is expressed in PSU points. Standard errors were calculated using the Delta Method. Assuming that $cov(\hat{\gamma}_1, \hat{\alpha}_1) = 0$, the variance of $\hat{\beta}_{TSIV}$ is giving by the following expression: $Var(\hat{\beta}_{TSIV}) = Var(\hat{\gamma}_1)/\hat{\alpha}_1^2 + (\hat{\gamma}_1^2/\hat{\alpha}_1^4) \cdot Var(\hat{\alpha}_1)$. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Proportion of parents with college education with children enrolled in primary schools in Arica, by distance from school to *site F*

Distance from school to site F	1989	1999	2008
Less than 1.5 kilometers	10.2%	9.7%	11.4%
Between 1.5 and 3 kilometers	1.7%	2.1%	12.0%
More than 3 kilometers	23.9%	30.3%	31.3%

Source: SIMCE 1989-2008. Notes: We present the proportion of parents with college for three different categories of distance from home to *site F*.

Table 9: PSU 2004 Scores on Average Monthly Wages 2010-2011 (National Level)

Variable	(1)
Math PSU 2004 score	0.0017***
	(0.000)
Language PSU 2004 score	0.0001***
	(0.000)
Constant	11.7713***
	(0.106)
Region fixed effect	Yes
Family background	Yes
Age and gender	Yes
School-level variables	Yes
N. of Obs.	65,497

Source: PSU 2004 and Unemployment Insurance Database. *Region* corresponds to the region in which the student took the test in 2004. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 10: Marginal effect of PSU points on monthly wages

		Math	Language
$\frac{\partial W}{\partial PSU}$	(Pesos/PSU points)	745	60

Notes: Using the estimated coefficients of the relationship between wages and PSU scores we can obtain the marginal effect of an increase in language and math PSU scores on the monthly wages of the students who took PSU in 2004 and for whom we observe their wages in 2010 and 2011. Due the fact that the dependent variable in the regression estimated in Table 9 is expressed as the natural logarithm of the monthly wages, under the assumption of normality, we can write the marginal effect of an increase in PSU score as $\hat{\beta}_1 \times \exp(\hat{\beta}_0 + \hat{\beta}_1 P\bar{S}U + \hat{\beta}_2 \bar{X} + \hat{\sigma}^2/2)$. As we discussed in the text, PSU corresponds to the scores obtained in math or language tests, X corresponds to a vector of observed characteristics (such as gender, age at the moment of which we observe the wages, parental education and the region in which they took the test). Then, replacing the coefficients presented in Table 9, we find that on average, for each extra point in math PSU scores, the monthly wage increases in 745 CLP (or 1.49 dollars), while in the case language PSU this effect corresponds to 60 CLP (or 0.12 dollars). We use 1 dollar=500 pesos.

Table 11: Marginal effect of lead on monthly wages

			Math	Language	Total Effect
(1)	$\frac{\partial W}{\partial Pb}$	$\frac{Pesos}{\mu g/dL}$	-10,296	-1,162	-11,458
(2)	$\%\frac{\partial W_{mean}}{\partial Pb}$	$\%\left(\frac{Pesos}{\mu g/dL}\right)$	-2.6%	-0.3%	-2.9%

