LEAD EXPOSURE AND ITS EFFECTS ON ACADEMIC ACHIEVEMENT:

EVIDENCE FROM AN ENVIRONMENTAL NEGLIGENCE

Tomás Rau
Universidad Catolica

Loreto Reyes
Ministry of Finance
Chile

Sergio Urzúa
Ministry of Finance
Chile

Network on Inequality and Poverty
Cuernavaca, Mexico - 2011

Introduction

- Between 1984 and 1989, more than 20,000 tons of toxic chemicals were brought to the city of Arica (northern Chile).
- The chemical waste was located several kilometers from the city (*site F*).
- The rapid expansion of Arica in the early 1990s, which included the construction of public housing projects meters away from the waste deposit, put at risk a large number of families.
- For more than a decade, individuals living in the vicinity of contaminated areas were exposed to critical levels of lead.

Introduction

- We use these unfortunate events to evaluate the effect of lead exposure on academic achievement.
- We exploit the relationship between household and school distance to contaminated areas, lead blood concentrations and their consequences on cognitive development.

Figure 1: *Site F* before the construction of housing projects

Motivation

Table 1: Construction Dates of Housing Projects in Arica

Housing Project (1)	**Construction date** (2)	**Number of houses** (3)
Los Industriales I	1985	536
Los Industriales II	1994	316
Los Industriales III	1990	279
Los Industriales IV	1991	405
Cerro Chuño I	1993	149
Cerro Chuño II	1994	60
Cerro Chuño III-a	1994	203
Cerro Chuño III-b	1995	38
Cerro Chuño IV	1994	185
Cerro Chuño V	1995	110
Cerro Chuño VII	1996	134
	Total	**2,415**

Source: CONGRESS, 2003.

Figure 2: *Site F* and neighborhoods exposed to lead pollution

Figure 3: Distances from *Site F*

Toxicological effects of lead

- Depending on the dose and length of exposure, lead is associated with neurological damage and can even cause death in very high concentrations (ATDSR, 2007).
- 10 micrograms per deciliter of blood ($10\mu g/dL$) is the currently accepted threshold of risk. At this value, lead has been documented to have negative effects on IQ, hearing capability, behavior (agression) and bone growth (CDC, 1991),
- Nevertheless, there is evidence about the negative effects of lead at lower levels (even less to $10\mu g/dL$) (Nilsson et al., 2009).
- This suggests that there may not be a safe level of lead exposure.

Toxicological effects of lead

Children are especially vulnerable to the effects of lead for (at least) three reasons:

- tendency to put objects in their mouths, especially during teething
- their digestive systems have a higher degree of absorption of lead
- brain function children is still developing

Toxicological effects of lead

A large literature documents the detrimental effects of lead on cognitive development and academic achievement:

- Needleman H. et al. (1979)
- Bellinger D. et al. (1992)
- Wasserman G. et al. (2000)
- Lanphear B. et al. (2000)
- Koller K. et al. (2004)
- Chen A. et al. (2007)

Identification Strategy

- We analyze the effect of proximity to a source of lead on academic performance of students in Arica using production functions of academic achievement.
- We exploit the relationship between household and school distance to contaminated areas to estimate the impact of lead exposure on cognitive development measured through standardized test scores.
- We suppose that the students who took the tests were exposed to lead for a long time. Then, we capture the effect of lead using the distances to the sources of pollution.

Threats to Identification Strategy

- Distance to a source of lead pollution is a proxy for lead exposure.
- We do not know if the students always lived at the same address that they reported when took test or whether they have moved.
- We cannot accurately identify if changes between schools are explained by normal transitions from primary to secondary education or due to parents' decisions that could be influenced by the proximity of the schools to polluted areas.

Addressing problems with the Identification Strategy

- Using blood tests taken from residents of *Cerro Chuño* and *Los Industriales* during 2000, we obtain a significant correlation between distance to source of lead pollution and blood lead levels.
- It was prohibited to sell the houses built in *Cerro Chuño* and *Los Industriales* (Art. 43° Decreto Supremo N°140. *Reglamenta Programas de Viviendas Progresivas*).
- Scarce evidence of regional migration in Chile. Soto and Torche (2004): 0.6% per year.

Measures of Academic Achievement

In order to measure the academic performance of the students in Arica, we used data from two national standardized tests:

- SIMCE (System for Measuring the Quality of Education): 4th, 8th and 10th grades (2003-2006). This database contains information about family characteristics and background of the establishment.

- PSU (University Admission Test): 12th grade schools. In addition to test scores this database contains information on student's socio-economic background as well as the name and address of his/her school.

  Also, for those students who took PSU we observe their scores in math and language as well as a rich set of covariets including student's address (at the time of the test), income and GPA.

  In relation to SIMCE, PSU has the advantage of linking student's academic results (SIMCE 2004-2006 and PSU 2008).

Academic Achievement

- Initially, we use linear regression to estimate production functions of academic achievement of the effect of distance to *Site F* on SIMCE 2003-2006 scores.
- Then we restrict our sample of test scores to student who took SIMCE 2004 and PSU 2008.
- The estimation method is OLS and the inference is performed using school-cluster standard errors

SIMCE 2003-2006

- We estimate the effect of distance between school and *Site F* on SIMCE 2003-2006 scores, through the following expression:

$$\text{SIMCE}_{i,j} = \alpha_0 + \alpha_1 X_i + \alpha_2 D_{school(j),i} + \alpha_3 D^2_{school(j),i} + \alpha_4 W_j + \alpha_5 Z_j + \epsilon_{i,j} \quad (1)$$

Figure 4: School distances to *Site F*

Pre-treatment SIMCE's

Table 2: Pre-treatment SIMCE 1989-1990

	Math		Language	
Distance from school to Site F	**1989**	**1990**	**1989**	**1990**
	(1)	(2)	(3)	(4)
Less than 1.5 kilometers	61.22	67.92	60.45	67.71
	(11.26)	(7.70)	(6.86)	(6.74)
Between 1.5 and 2.5 kilometers	61.96	64.50	60.93	66.46
	(7.69)	(7.51)	(6.99)	(6.83)
Between 2.5 and 3.5 kilometers	59.24	69.17	64.10	69.42
	(10.63)	(9.01)	(8.70)	(6.28)
Between 3.5 and 4.5 kilometers	58.74	65.24	60.45	66.48
	(12.54)	(8.97)	(8.61)	(9.75)
More than 4.5 kilometers	59.57	64.74	61.20	65.66
	(14.29)	(12.00)	(13.65)	(10.70)

Note: Distances are expressed in kilometers. Standard deviations in parentheses.

