

Microeconomía III. Problem Set 2

Gonzalo Escalona, Pablo Herrera

Semestre Otoño 2017

1. Ensayo

Hace aproximadamente un mes fueron publicados los resultados de la prueba SIMCE 2016, que mide a estudiantes de todo el país en las áreas de Comprensión de Lectura, Matemáticas e Historia y Ciencias. Según consigna el diario digital El Mostrador, "...existe casi una nula brecha que entre hombres y mujeres en Matemática", cuya diferencia fue entre uno y tres puntos. Además de que en comprensión lectora, las mujeres están por sobre los hombres en alrededor de diez puntos.

Surge entonces, un legítimo cuestionamiento acerca de la asociación de las matemáticas con los hombres y su constante mayor postulación a aquellas carreras "más exigentes". Lo anterior podría estar fuertemente vinculado a los roles de género y a la "incapacidad aprendida". (El Mostrador, 2017)

Priscila González, parte de la Red Chilena contra la Violencia hacia las Mujeres, define la incapacidad aprendida como el "fenómeno que afecta principalmente a las estudiantes mujeres y consiste en una creencia consciente e inconsciente, implícita o explícita, de la imposibilidad de realizar una tarea o asumir previa e infundadamente que no se tendrá un buen rendimiento en determinado asunto o área" (González, 2014).

Pero, ¿es esto un resultado de nuestros rasgos culturales como país? ¿Qué rol juega hoy el sistema educacional chileno en esta materia? En el siguiente ensayo se analizarán algunos estudios que apuntan específicamente a dicho cuestionamiento, con el fin de identificar qué papel está jugando la formación educacional en niños y niñas. Además de nombrar las consecuencias que tienen los roles de género, lo que suponemos es una realidad, en el mundo del trabajo.

Comunidad Mujer en su informe GET, habla sin rodeos acerca de la existencia de roles de género: "aquellos atributos que en los niños son considerados positivos y conducentes al "éxito", como la audacia, la ambición, el riesgo, el "don de mando", en ellas son considerados negativos". Esto tiene como resultado, una constante subestimación de las mujeres y con ello proyecciones consistentes en su vida escolar y profesional.

La organización, trae a colación también factores cotidianos como los "cuentos infantiles, dibujos animados, entre otros agentes de socialización temprana, que son tremendamente estereotipados y reproducen la tradicional división sexual del trabajo". Estableciendo que se generan tempranas desigualdades de condiciones entre niñas y niños lo que tendrá posteriormente efectos negativos durante la etapa escolar (Comunidad Mujer, 2016).

Valentina Paredes explica en su estudio "A teacher like me or a student like me? Role model versus teacher bias effect", que analiza los roles de género en la edad escolar, indica algo que no debiese ser sorpresa. "Las mujeres en Chile tienden a estudiar campos que conducen a trabajos relacionados con educación y salud, mientras que los hombres tienden a estudiar áreas

que conducen a carreras en ciencias y matemáticas, las que en promedio están asociadas con mayores salarios. Lo cual podría tener consecuencias en los retornos de la escolaridad de las mujeres”.

Paredes, además, obtiene resultados en base a datos del Simce 2009, los cuales indican que mayor presencia de mujeres entre el profesorado tienen efectos positivos en los resultados de las estudiantes mujeres. La académica también menciona, entre otros, un estudio realizado por Nixon and Robinson en 1999 en Estados Unidos, el cual estima efectos positivos de la presencia de mujeres entre los profesores de una institución, en la posterior inscripción y graduación de la universidad de mujeres, pero sin efectos en hombres (Paredes, 2013).

Ambos análisis llevan a caracterizar un problema que crece por un lado desde los roles de género culturales, y por otro lado desde las condiciones de los establecimientos educacionales que no dan soporte a contrarrestar esta dispareja situación.

Lo anterior es discutido en una de las recomendaciones del Informe GET, que promueve la participación de mujeres en Ciencia, Tecnología, Ingeniería y Matemáticas: “Las acciones afirmativas, como cupos reservados, son adecuadas en una primera instancia, mientras se genera una masa crítica de estudiantes que luego se transforme en modelo para las futuras generaciones” (Comunidad Mujer, 2016).

Según lo expuesto, la desigual situación entre hombres y mujeres, comienza desde temprano, por rasgos culturales que luego no son solucionados durante la formación escolar. Estas soluciones son prioridad si se busca saldar la brecha existente en el mundo del trabajo y en otros aspectos culturales que segregan entre hombres y mujeres.

2. Caracterización del sistema escolar chileno

2.1. Estadística Descriptiva

Los resultados obtenidos como Tablas y Gráficos de esta sección son adjuntados en los Anexos.