Notes: In order to analyze the effect of lead exposure on wages, we develop an expression that relates: (1) the marginal effect of blood lead concentration on PSU scores and (2) the marginal effect of PSU scores on wages (i.e, $\frac{\partial W}{\partial Pb} = \frac{\partial W}{\partial PSU_{lang}} \times \frac{\partial PSU_{lang}}{\partial Pb} + \frac{\partial W}{\partial PSU_{math}} \times \frac{\partial PSU_{math}}{\partial Pb}$). Using data from blood tests taken to residents of the neighborhoods nearest to *site F* (main source of exposure to lead in Arica) during 2000 and PSU scores in the subjects of math and language, we obtain the effect in terms of PSU points of an increase of 1 microgram of lead per deciliter of blood (i.e, $\frac{\partial PSU}{\partial Pb}$). Table 7 presents these results differentiated by subject (math and language). Then, using a sample of wages and PSU 2004 scores, we explore the relationship between scores PSU and wages (i.e, $\frac{\partial W}{\partial PSU}$). In particular, we estimate a Mincer equation taking as dependent variable the natural logarithm of the average monthly wages obtained between 2010 and 2011 by the students who took PSU in 2004. The estimated coefficients of this relationship are presented in Table 9 while the marginal effect of the increase in 1 PSU point on monthly wages is presented in Table 10. Then, combining both estimated effects ($\frac{\partial W}{\partial PSU}$ and $\frac{\partial PSU}{\partial Pb}$), we obtain the effect of an increase of 1 microgram of lead per deciliter of bood ($1\mu g/dL$) on monthly wages. The first row of the table shows these results. Then, for each increase of 1 microgram on individual's blood lead level, on his (her) monthly wage will decrease in CLP 11,458 pesos (or equivalently US 22.92 dollars). Finally, the row (2) shows the marginal effect of lead on annual wages as percentage of the average monthly wage of the sample used (CLP 395,447 pesos per month).

Table 12: Marginal effects of distance from home to *Site F* on monthly wages

Distance	Math		Language		Math+Lang		Percentage
	(CLP)	(US$)	(CLP)	(US$)	(CLP)	(US$)	%
0.1 kms	65,480	130.96	7,393	14.79	72,873	145.75	18.4
0.5 kms	13,096	26.19	1,479	2.96	14,575	29.15	3.7
1.5 kms	4,365	8.73	493	0.99	4,858	9.72	1.2
2.5 kms	2,619	5.24	296	0.59	2,915	5.83	0.9

Notes: In order to analyze the effect of lead exposure (measured through the distance from home to the main source of pollution in Arica) on wages, we develop an expression that relates: (1) the marginal effect of distance from home to *site F* on PSU scores and (2) the marginal effect of PSU scores on wages (i.e, $\frac{\partial W}{\partial D_{home}} = \frac{\frac{\partial W}{\partial PSU_{lang}} \times \frac{\partial PSU_{lang}}{\partial \ln(D_{home})}}{D} + \frac{\frac{\partial W}{\partial PSU_{math}} \times \frac{\partial PSU_{math}}{\partial \ln(D_{home})}}{D}$ Using the estimated coefficients of equation 2 , we obtain the effect in terms of PSU points of an increase of 1 kilometer of the distance from student's home to *site F*. Then, using a sample of wages and PSU 2004 scores, we explore the relationship between scores PSU and earnings (i.e, $\frac{\partial W}{\partial PSU}$). In particular, we estimate a Mincer equation taking as dependent variable the natural logarithm of the average monthly earnings obtained between 2010 and 2011 by the students who took PSU in 2004. The estimated coefficients of this relationship are presented in Table 9 while the marginal effect of the increase in 1 PSU point on monthly wages is presented in Table 10. Then, combining both estimated effects ($\frac{\partial W}{\partial PSU}$ and $\frac{\partial PSU}{\partial D_{home}}$), we obtain the effect of an increase of 1 kilometer in the distance from student's home to *site F* on monthly wages. The first row of the table shows these results for Distance equal to 0.1km (people living very close to *site F*). Then, for each increase of 1 kilometer in the distance to *site F*, on his (her) monthly wage will increase in CLP 72,873 pesos (or equivalently US 145.75 dollars).

Table 13: Estimated coefficients for Equation (8)

	Math	Language
Natural logarithm of the distance from site F to home	8.881*	12.436***
	(5.157)	(3.975)
Family Background and characteristics ($P\hat{S}U_{-D}$)	0.524***	0.555***
	(0.133)	(0.145)
Home located to the east of *site F*	-3.747	-1.681
	(12.271)	(10.907)
Constant	237.515*	203.567**
	(75.890)	(76.208)
Observations	1,076	1,076

Notes: Robustness check. Ordinary Least Squares estimates of the effect of distance to *site F* on 2008 PSU scores. Instead of family background and characteristics, we include the predicted value of the PSU using the parameters of three northern provinces not affected by lead (Parinacota, Iquique, and Tamarugal), hence these parameters are consistent and not affected by correlation between unobservables and distance from *site F*.