PSU 2008

- We geo-referenced address of students' place of residence who took PSU 2008 in Arica.
- We included the distance from students' home to *Site F* as a covariate to analyze the effect of distance from home and school to *Site F* on PSU 2008 scores. Formally, we estimate:

$$\text{PSU}_{i,j} = \beta_0 + \beta_1 X_i + \beta_2 D_{home,i} + \beta_3 D^2_{home,i} + \beta_4 D_{school(j),i} + \beta_5 D^2_{school(j),i} + \beta_6 W_{i,j} + \beta_6 Z_j + \varepsilon_{i,j} \quad (2)$$

- We extended the analysis to estimate semiparametric models.

PSU 2008

- In both equations we control for a rich set of covariates for students and their families (gender, parent absence, parent education) and a set of dummy variables for school dependence.
- We include quadratic terms to capture nonlinear effects of the measures of distance.
- In the case of PSU scores, we exploit the advantage of observing individuals took SIMCE 2004 and PSU 2008, which allows us have longitudinal of student performance.
- We estimate equations (1) and (2) using OLS and we construct clusters for standard errors.

SIMCE 2003-2006

Table 3: SIMCE 2003-2006

Variable	Math				Language			
	2003	2004	2005	2006	2003	2004	2005	2006
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance from school to Site F	7.189	3.681	8.814**	12.365	10.418*	5.654*	7.138**	10.586
	(7.371)	(3.646)	(3.660)	(9.699)	(5.183)	(3.105)	(3.496)	(6.554)
Distance from school to Site F^2	-0.409	-0.411	-0.527**	-1.176	-0.675	-0.556**	-0.550**	-0.916*
	(0.576)	(0.254)	(0.247)	(0.758)	(0.408)	(0.225)	(0.240)	(0.517)
School located to the east of Site F	-9.736	-0.438	-14.746**	9.124	-5.103	-0.538	-13.077**	4.832
	(14.060)	(8.252)	(5.521)	(15.388)	(8.347)	(6.471)	(5.748)	(7.667)
Gender	7.985**	11.255***	7.944***	8.687**	-1.636	-4.019	-4.147*	-5.509*
	(3.221)	(2.436)	(2.829)	(4.142)	(2.122)	(2.484)	(2.373)	(3.026)
Mother with more than 12 years of education	18.507**	21.195***	11.751**	23.121***	15.971	26.794***	14.485***	22.180***
	(8.859)	(4.383)	(5.281)	(7.015)	(9.514)	(5.003)	(4.663)	(5.686)
Father with more than 12 years of education	25.706***	15.396***	1.101	28.953***	15.231**	16.914***	3.432	21.375***
	(8.953)	(4.846)	(3.675)	(9.747)	(7.310)	(5.418)	(3.793)	(5.247)
Voucher school	15.215	10.807**	14.239**	6.616	9.371	9.310**	13.738**	2.270
	(15.883)	(4.465)	(6.610)	(19.093)	(11.987)	(4.610)	(5.467)	(12.777)
Private school	56.841**	35.397***	43.805***	50.676*	42.422**	24.880***	36.913***	29.145**
	(20.619)	(5.805)	(7.232)	(27.608)	(15.949)	(8.273)	(9.014)	(12.715)
Observations	2459	1946	2786	2470	2459	1946	2786	2470

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

PSU 2008

Table 4: PSU 2008

	Math			Language		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance from *Site F* to home	14.265*	19.718***	16.467**	19.191	9.706*	6.541
	(8.086)	(6.585)	(6.719)	(12.277)	(5.718)	(5.772)
Distance from *Site F* to home^2	-2.962*	-3.734***	-3.348***	-3.612	-1.805*	-1.443
	(1.442)	(1.018)	(1.073)	(2.379)	(1.074)	(1.142)
Distance from *Site F* to school	8.699	7.535	10.702**	4.095	2.309	5.814
	(9.437)	(6.410)	(3.874)	(10.003)	(7.144)	(4.903)
Distance from *Site F* to school^2	-0.338	-0.176	-0.445	-0.182	0.250	-0.046
	(0.665)	(0.480)	(0.339)	(0.690)	(0.522)	(0.369)
Home located to the east of *Site F*	yes	yes	yes	yes	yes	yes
Simce 2004	no	yes	yes	no	yes	yes
Parent's absence	yes	yes	yes	yes	yes	yes
Parent's education	yes	yes	yes	yes	yes	yes
School dependence	no	no	yes	no	no	yes
Observations	1243	1243	1243	1243	1243	1243