1. Para la construcción de los niveles de ingreso de los hogares se cuenta únicamente con el tramo de ingresos al que pertenece cada hogar. Ésto ofrece una dificultad para la creación de deciles balanceados en la cantidad de datos.

Para solucionar lo anterior, se supondrá que los ingresos de los individuos de cada tramo se comportarán como una normal con media en torno al valor medio entre los extremos de cada tramo, e.g. para el segundo tramo, el cual incluye entre \$100.001 y \$200.000, los ingresos de los hogares se distribuyen como una normal en torno a los \$150.000.

Es interesante resaltar que para ambas pruebas, y sin excepción, a medida que se sube en el decil de ingresos, los resultados mejoran. Lo mismo ocurre con la disponibilidad de internet y acceso a computadores.¹

2. Dentro de los resultados obtenidos, merece mención en primer lugar que para aumentos en los deciles de puntajes de ambas pruebas, siempre existe también aumento en los niveles

¹Cabe mencionar además que para identificar la disponibilidad de computadores en el hogar se buscó dicha información en el cuestionario a padres del Simce. En él, se indicaba que la pregunta cpad p05 01 se refería a lo anterior. Se corrige así el problema de las glosas incluidas en las bases de datos donde se indicaba que dos preguntas respondían el acceso a internet del estudiante.

de ingreso. Es casi tautológico, luego de las primeras tres tablas, decir que a mayores ingresos, mejores resultados de los estudiantes en las pruebas estandarizadas.

Además es consistente con dicha información que para aumentos en los deciles, el porcentaje de alumnos que pertenecen a una institución de dependencia municipal disminuyen permanentemente, lo que entrega luces acerca de los precarios resultados de las instituciones públicas de educación.

3. Se adjuntan los resultados en Anexos.
4. Los resultados obtenidos, tienen relación con el efecto par. Dicho efecto indica que las características del curso tendrán directo impacto en los resultados de los alumnos de manera individual
5. Se adjuntan los resultados en Anexos.

2.2. Análisis Gráfico

1. Por un lado en los resultados de la prueba de comprensión lectora, es posible identificar que los puntajes han ido disminuyendo en cada año, lo cual no ofrece mucha interpretación, pues esto ocurre para los tres tipos de dependencia. La disminución podría deberse a condiciones externas a la dependencia de los establecimientos como mayor dificultad en las pruebas.

Por otro lado, para la prueba de matemáticas, es posible notar que mientras los establecimientos Privados mantienen resultados sin mucha variación, los establecimientos Particulares Subvencionados y sobre todo Municipales presentan grandes mejoras en sus puntajes.

Sin embargo, de manera transversal existen fuertes brechas entre establecimientos Privados y las otras dos dependencias. Evidencia inequívoca de la desigual calidad entre dependencias.

2. Se adjuntan los resultados en Anexos.
3. Se adjuntan los resultados en Anexos.
4. Se adjuntan los resultados en Anexos.
5. Se adjuntan los resultados en Anexos.

3. SIMCE: Corte transversal

3.1.

En esta sección, se presentará estadística descriptiva para caracterizar y conocer de mejor manera a quienes sufren de Bullying y quienes afortunadamente no.

Características	Víctimas bullying	No víctimas bullying	Diferencia	Significancia
SIMCE Lenguaje	251,2	257,5	6,3	Si
SIMCE Matemáticas	265	269,5	4,5	Si
Promedio SIMCE	258,2	263,7	5,5	Si
Ingreso Familiar	\$525.346,3	\$552.783,5	\$27.437,2	Si
Proporción mujeres	53 %	45,8 %	7,2 %	Si
Municipal	24.070	40.340	-	-
Subvencionado	32.783	58.399	-	-
Particular	4.472	9.398	-	-

Cuadro 1: Caracterización de los grupos

Vemos que, en promedio (como serán todos los datos que analizaremos acá), las víctimas de *bullying* tienen menores puntajes en el SIMCE, así como un menor ingreso familiar. A su vez, es mayor la proporción de mujeres en quienes reciben agresiones físicas, verbales, sociales y/o electrónicas (*ciber-bullying*) por sobre el grupo que no es víctima de maltratos. El 37 % de los y las estudiantes de colegios municipales ha sido víctima de *bullying*, mientras que el 36 % de quienes están en subvencionados y el 32 % de los y las estudiantes de colegios particulares, ha sido víctima.