Table 14: Constrained estimated coefficients for Equation (9)

Variable	Math			Language		
	γ_2=0.8	γ_2=1	γ_2=1.2	γ_2=0.8	γ_2=1	γ_2=1.2
	(1)	(2)	(3)	(4)	(5)	(6)
Natural logarithm of the distance from *site* F to home	6.344	4.507	2,670	10.283***	8.524***	6.765***
	(5.874)	(5.975)	(6.149)	(3.422)	(3.218)	(3.129)
Family Background and characteristics ($P\hat{S}U_{-D}$)	0.8	1	1.2	0.8	1	1.2
	(.)	(.)	(.)	(.)	(.)	(.)
Home located to the east of *site F*	-0.907	1.149	3.2050	0.546	2.367	4.187
	(11.407)	(10.914)	(10.460)	(10.532)	(10.062)	(9.616)
Constant	102.66***	4.005	-94.057*	85.251***	-11.473	-108.198***
	(13.384)	(13.158)	(12.956)	(9.521)	(9.231)	(8.964)
Observations	1.076	1.076	1.076	1.076	1.076	1.076

Notes: Constrained Least Squares for different values of γ_2. Thus, by fixing γ_2 and estimating by CLS, we are avoiding any residual correlation between the family background and unobservables determining the PSU score.
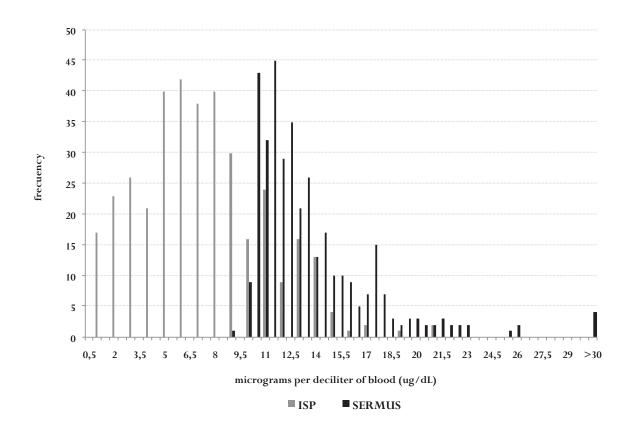
Figure 1: *Site F* before the construction of housing projects

Figure 2: *Site F* after the construction of housing projects

Figure 3: Histogram of blood lead levels: ISP and SERMUS samples

Note: This figure presents histograms of blood lead level samples collected from the residents of *Cerro Chuño* and *Los Industriales* during 2000. For each individual, we have two measures: SERMUS and ISP.
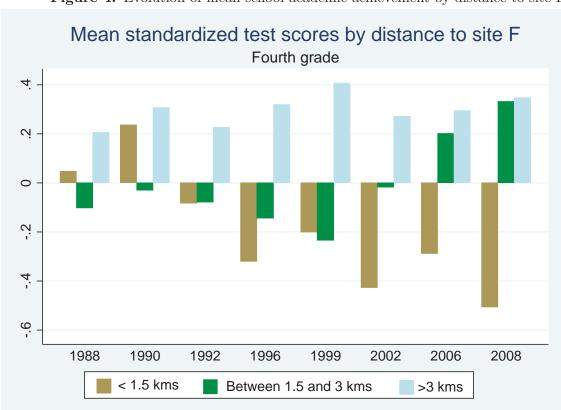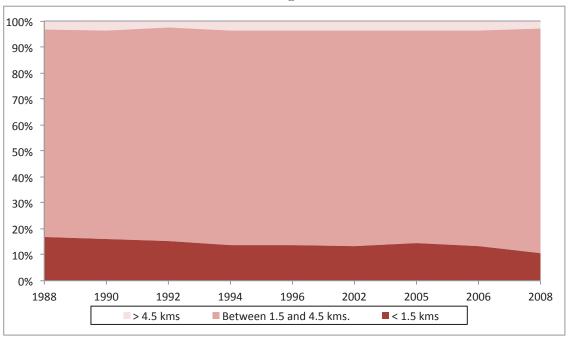
Figure 4: Evolution of mean school academic achievement by distance to site F

Note: This figure shows mean standardized test scores of schools located at 1.5 or less kilometers from site F, schools between 1.5 and 3 kilometers from site F and schools 3 or kilometers from site F. The results consider 4th grades test score results.