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

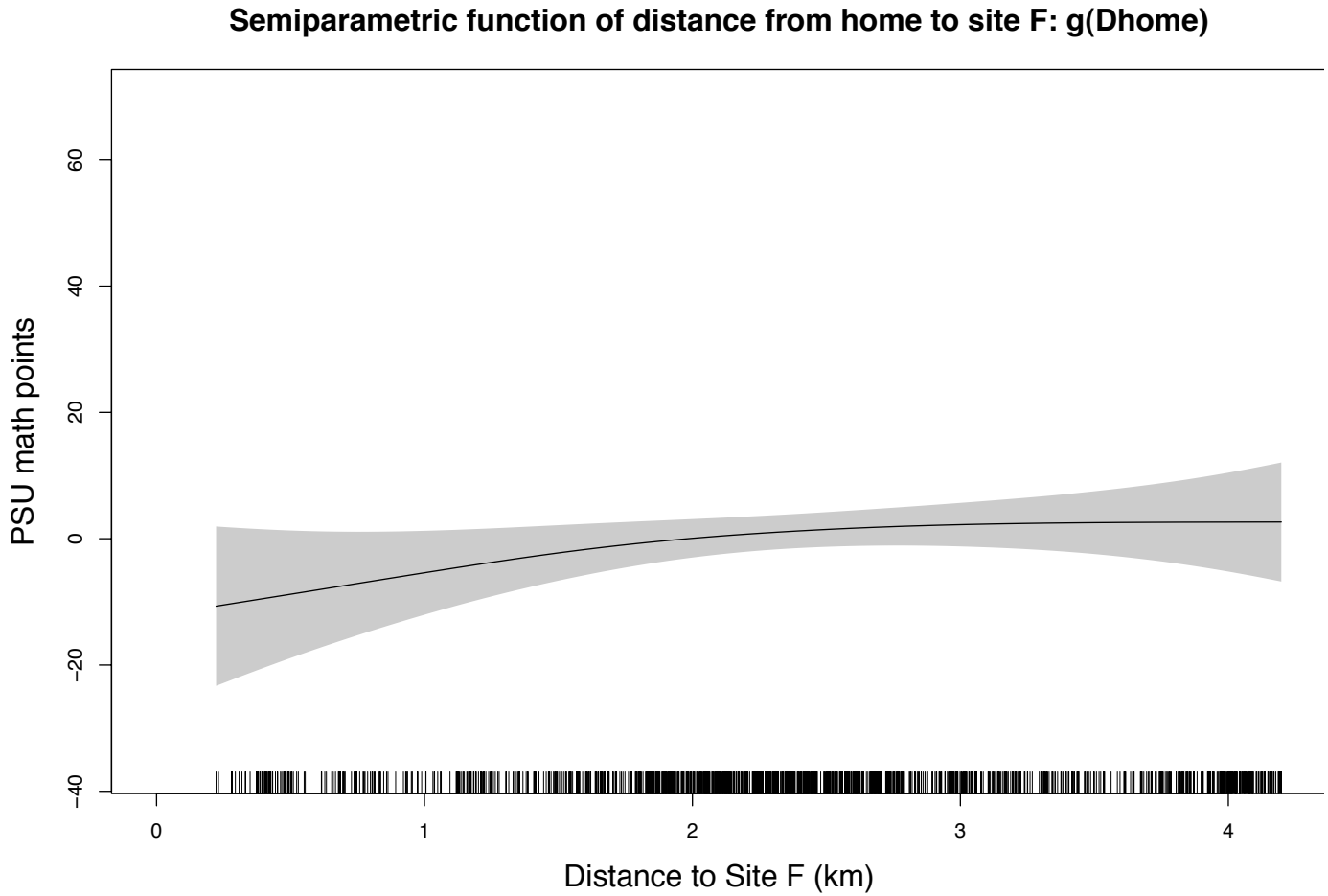
Semiparametric Estimation

We extend the analysis to a semiparametric model using Penalized Splines. Formally, we estimate:

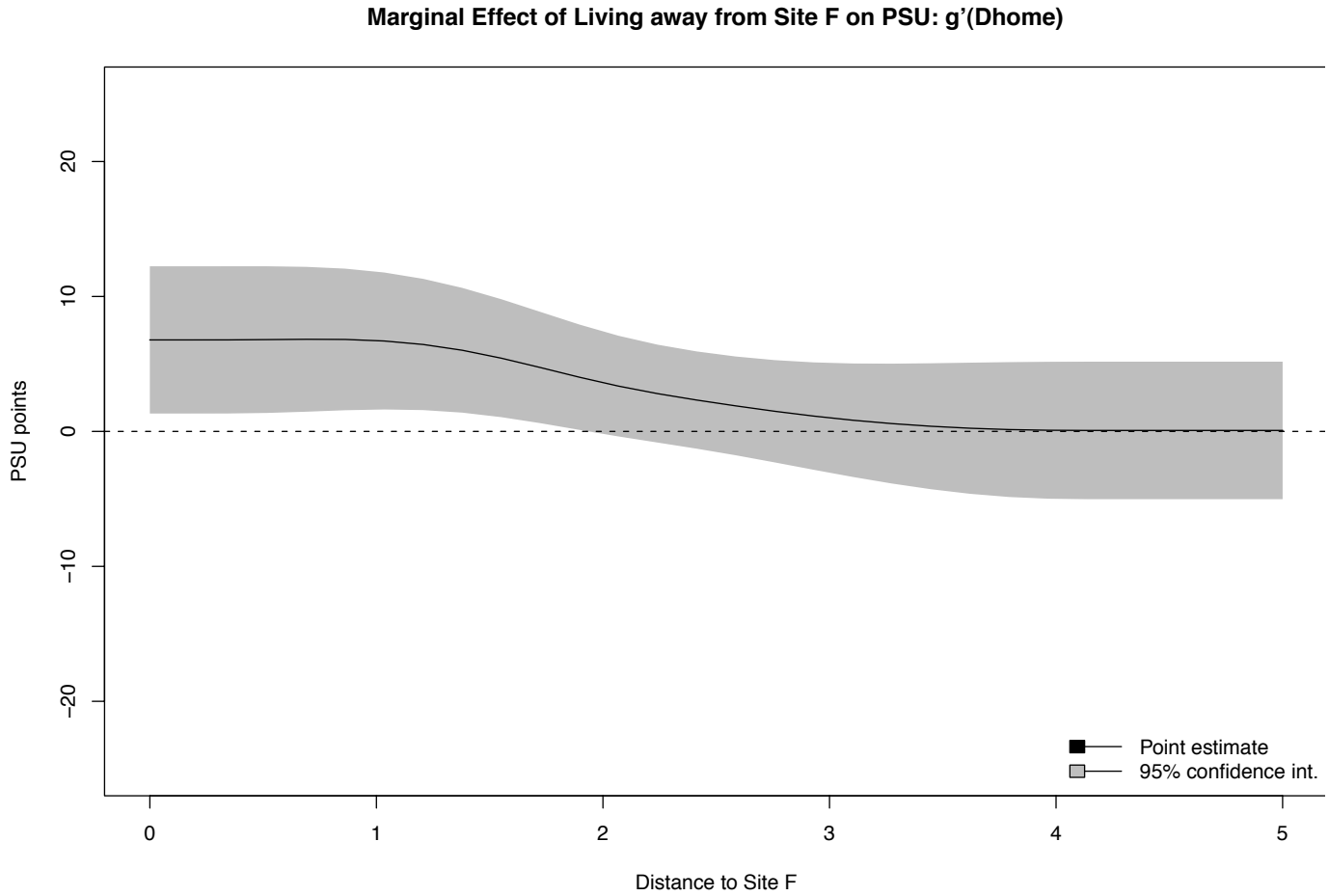
$$\text{PSU} = \beta_0 + \beta_1 X + \beta_2 Z + f(D_{home}) + g(D_{school}) + \varepsilon \tag{3}$$

We obtain $f(\cdot)$ and $g(\cdot)$ functions and we take their derivates.