El análisis de regresión multivariado permite analizar y controlar por una serie de observables que están determinando, al mismo tiempo que el bullying, el puntaje en el SIMCE. Dá pie a ver posibles interacciones y a estimar el efecto de sufrir bullying de la manera más "limpia" posible. Si bien es probable que tengamos variables omitidas relevantes, como habilidad, cariño del hogar o la resiliencia del niño o niña, añadir las observables relevantes nos acercará a una estimación causal de manera más precisa.

3.2.

Se estimará el siguiente modelo, que conoceremos como **modelo o regresión (1)**

$$PtjeLeng_{ij} = \beta_0 + \beta_1 * Sexo_i + \beta_2 * Ln(Ingreso)_i + \beta_3 * dependencia_j + \beta_4 * Bullying_i + \mu_{ij}$$

```
. reg leng gen_alu lny dependencia bullying, r
```

Linear regression

Number of obs = 133167
F(4,133162) = 4599.86
Prob > F = 0.0000
R-squared = 0.1124
Root MSE = 51.66

leng	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
gen_alu	11.01061	.284616	38.69	0.000	10.45276	11.56845
lny	14.63838	.198192	73.86	0.000	14.24993	15.02683
dependencia	12.56528	.2795581	44.95	0.000	12.01735	13.11321
bullying	-4.637399	.2991527	-15.50	0.000	-5.223732	-4.051065
_cons	44.37312	2.322014	19.11	0.000	39.82202	48.92423

Figura 1: Regresión lineal del modelo propuesto

Acá el efecto de β_4 es *cómo cambia el promedio del SIMCE de lenguaje al pasar de ser un niño que no sufre bullying, a uno que si sufre*. Es decir, captura el efecto de ser víctima de Bullying. Vemos que el efecto de sufrir bullying es, en promedio, negativo, restando entre 4 y 5 puntos en el SIMCE. Además, vemos que con estos controles, la dependencia del establecimiento es muy importante (mejora en 12 puntos cuando se pasa de municipal a subvencionado y de este último a particular). Igual de relevantes al alza son el ingreso familiar y el género del individuo, donde el ser mujer afecta positivamente en 11 puntos en promedio. Todas son significativas al 5 % y al 1 %.

No sabemos si el efecto que recoge β_4 es un efecto causal, pues podemos estar en el caso en que la variable Bullying esté correlacionada con el error, sesgando la estimación que se busca.

3.3.

Estimamos un modelo probit para estimar los determinantes de sufrir bullying, considerando el género, el logaritmo natural del ingreso, el tipo de dependencia del colegio (municipal, subvencionado o particular pagado), promedio del individuo, efecto-par, capital cultural, la disciplina de los estudiantes del establecimiento y el consumo de alcohol y drogas.

Utilizando la base de datos, se construyen proxys sobre disciplina (mediante los comportamientos observados por los y las estudiantes) y capital cultural (número de libros en la casa). Como proxy de efecto-par, se considera el puntaje promedio del establecimiento.

```
. probit bullying gen_alu lny dependencia promedio libros promest disciplina alcdro
```

```
Iteration 0: log likelihood = -81799.514
Iteration 1: log likelihood = -80033.717
Iteration 2: log likelihood = -80032.915
Iteration 3: log likelihood = -80032.915
```

```
Probit regression                                Number of obs   =    125668
                                                LR chi2(8)      =    3533.20
                                                Prob > chi2     =    0.0000
Log likelihood = -80032.915                    Pseudo R2      =    0.0216
```

bullying	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
gen_alu	-.1770718	.007342	-24.12	0.000	-.1914618	-.1626818
lny	-.0167793	.0057443	-2.92	0.003	-.0280379	-.0055208
dependencia	.0285746	.0075699	3.77	0.000	.0137378	.0434114
promedio	-.0009624	.0000939	-10.25	0.000	-.0011464	-.0007785
libros	.0182334	.0042939	4.25	0.000	.0098175	.0266494
promest	.0002984	.0001539	1.94	0.052	-3.15e-06	.0006
disciplina	.448792	.012507	35.88	0.000	.4242787	.4733052
alcdro	.3369006	.010716	31.44	0.000	.3158976	.3579035
_cons	-1.088747	.07004	-15.54	0.000	-1.226023	-.9514712

Figura 2: Estimación Probit

En línea con los resultados de Harlow Roberts (2010), BLABLABLA

3.4.