Figure 5: The trends in the distribution of enrollment by school distance to *site F*: Fourth graders

Source: SIMCE.

Appendix

Table A.1: Effects of lead on IQ

	IQ Measurement Test	Sample IQ Scores	Effects of lead on IQ
Needleman et al. (1979)	WISC-R	WISCR Full Scale: M: 106.6 pts. for LL and 102.1 pts. for HL children. WISCR Verbal: M: 103.9 pts. for LL and 99.3 pts. for HL children. WISCR Performance: M: 108.7 pts. for LL and 104.9 pts. for HL children.	Authors report Pvalue of the difference between IQ scores for high and low lead children. Full Scale Pvalue: 0.03, Verbal IQ: 0.03, Performance: 0.08.
Canfield et al. (2003)	SBIS- 4th Ed. (Std. M: 100 pts., Std. Dev.: 16 pts.)	M: 89.9 pts, Std. Dev.: 11.4 pts.	For each increase of 1 μg/dL: -0.46 pts. and 0.74 pts. with a linear and non linear model, resp.
Bellinger et al. (1992)	WISC-R - 3rd Version (M: 100 pts., Std. Dev.: 15 pts.) KTEA (M: 100 pts., Std. Dev.: 15 pts.)	WISCR Full Scale: M: 119.1 pts., Std. Dev.: 14.8 pts. KTEA Battery Composite: M: 118.18 pts., Std. Dev: 16.3 pts.	For each increase of 1 μg/dL: -0.58 pts. in WISCR Full Scale and -0.89 pts. in KTEA Battery Composite.
Lanphear et al. (2002)	WRAT: Arithmetic and Reading (Std.M: 100 pts., Std. Dev.: 15 pts.) and WISC-R: Block Design and Digit Span subsets (Std. M: 10 pts., Std.Dev.: 3 pts.)	WRAT Arithmetic: M: 93.5 pts., Std. Dev.: 0.5 pts. WRAT Reading: M: 91.9 pts., Std. Dev: 0.5 pts. WISC-R: Block Design: M: 9.5 pts., Std.Dev.: 0.1 pts.	For each increase of 1 μg/dL: -0.7 points in Arithmetic, -0.99 pts. in Reading, -0.1 points in Block Design and -0.05 pts. in Digit Span scores.
Wasserman et al. (2000)	WISC-R - 3rd Version (M: 100 pts., Std.Dev.: 15 pts.)	WISC-R: Digit Span: M 8.7 pts., Std. Dev.: 0.1 pts.	After an increase form 10 μg/dL to 30 μg/dL authors document: -4.3 pts. in WISC III Full Scale, -3.4 pts. in Verbal WISC III and -4.5 pts. WISC III.

Notes: (a) IQ Test correspond to: WISC-R: Wechsler Intelligence Scale for Children - Revised, SBIS: Stanford-Binet Intelligence Scale, KTEA: Kaufman Test of Educational Achievement and WRAT: Wide Range Achievement Test - Revised. (b) Needleman et al. (1979) define two groups of study depending on the concentration of lead on children's dentine: low levels (LL) <100 ppm and high levels (HL)>20 ppm.

Table A.2: Summary Statistics. Blood lead levels (BLLs) of the residents of *Cerro Chuño* and *Los Industriales* and distances between their homes and different sources of pollution

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Age	376	9.305	10.599	1	78
Male	376	0.638	0.481	0	1
BLL: IPS Measure	376	7.281	3.836	1	21
BLL: SERMUS Measure	376	13.593	3.966	8.6	42.8
BLL: Average ISP and SERMUS Measures	376	10.437	3.047	5.55	24.4
Distance from *site F* to home	376	0.668	0.316	0.166	1.300
Distance from railway to home	376	1.066	0.419	0.243	1.776
Distance from leaded soil to home	376	0.505	0.382	0.009	1.225