Semiparametric Estimation of Math PSU scores



Semiparametric Estimation



Conclusions

- We estimate the effect of distance to a source of contamination on the academic performance of students in Arica using production functions of academic achievement.
- After controlling for individuals' characteristics such as gender, educational level of both parents and school characteristics, we find that distances from school and place of residence to *Site F* have significant effects on test scores (PSU and SIMCE).
- In particular, we estimate that a one kilometer reduction in the distance between place of residence and *Site F*, is associated with a 0.11 and 0.04 standard deviation reduction in math and language PSU score, respectively.

Conclusions

- In the case of SIMCE, the estimated effects are in the range of [0.06,0.15] and [0.09,0.18] standard deviations in math and language, respectively.
- We extended the regression analysis to semiparametric estimations that confirm the nonlinear and negative effect of distance from home to *Site F*.

Buckup Slides

Sources of lead pollution

Initially we consider three possible sources of pollution:

- Site F
- Line Railroad Arica-La Paz
- 26 points whose lead levels exceed the EPA's international standard of lead concentration in soil

These sources of pollution have particular characteristics in terms of time spent in city and volume of waste deposited.

To establish the relative importance of these sources of contamination, we use 390 blood test taken to to *Cerro Chuño* and *Los Industriales* during 2000. This sample reports two different and independent measures of lead on blood.

Blood Lead Levels

Figure 5: Histogram of blood lead levels: ISP and SERMUS samples

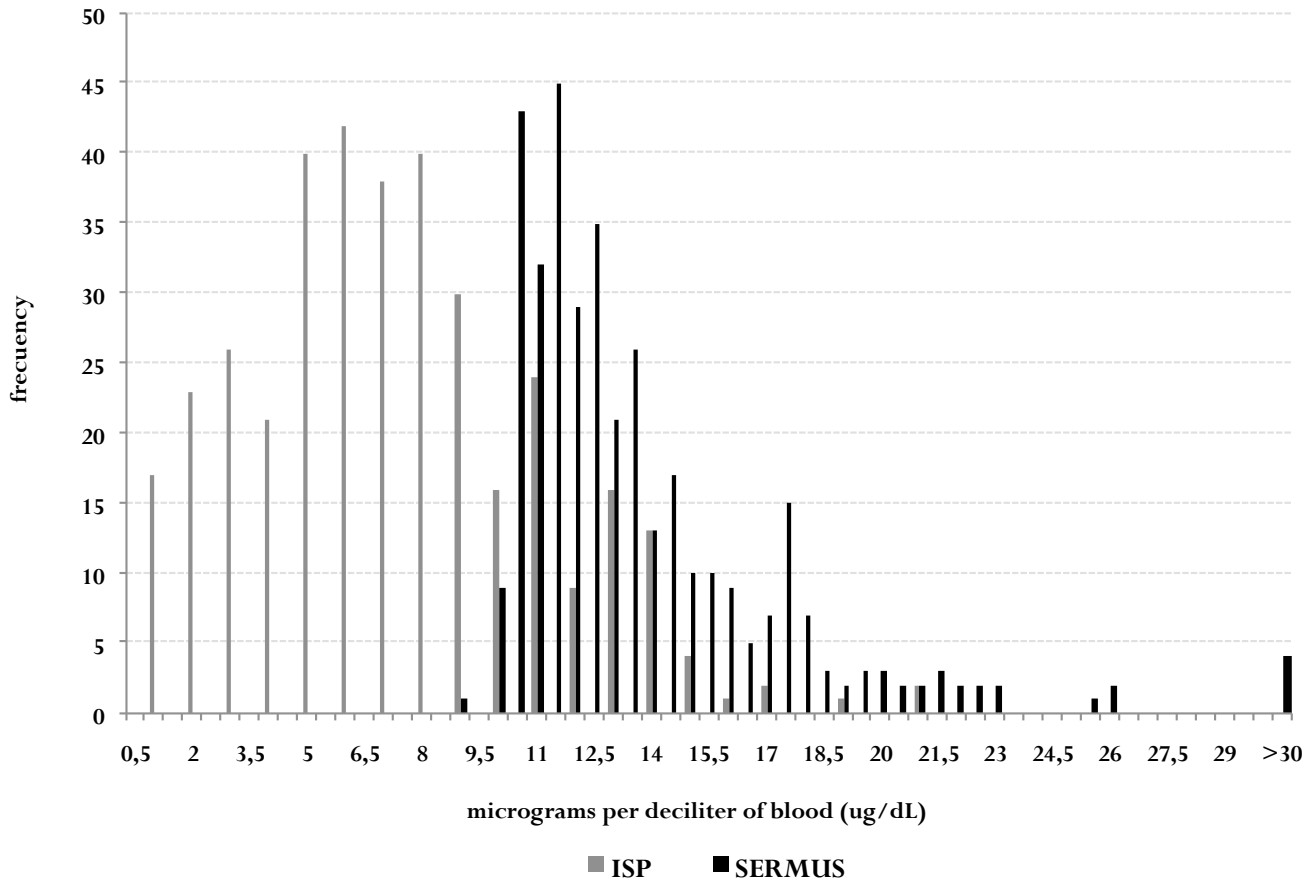
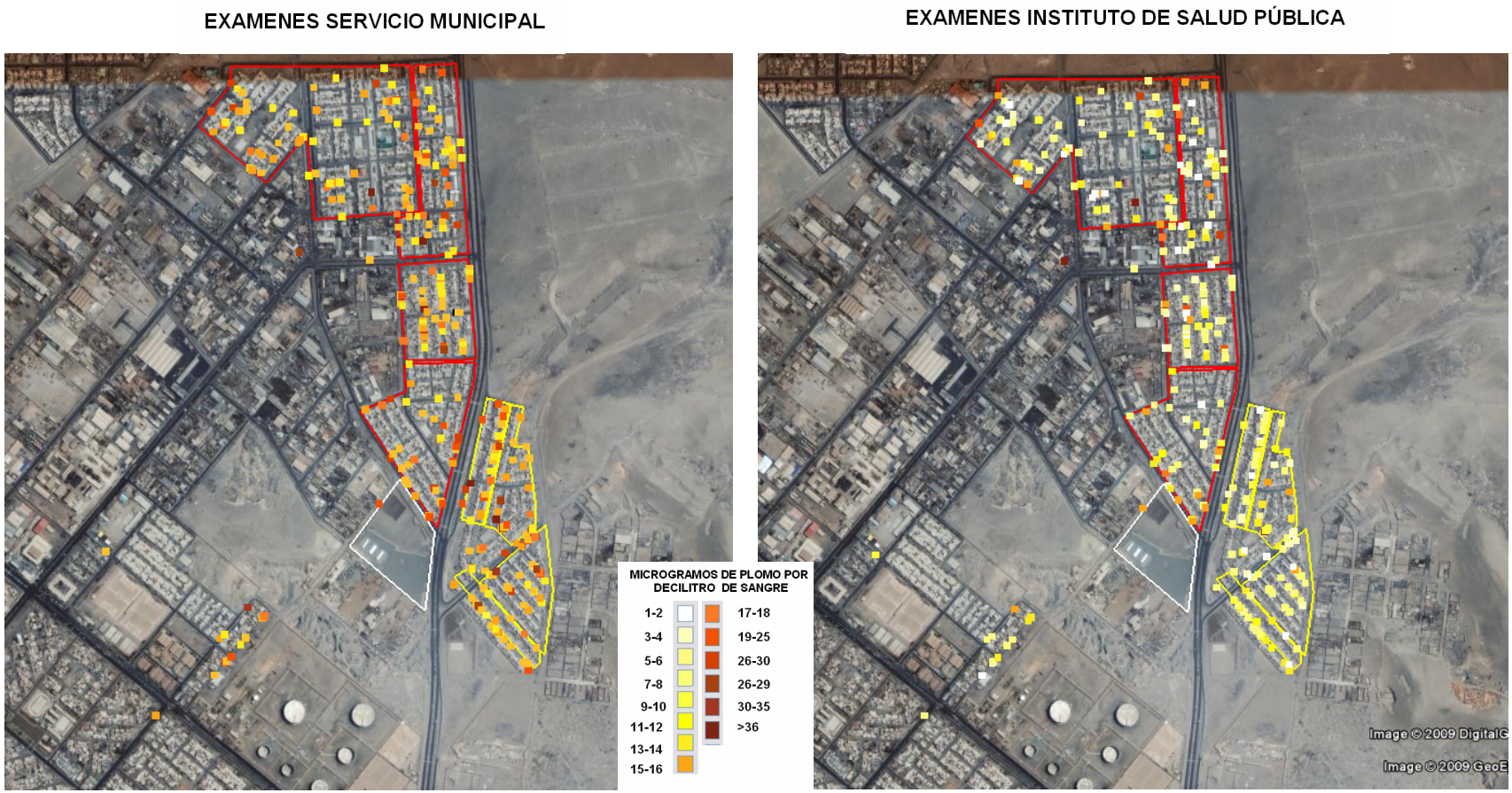