Volveremos a estimar el modelo anterior, pero agregando controles:

$$Leng_{ij} = \beta_0 + \beta_1 * Género_i + \beta_2 * Ln(Ingreso)_i + \beta_3 * dependencia_j + \beta_4 * Bullying_i + \beta_5 * Libros_i + \beta_6 * Promedio_j + \beta_7 * Disciplina_j + \beta_8 * AlcDro_j$$

Donde la cantidad de libros existentes en el hogar son una proxy de capital cultural, el promedio del curso es una proxy del efecto par y alcdro es una dummy que se activa cuando en el establecimiento existe consumo de drogas y alcohol por parte del estudiantado. La disciplina, medida de 1 a 4, indica que a mayor número, existe mayor disciplina. El análisis resulta en la siguiente estimación de los parámetros:

```
. reg leng gen_alu lny dependencia bullying libros promest disciplina alcdro, r
```

Linear regression

Number of obs = 126855
F(8,126846) = 8447.67
Prob > F = 0.0000
R-squared = 0.3200
Root MSE = 45.137

leng	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
gen_alu	8.348017	.2558054	32.63	0.000	7.846643	8.849391
lny	-.6555578	.2004138	-3.27	0.001	-1.048365	-.2627503
dependencia	-4.158535	.2607457	-15.95	0.000	-4.669592	-3.647478
bullying	-3.121211	.2709345	-11.52	0.000	-3.652238	-2.590184
libros	5.039119	.1503926	33.51	0.000	4.744352	5.333886
promest	.8000251	.0041909	190.89	0.000	.7918109	.8082392
disciplina	9.980804	.4407897	22.64	0.000	9.116864	10.84474
alcdro	-4.857833	.4003485	-12.13	0.000	-5.642509	-4.073157
_cons	21.33929	2.409538	8.86	0.000	16.61664	26.06194

Figura 3: Análisis de regresión

Vemos que, al controlar por las variables añadidas al modelo, el valor de β_4 disminuye, teniendo un efecto menos negativo que en la estimación anterior. Esto se explica por la inclusión de más controles, por lo que era esperable la atenuación del efecto. Al incorporar la disciplina de un establecimiento, podríamos pensar que con menor disciplina existirá una mayor probabilidad de que un compañero violento a otro. A su vez, el controlar por el uso de alcohol y drogas dentro del establecimiento, nos hablaría sobre ciertas características no observables de los y las estudiantes que permitirían entender la presencia de bullying y disminuir el efecto propio de éste sobre el rendimiento medido por puntaje SIMCE.

3.5.

Un cluster es un grupo que contiene elementos, observaciones o individuos parecidos en su interior. Para realizar análisis por clusters, requerimos que cada grupo sea homogéneo en su interior y lo suficientemente heterogéneo a los otros como para poder realizar una separación efectiva. En nuestro caso, los clusters estarán conformados por la dependencia del establecimiento, siendo los 3 clusters: Municipal, Subvencionado y Particular. Podemos aplicar esta técnica en nuestro análisis, ya que es similar a la evaluación por tramos, donde creemos que no existen efectos lineales entre los distintos tipos de establecimientos.

En este sentido, el modelo anteriormente mencionado asume que cambiar de un colegio municipal a uno subvencionado, tiene el mismo efecto en el rendimiento escolar que pasar de un colegio subvencionado a uno particular. Un efecto lineal, que entendemos no es posible analizar con datos de corte transversal, donde acá analizamos la diferencia del valor esperado del puntaje, a distintos tipos de dependencia del establecimiento. Para nosotros existen fuertes intuiciones de que no existe un efecto lineal, por lo que separamos en clusters según el criterio anteriormente señalado. La estimación es como sigue:

$$Leng_{ij} = \beta_1 * Municipal_j + \beta_2 * Subvencionado_j + \beta_3 * Particular_j + \beta_4 * Género_i + \beta_5 * Ln(Ingreso)_i + \beta_6 * Bullying_i + \beta_7 * Libros_i + \beta_8 * Promedio_j + \beta_9 * Disciplina_j + \beta_{10} * AlcDro_j + \mu_{ij}$$

Resultando en:

```
. reg leng municipal subvencionado particular gen_alu lny bullying libros promest disciplina alcdro, r noconst
```

Linear regression

Number of obs = 126855
F(10,126845) = .
Prob > F = 0.0000
R-squared = 0.9707
Root MSE = 45.135

leng	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
municipal	14.13097	2.644842	5.34	0.000	8.947128	19.31482
subvencionado	10.4321	2.705098	3.86	0.000	5.130151	15.73404
particular	4.454782	2.999091	1.49	0.137	-1.423384	10.33295
gen_alu	8.339006	.2558093	32.60	0.000	7.837624	8.840387
lny	-.4474016	.209552	-2.14	0.033	-.85812	-.0366833
bullying	-3.130439	.2709263	-11.55	0.000	-3.66145	-2.599428
libros	5.050697	.1504288	33.58	0.000	4.755859	5.345535
promest	.8006823	.0041922	190.99	0.000	.7924657	.8088989
disciplina	9.990616	.4407545	22.67	0.000	9.126744	10.85449
alcdro	-4.809158	.4006268	-12.00	0.000	-5.59438	-4.023936

Figura 4: Análisis de Regresión

Donde vemos que, sorprendentemente, pertenecer a un colegio municipal tiene mejores efectos que uno subvencionado, que a la vez tiene mejores efectos que uno particular, teniendo este último un efecto no significa. Esto se debe a que, *ceteris paribus*”, los colegios municipales no son malos como se cree, si no que son otros factores como el capital cultural o la disciplina los determinantes últimos del rendimiento escolar.