Notes: BLLs are expresed in micrograms per deciliter of blood (μg/dL). All distances are expressed in kilometers. Male is equal to 1 for male, 0 for female. Blood lead levels correspond to 390 blood tests taken to *Cerro Chuño* and *Los Industriales* residents during 2000. As described in the text, each sample reports two measures: one taken by Municipal Health Service (SERMUS) and another, which is the confirmation of these findings made by the Public Health Institute (ISP). From this data we consider distances beetween his/her home and three sources of pollution: the *site F*, Line Railroad Arica-La Paz and a set of 26 points whose lead levels exceeded the EPA's international standard for soil (400 miligrams of lead for kilo of soil in analysis). Specifically, we geo-reference residential addresses and calculate the minimum distance between home and 33 centroids of pollution covering the full extent of the railway line and the 26 points of contamination in soil. In the case of *site F*, we consider *Chapiquiña 3066* as reference direction.

Table A.3: Distance from home to *Site F* and parent's education

Variable	
Mother absence	0.101
	(0.141)
Father absence	0.087
	(0.103)
Mother with 8-11 years of education	0.113
	(0.103)
Mother with 12 years of education	0.118
	(0.095)
Mother with more than 12 years of education	0.241***
	(0.091)
Father with 8-11 years of education	-0.093
	(0.120)
Father with 12 years of education	-0.055
	(0.110)
Father with more than 12 years of education	0.202*
	(0.105)
Male	0.001
	(0.057)
Constant	2.110***
	(0.095)
Observations	1,076
R^2	0.0284

Source: PSU 2008. Notes: Distances are expressed in kilometers. Parent's absence is equal to 1 if the biological or stepfather/stepmother parent does not live with the student. Parental education was tabulated only if parents were present. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.4: Summary Statistics. SIMCE 2003-2006

Variable	2003				2004				2005				2006			
	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.
Distance from school to *site F*	2.845	1.826	0.774	12.899	2.957	1.555	0.774	12.899	2.962	1.445	0.774	12.899	2.898	1.815	0.774	12.899
Natural logarithm of the distance from school to *site F*	0.894	0.548	-0.255	2.557	0.975	0.470	-0.255	2.557	0.982	0.468	-0.255	2.557	0.905	0.571	-0.255	2.557
School located to the east of *site F*	0.084	0.278	0	1	0.160	0.367	0	1	0.167	0.373	0	1	0.109	0.312	0	1
Male	0.501	0.500	0	1	0.526	0.499	0	1	0.502	0.500	0	1	0.492	0.500	0	1
Mother absence	0.088	0.284	0	1	0.083	0.277	0	1	0.001	0.042	0	1	0.006	0.077	0	1
Father absence	0.248	0.431	0	1	0.271	0.445	0	1	0.031	0.173	0	1	0.041	0.199	0	1
Mother with less than 8 years of education	0.138	0.345	0	1	0.138	0.345	0	1	0.095	0.293	0	1	0.122	0.328	0	1
Mother with 8-11 years of education	0.315	0.465	0	1	0.325	0.468	0	1	0.267	0.442	0	1	0.287	0.452	0	1
Mother with 12 years of education	0.288	0.453	0	1	0.287	0.452	0	1	0.362	0.480	0	1	0.351	0.477	0	1
Mother with more than 12 years of education	0.168	0.374	0	1	0.164	0.370	0	1	0.272	0.445	0	1	0.232	0.422	0	1
Father with less than 8 years of education	0.096	0.295	0	1	0.082	0.275	0	1	0.076	0.266	0	1	0.097	0.297	0	1
Father with 8-11 years of education	0.230	0.421	0	1	0.247	0.431	0	1	0.255	0.436	0	1	0.273	0.445	0	1
Father with 12 years of education	0.269	0.443	0	1	0.249	0.432	0	1	0.357	0.479	0	1	0.342	0.474	0	1
Father with more than 12 years of education	0.155	0.362	0	1	0.149	0.356	0	1	0.279	0.448	0	1	0.244	0.429	0	1
Public school	0.636	0.481	0	1	0.586	0.492	0	1	0.430	0.495	0	1	0.538	0.498	0	1
Voucher school	0.311	0.463	0	1	0.386	0.487	0	1	0.535	0.498	0	1	0.430	0.495	0	1
Private school	0.051	0.220	0	1	0.027	0.162	0	1	0.034	0.181	0	1	0.030	0.172	0	1
Language SIMCE score	250.86	49.255	129.62	397.16	245.63	49.340	101.68	380.8	254.191	51.464	108.63	364.77	250.19	50.781	130.27	394.14
Math SIMCE score	236.82	56.889	104.52	428.92	252.32	46.230	124.71	391.04	251.882	53.451	91.23	363.62	245.62	63.055	93.48	426.58
Observations	2,459	2,459	2,459	2,459	1,946	1,946	1,946	1,946	2,786	2,786	2,786	2,786	2,470	2,470	2,470	2,470