Figure 6: Blood lead test of *Cerro Chuño* and *Los Industriales* residents (2000)



Results of Medical Exams as a Function of Distances to Sources of Pollution

Table 5: Results of Medical Exams as a Function of Distances to Sources of Pollution

Variable	ISP Measure			SERMUS Measure		
Age	-0.142***	-0.153***	-0.156***	0.006	0.001	0.004
	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.020)
Gender	0.965***	0.834**	0.725*	1.011	1.105**	0.841**
	(0.379)	(0.376)	(0.379)	(0.426)	(0.440)	(0.398)
Distance from Site F to home	-13.446***			-9.660**		
Distance from Site F to home^2	9.031***			5.603**		
	(2.495)			(2.787)		
Distance from railway to home		-1.055			1.786	
Distance from railway to home^2		0.243			-0.576	
Distance from leaded soil to home			1.286			5.231**
Distance from leaded soil to home^2			-1.243			-4.052**
Constant	11.837***	8.720***	7.875***	16.456***	11.945***	12.039***
Observations	371	371	371	371	371	371
P-value Test	0.001	0.830	0.7616	0.0189	0.487	0.032

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In this case we control for influential observations using Cook's distance. All the numbers in the table are robust to this procedure.

Sources of lead pollution

- The results indicate a significant correlation between concentration of lead in blood and distance to *Site F*.
- On average, for every additional 100 meters further away from *Site F*, concentrations of lead in blood decrease by 1μg/dL.
- The above results are conclusive in pointing to *Site F* as the main source of lead pollution in Arica. So it is the distance to this point we will use henceforth to evaluate the effect of distance to a source of lead exposure on school performance.

Blood Lead Levels

Figure 7: Histogram of blood lead levels: ISP and SERMUS samples

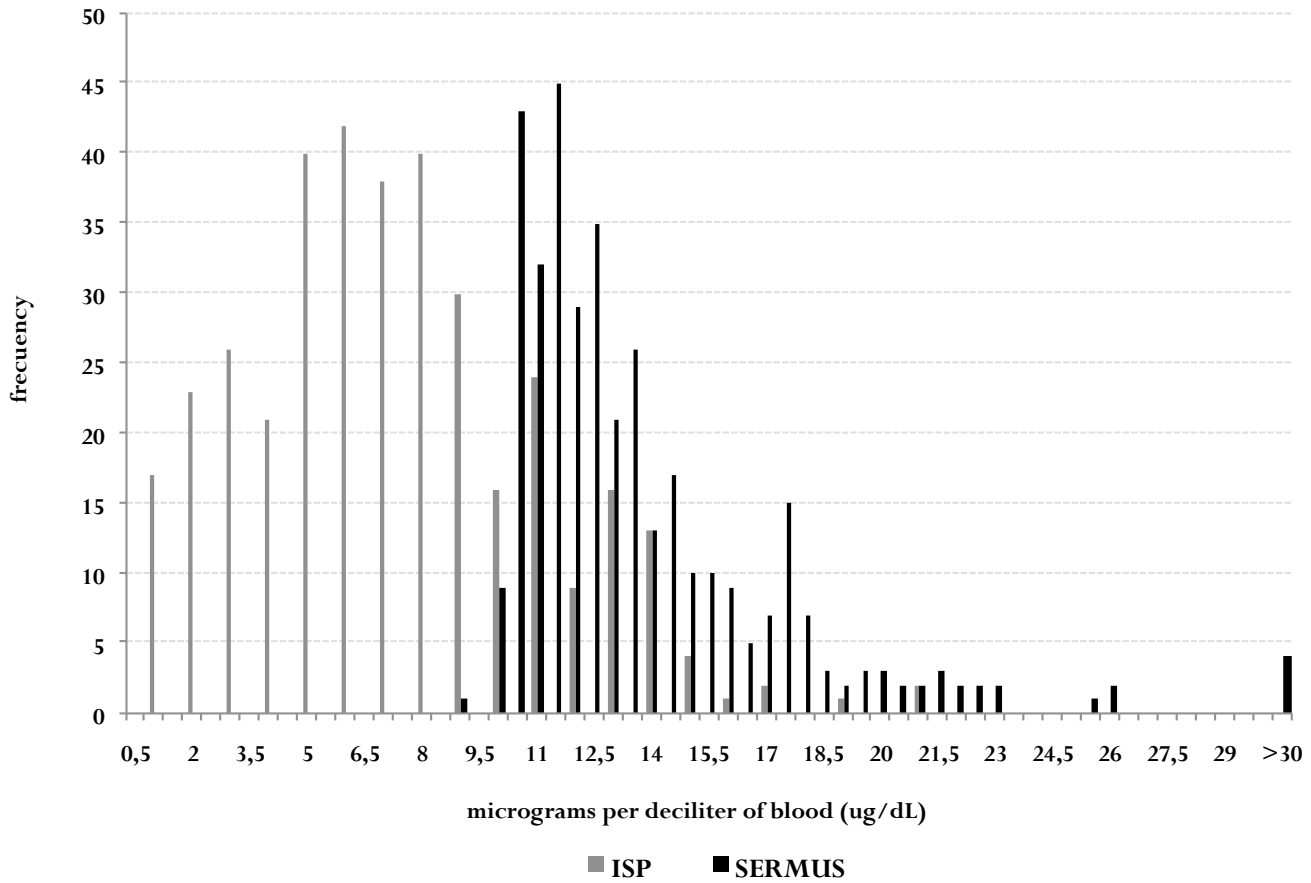
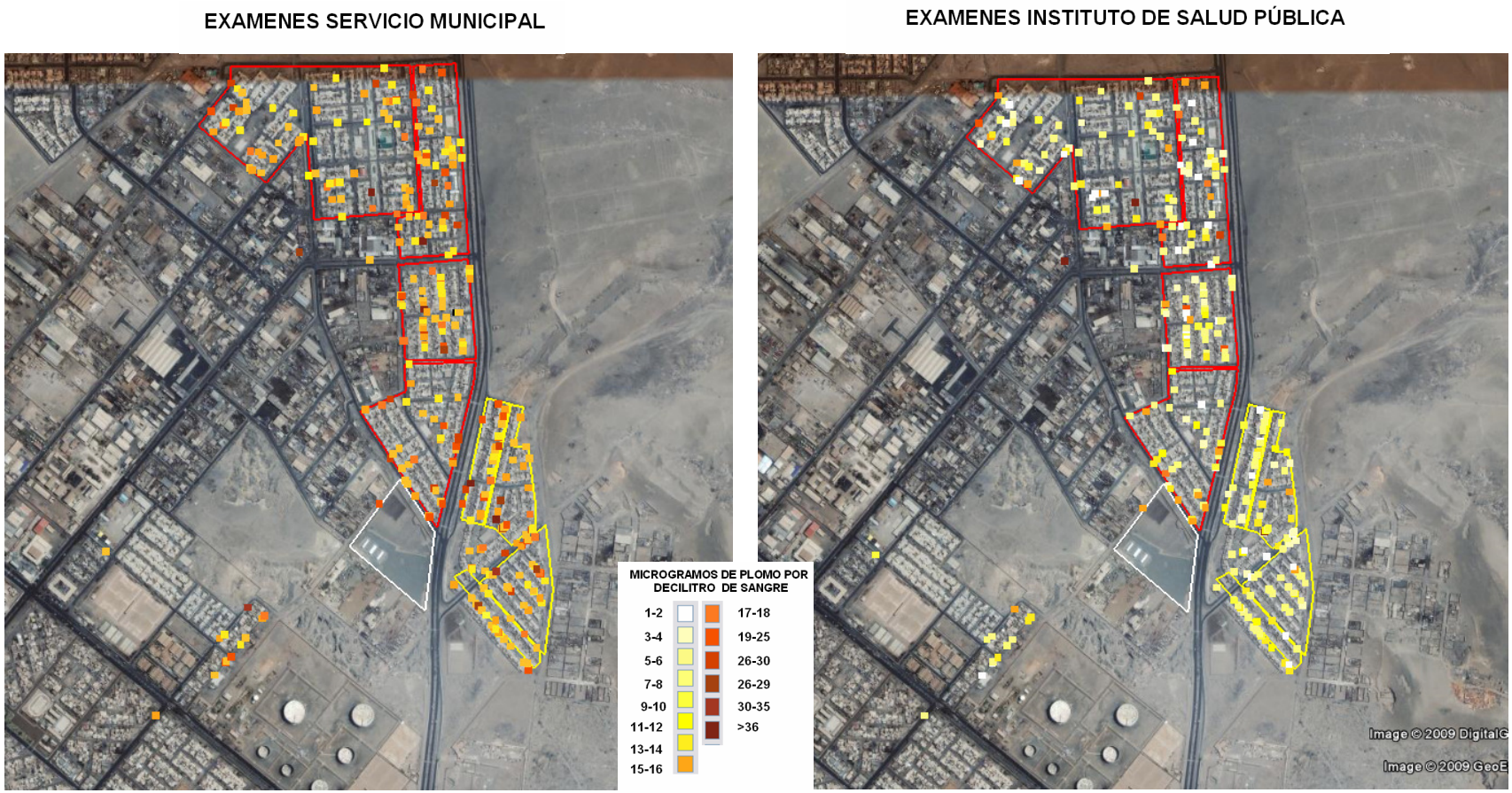


Figure 8: Blood lead test of *Cerro Chuño* and *Los Industriales* residents (2000)



Results of Medical Exams as a Function of Distances to Sources of Pollution

Table 6: Results of Medical Exams as a Function of Distances to Sources of Pollution

Variable	ISP Measure			SERMUS Measure		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance from Site F to home	-13.446***			-9.660**		
	(3.761)			(4.212)		
Distance from Site F to home^2	9.031***			5.603**		
	(2.495)			(2.787)		
Distance from railway to home		-1.055			1.786	
		(2.449)			(2.851)	
Distance from railway to home^2		0.243			-0.576	
		(1.144)			(1.336)	
Distance from leaded soil to home			1.286			5.231**
			(2.017)			(2.088)
Distance from leaded soil to home^2			-1.243			-4.052**
			(1.756)			(1.809)
Constant	11.837***	8.720***	7.875***	16.456***	11.945***	12.039***
	(1.225)	(1.212)	(0.503)	(1.381)	(1.384)	(0.519)
Observations	371	371	371	371	371	371
P-value Test	0.001	0.830	0.7616	0.0189	0.487	0.032

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. In this case we control for influential observations using Cook's distance. All the numbers in the table are robust to this procedure.