Source: SIMCE 2003-2006. Notes: Distances are expressed in kilometers. In order to consider the winds currents and the dispersion of pollutants around the city, we include the variable school located to the east of *site F*. It is equal to 1 if the school is located to the east of the *site F* (main contaminated area) and 0 in otherwise. Male is equal to 1 for male, 0 for female. Parent's absence is equal to 1 if the biological or stepfather/stepmother parent does not live with the student. Parent education was tabulated only if parents were present.

Table A.5: Distance from school to *site F* on SIMCE 2003-2006 scores

Variable	Math				Language			
	2003 (1)	2004 (2)	2005 (3)	2006 (4)	2003 (5)	2004 (6)	2005 (7)	2006 (8)
Natural logarithm of the distance from school to *site F*	9.016	2.083	13.998***	6.070	12.227*	3.461	8.530*	7.005
	(8.434)	(5.727)	(4.780)	(11.126)	(5.932)	(4.872)	(4.962)	(7.456)
School located to the east of *site F*	-10.548	-3.215	-14.927***	-2.401	-9.481	-3.726	-14.592**	-3.049
	(12.827)	(8.008)	(4.889)	(16.883)	(8.957)	(6.665)	(5.533)	(10.016)
Male	7.948**	11.406***	7.799***	8.746**	-1.645	-3.886	-4.347*	-5.491*
	(3.259)	(2.447)	(2.837)	(3.877)	(2.146)	(2.498)	(2.383)	(2.793)
Mother absence	2075	2.962	-14.393	13.280	0.664	2.731	-17.796	21.243*
	(3.456)	(4.939)	(24.067)	(12.615)	(4.840)	(5.251)	(13.677)	(11.487)
Father absence	10.324**	0.835	-11.383*	8.255	5.645	5.674	-4.771	5.144
	(4.917)	(4.265)	(6.219)	(7.451)	(3.358)	(4.761)	(5.454)	(4.682)
Mother with 8-11 years of education	3973	3134	3.346	3.731	3.817	1.619	3.529	5.959
	(2.645)	(3.394)	(4.237)	(3.858)	(3.545)	(3.437)	(3.998)	(3.706)
Mother with 12 years of education	15.085**	10.991***	8.183*	14.429***	13.186*	13.893***	10.102**	13.855***
	(5.393)	(3.631)	(4.111)	(5.091)	(6.344)	(3.966)	(4.062)	(4.720)
Mother with more than 12 years of education	19.207**	21.503***	11.844**	23.642***	17.015*	27.212***	14.948***	22.563***
	(8.927)	(4.439)	(5.295)	(6.567)	(9.482)	(5.067)	(4.654)	(5.288)
Father with 8-11 years of education	7130	1456	-6.821	6.984	2.224	2.662	-7.074	-0.202
	(4.162)	(4.351)	(4.667)	(5.567)	(4.612)	(4.499)	(4.667)	(2.292)
Father with 12 years of education	10.555*	4972	-2.430	16.372**	4.841	10.405**	-1.715	10.024***
	(5.922)	(4.680)	(3.665)	(6.050)	(4.636)	(4.684)	(4.078)	(2.687)
Father with more than 12 years of education	26.122***	16.592***	1.252	31.988***	15.987**	18.469***	3.897	23.530***
	(9.151)	(5.044)	(3.685)	(10.225)	(7.324)	(5.521)	(3.771)	(5.829)
Voucher school	15.980	11.657**	14.808***	11.804	11.157	10.291**	14.619***	6.069
	(14.132)	(4.511)	(6.450)	(18.355)	(10.641)	(4.698)	(5.403)	(12.223)
Private school	57.119***	35.123***	46.270***	52.492**	43.647***	24.115***	38.227***	30.660**
	(19.477)	(6.172)	(7.468)	(23.745)	(14.839)	(7.854)	(9.194)	(11.529)
Constant	197.612***	227.083***	222.753***	201.995***	222.198***	223.507***	234.134***	222.040***
	(9.727)	(7.850)	(8.303)	(17.440)	(7.949)	(6.866)	(7.974)	(10.824)
Observations	2,459	1,946	2,786	2,470	2,459	1,946	2,786	2,470
R²	0.1481	0.1170	0.0829	0.1346	0.1236	0.1107	0.0819	0.1295