Sources of lead pollution

- The results indicate a significant correlation between concentration of lead in blood and distance to *Site F*.
- On average, for every additional 100 meters further away from *Site F*, concentrations of lead in blood decrease by 1μg/dL.
- The above results are conclusive in pointing to *Site F* as the main source of lead pollution in Arica. So it is the distance to this point we will use henceforth to evaluate the effect of distance to a source of lead exposure on school performance.

Semiparametric Estimation

$$\begin{aligned} PSU &= \beta_0 + \beta_1 X + \beta_2 Z + f(D_{home}) + g(D_{school}) + \varepsilon \\ &\cong \beta_0 + \beta_1 X + \beta_2 Z + \sum_{i=1}^{J} \alpha_i \rho_i(D_{home,i}) + \sum_{j=1}^{L} \gamma_j \psi_j(D_{school(j)}) + \varepsilon \end{aligned}$$

where $\{\rho_i\}$ y $\{\psi_j\}$ are polynomic funcions of third degree (cubic splines)

We will estimate using Penalized Splines. In other words,

Let $X = [X, Z, ...\rho_i..., ...\psi_j]$ and $\delta = [\beta, \alpha, \gamma]$

Semiparametric Estimation

$$\min_{\{\delta\}}(y - X\delta)'(y - X\delta) + \sum_{i=1}^{2} \theta_i \mathcal{P}_i$$

where the values of θ_i are the *smoothing parameters* and $\mathcal{P}_i$ the penalized matrix.

$$\mathcal{P}_1 = \int [f''(D_{home})]^2 dD_{home}; \quad \mathcal{P}_2 = \int [g''(D_{school})]^2 dD_{school}$$

These penalized matrix can be written as quadratic forms:

$$\mathcal{P}_1 = \alpha' \mathcal{S}_1 \alpha; \quad \mathcal{P}_2 = \gamma' \mathcal{S}_2 \gamma$$

Semiparametric Estimation

$$\min_{\{\delta\}} \; (y - X\delta)'(y - X\delta) + \sum_{i=1}^{2} \theta_i \delta' H_i \delta$$

where H_i is a block matrix that contains S_i in the diagonal and zeros out there. Then, we have a *Ridge Regression*

$$\hat{\delta} = \left(X'X + \sum_{i=1}^{2} \theta_i H_i \right)^{-1} X'y$$

where θ_i is obtain through generalized cross validation (GCV).

Robustness checks

- As a form to evaluate the robustness of our results, we analyze:
  - The effect of distance from home to three sources of pollution (*F*, Railroad Arica-La Paz and points to pollution on soil) on the probability of take PSU 2008 conditional on the student took SIMCE 2004 in schools located in Arica.
  - The effect of distance from school to *Site F* for those student who took PSU 2008 in schools located in Arica but who took SIMCE 2004 in other region.

Robustness checks

In order to evaluate the effect of proximity to a source of lead on the probability of take math and language PSU, we define:

$$Y_i = \begin{cases} 1 & \text{if the student i take PSU 2008} \\ 0 & \text{otherwise} \end{cases}$$

Then, the probability of take PSU conditional to have took SIMCE 2004, is given by the following expression:

$$\text{Prob}(Y_i = 1 | S_{2004,i} > 0, X_i) \quad = \quad \alpha_0 + \alpha_1 X_i + \alpha_2 D_{ij} + + \epsilon_i \tag{4}$$

where X_i y D_{ij} is a vector of individual's observable characteristics of the student i, and the minimum distance from home to a nearest source of pollution, respectively.

Probability of take PSU 2008

	Math			Language		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Distance from home to Site F	0.3114			0.3190		
	(0.2691)			(0.2734)		
Distance from home to Site F^2	-0.0587			-0.0593		
	(0.0480)			(0.0490)		
Distance from railway to home		-0.0400			-0.0079	
		(0.2032)			(0.1943)	
Distance from railway to home^2		0.0212			0.0159	
		(0.0477)			(0.0462)	
Distance from leaded soil to home			-0.1475			-0.1775
			(0.4160)			(0.3982)
Distance from leaded soil to home^2			0.1174			0.1334
			(0.2985)			(0.2798)
Age	-0.1384	-0.1452	0-.1483	-0.1420	-0.1492	-0.1519
	(0.0887)	(0.0870)	(0.0898)	(0.0912)	(0.0894)	(0.0923)
Gender	-0.0988	-0.0857	-0.0896	-0.1234	-0.1107	-0.1138
	(0.1001)	(.1022)	(0.1004)	(0.0998)	(0.0894)	(0.0976)
Parents education	si	si	si	si	si	si
Observations	1079	1079	1079	1079	1079	1079

Standard errors in parentheses

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

Students who take PSU 2008 in Arica and SIMCE 2004 in other place

- Additionaly, we estimate the equation (4) restricts the sample to those students who took PSU 2008 in schools of Arica but SIMCE 2004 in other place. Formally,

$$PSU_{i,j} = \beta_0 + \beta_1 X_i + \beta_2 D_{home,i} + \beta_3 D^2_{home,i} + \beta_4 D_{school(j),i} + \beta_5 D^2_{school(j),i} + \beta_6 Z_j + \varepsilon_{i,j} \quad (5)$$

Students who take PSU 2008 in Arica and SIMCE 2004 in other place

	Math			Language		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance from home to Site F	-9.101	33.561	28.395	-27.198	-20.113	-15.352
	(32.093)	(32.657)	(27.755)	(26.416)	(27.664)	(28.074)
Distance from home to Site F^2	1.111	-5.932	-4.506	3.540	1.341	-0.256
	(6.344)	(6.120)	(5.243)	(5.360)	(5.867)	(6.180)
Distance from school to Site F	-24.857	-69.577	-79.075	9.907	-31.941	-12.057
	(47.907)	(52.156)	(74.694)	(44.723)	(42.532)	(52.562)
Distance from school to Site F^2	3.384	12.276	13.120	-0.914	7.249	4.082
	(8.372)	(8.500)	(11.789)	(7.773)	(7.642)	(8.873)
Math Simce 2004 score		1.256***	1.192***			
		(0.272)	(0.285)			
Private school			28.767			51.302
			(38.943)			(47.723)
Voucher school			46.059			-0.730
			(34.904)			(25.795)
Language Simce 2004 score					1.138***	1.160***
					(0.163)	(0.211)
Constant	504.352***	224.699*	250.060*	415.616***	181.120*	148.642
	(66.341)	(113.601)	(130.559)	(72.880)	(90.176)	(120.114)
Parent absence	si	si	si	si	si	si
Parent education	si	si	si	si	si	si
Observations	78	78	78	78	78	78
P-value Test: distance from Site F to home	0.9241	0.5974	0.5683	0.3357	0.2078	0.1892
P-value Test: distance from Site F to school	0.8485	0.3630	0.5358	0.9290	0.5660	0.6845

Results of Medical Exams as a Function of Distances to Sources of Pollution

Table 7: Results of Medical Exams as a Function of Distances to Sources of Pollution

Variable	ISP Measure	SERMUS Measure
	(1)	(2)
Distance from Site F to home	-13.264***	-11.433***
	(3.829)	(3.956)
Distance from Site F to home^2	8.875***	6.714**
	(2.539)	(2.623)
Age	-0.147***	0.009
	(0.024)	(0.024)
Gender	0.937**	0.974**
	(0.385)	(0.398)
Constant	11.876***	16.801***
	(1.250)	(1.291)
Observations	365	365
R^2	0.1416	0.0491
P-value	0.0025	0.0027

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. In this case we control for influential observations using Cook's distance. All the numbers in the table are robust to this procedure.

Probability of take PSU 2008

Table 8: Probability of take PSU 2008

Distance from home to *Site F*	0.355*
	(0.185)
Distance from home to *Site* F^2	-0.067*
	0.036)
Gender	-0.116
	(0.110)
Parent's absence	yes
Parent's education	yes
School dependence	yes
Observations	1279

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

Students who took PSU 2008 in Arica and SIMCE 2004 in another place

Table 9: Students who took PSU 2008 in Arica and SIMCE 2004 in another place

	Math	**Language**
	(1)	(2)
Distance from home to *Site F*	28.395	11.503
	(27.755)	(24.869)
Distance from home to *Site* F^2	-4.506	-2.764
	(5.243)	(4.994)
Distance from school to *Site F*	-79.075	-31.798
	(74.694)	(68.628)
Distance from school to *Site* F^2	13.120	7.078
	(11.789)	(10.802)
Simce 2004	1.192***	1.150***
	(0.285)	(0.219)
Parent's absence	yes	yes
Parent's education	yes	yes
School dependence	yes	yes
Observations	78	78

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

Probability of take PSU 2008

Table 10: Probability of take PSU 2008

	Math	**Language**
Distance from home to *Site F*	0.367*	0.303
	(0.190)	(0.189)
Distance from home to *Site* F^2	-0.069*	-0.058
	(0.037)	(0.037)
Simce 2004	0.008***	0.006***
	(0.001)	(0.001)
Gender	-0.247**	-0.113
	(0.116)	(0.112)
Parents absence	yes	yes
Parents education	yes	yes
School dependence	yes	yes
Observations	1279	1279

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

Probability of take PSU 2008

Table 11: Probability of take PSU 2008

	Math	**Language**
Distance from home to *Site F*	0.360*	0.278
	(0.212)	(0.209)
Distance from home to *Site F*^{2}	-0.068*	-0.054
	(0.040)	(0.039)
Home located to the east of *Site F*	-0.013	-0.047
	(0.168)	(0.166)
Simce 2004	0.008***	0.006***
	(0.001)	(0.001)
Gender	-0.247**	-0.114
	(0.116)	(0.112)
Parents absence	yes	yes
Parents education	yes	yes
School dependence	yes	yes
Observations	1279	1279

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

Students who took PSU 2008 in Arica but SIMCE 2004 in another place

Table 12: Students who took PSU 2008 in Arica but SIMCE 2004 in another place

	Math	**Language**
	(1)	(2)
Distance from home to *Site F*	18.807	-36.471
	(29.909)	(24.550)
Distance from home to *Site F*^{2}	-3.054	2.891
	(5.420)	(5.595)
Distance from school to *Site F*	-72.307	2.705
	(75.778)	(52.325)
Distance from school to *Site F*^{2}	12.012	1.689
	(12.016)	(8.799)
Home located to the east of *Site F*	-23.843	-51.553
	(30.666)	(31.228)
Simce 2004	1.203***	1.194***
	(0.287)	(0.213)
Gender	18.663	24.921
	(22.252)	(17.305)
Parents absence	yes	yes
Parents education	yes	yes
School dependence	yes	yes
Observations	78	78

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

Students who took Simce 2004 in Arica but PSU 2009 in another place

Table 13: Students who took Simce 2004 in Arica but PSU 2009 in another place

	Math	**Language**
	(1)	(2)
Distance from home to *Site F*	-10.641	-10.325
	(28.499)	(31.997)
Distance from home to *Site* F^2	1.385	2.189
	(4.063)	(5.132)
Simce 20004	1.846***	1.327***
	(0.219)	(0.183)
Gender	-28.289*	14.426
	(15.503)	(15.645)
Parents absence	yes	yes
Parents education	yes	yes
School dependence	yes	yes
Observations	102	102

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

Students who took Simce 2004 in Arica but PSU 2009 in another place

Table 14: Students who took Simce 2004 in Arica but PSU 2009 in another place

	Math	**Language**
	(1)	(2)
Distance from home to *Site F*	-5.805	-9.773
	(28.164)	(32.063)
Distance from home to *Site F*^{2}	0.233	2.059
	(4.206)	(5.249)
School located to the east of *Site F*	16.545	1.792
	(20.918)	(22.364)
Simce 2004	1.835***	1.326***
	(0.218)	(0.186)
Gender	-28.721*	14.357
	(15.450)	(15.664)
Parents absence	yes	yes
Parents education	yes	yes
School dependence	yes	yes
Observations	102	102

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