Source: SIMCE 2003-2006. Notes: Distances are expressed in kilometers. In order to consider the winds currents and the dispersion of pollutants around the city, we include the variable school located to the east of *site F*. It is equal to 1 if the school is located to the east of the *site F* (main contaminated area) and 0 in otherwise. Male is equal to 1 for male, 0 for female. Parent's absence is equal to 1 if the biological or stepfather /stepmother parent does not live with the student. Parent education was tabulated only if parents were present. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.6: Individuals who took PSU 2008 in Arica but SIMCE 2004 in a different region

	Math	Language
	(1)	(2)
(ln) Distance from *site F* to home	12.127	5.188
	(19.158)	(21.028)
Age and Gender	Yes	Yes
Father/Mother absence	Yes	Yes
Parental Education	Yes	Yes
School type	Yes	Yes
N. of Obs.	78	77

Notes: Distances are expressed in kilometers. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: Individuals who took SIMCE 2004 in Arica but PSU 2009 in a different region

	Math	Language
	(1)	(2)
(ln) Distance from *site F* to school	-19.063	0.419
	(31.217)	(25.160)
Age and Gender	Yes	Yes
Father/Mother absence	Yes	Yes
Parental Education	Yes	Yes
School type	Yes	Yes
N. of Obs.	109	103

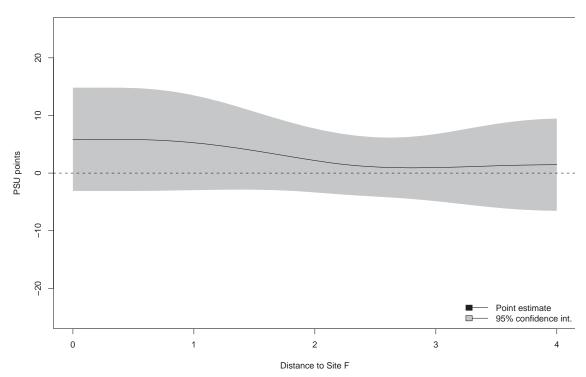Notes: Distances are expressed in kilometers. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure A.1: Semi-parametric derivative: Math PSU score

Note: Semi-parametric model using penalized splines:

$$\text{PSU} = \beta_0 + \beta_1 X + \beta_2 Z + g(D_{home}) + \varepsilon \tag{10}$$
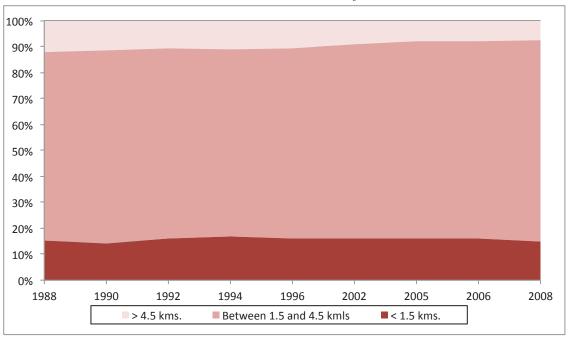
We plot the derivative of $g(\cdot)$.

Figure A.2: Semi-parametric derivative: Language PSU score

Note: Semi-parametric model using penalized splines:

$$\text{PSU} = \beta_0 + \beta_1 X + \beta_2 Z + g(D_{home}) + \varepsilon \tag{11}$$

We plot the derivative of $g(\cdot)$.

Figure A.3: The trends in the distribution of elementary schools by their distance to *site F* - City of Arica