3.6.

Esta estimación, por variadas razones, tiene problemas y limitantes que no permiten encontrar un efecto causal en todas las variables. En primer lugar, hay variables relevantes que han sido omitidas no por decisión propia, si no por su inexistencia en la base de datos. ¿Cómo podemos medir el cariño y la estimulación temprana y su efecto en el rendimiento escolar? Sabemos del área de la psicología que éstos aspectos de la niñez aportan positivamente, pero, no la magnitud.

Por otro lado, se podría argumentar endogeneidad en las variables libros (proxy de capital cultural) y de la dependencia del establecimiento. La distribución de los recursos dentro del hogar, correlacionado fuertemente con las preferencias, podría no estar enfocada en libros y sí a esparcimiento, alimentación, etc. Por otro lado, hay quienes manifiestan que el tipo de colegio (municipal, subvencionado o particular) es elegido por padres y madres. La discusión que escapa a esta entrega es si efectivamente existe la “libertad de elegir.”^{el} establecimiento o si sólo se envía al niño o niña al establecimiento que sus padres pueden costear.

Como curso de acción, podemos utilizar variables instrumentales o efectos fijos, para lo que se requieren datos de panel.

4. SIMCE: Panel

Se construye el panel para tener en una misma base de datos a los y las estudiantes que rindieron el SIMCE 2011, cuando cursaban 8vo básico, y el SIMCE 2013, cuando estaban en segundo medio.

4.1.

Realizamos una estimación pooled (de “agrupado”) y controlamos por el puntaje SIMCE 2011 para estimar su efecto en el puntaje del 2013. La estimación pooled nos indica crear una variable auxiliar que se active cuando el año es 2013, multiplicando posteriormente cada una de las variables por esta auxiliar, generando interacciones entre sí y agregando estas interacciones a la regresión.

```
. reg promedio2013 gen_alu lny dependencia bullying g1_lny g1_dependencia g1_bullying prom2011, r
note: g1_lny omitted because of collinearity
note: g1_dependencia omitted because of collinearity
note: g1_bullying omitted because of collinearity
```

```
Linear regression                               Number of obs = 129186
                                                F( 5,129180) =11257.25
                                                Prob > F      = 0.0000
                                                R-squared     = 0.2808
                                                Root MSE     = 45.99
```

promedio2013	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
gen_alu	.2845884	.2571994	1.11	0.269	-.2195179	.7886946
lny	17.65239	.1830085	96.46	0.000	17.29369	18.01108
dependencia	12.62999	.2614633	48.31	0.000	12.11753	13.14246
bullying	-3.586121	.2695841	-13.30	0.000	-4.114501	-3.057741
g1_lny	0	(omitted)				
g1_dependencia	0	(omitted)				
g1_bullying	0	(omitted)				
prom2011	.1482296	.0012699	116.73	0.000	.1457406	.1507186
_cons	-13.78387	2.091781	-6.59	0.000	-17.88373	-9.68402

Figura 5: Regresión Pooled

El mismo output de Stata nos indica que elimina las variables interactivas por presencia de colinealidad. El efecto que captura el beta del promedio del 2011, dice que ante un aumento de un punto del SIMCE 2011, el promedio 2013 aumentará en 0.14 puntos. Un número significativo, pero poco relevante por su magnitud pequeña.

4.2.

Para aplicar efectos fijos, debemos saber que el supuesto clave es que los no-observables se mantienen fijos en el tiempo, por lo que al aplicar esta estrategia, eliminamos los no-observables y podremos obtener efectos puros de los observables considerados. De esta manera, las únicas estimaciones que podemos realizar es sobre los observables que efectivamente cambian en el tiempo. Si vemos los problemas, en paneles cortos (como este) hay estimaciones que no se realizan de manera consistente.

En cuanto a efectos aleatorios, asume que los no-observables son aleatorios y que éstos no están correlacionados con los regresores. Se realizará una estimación por FGLS, pero que será inconsistente si se rompe el segundo supuesto mencionado. Sabemos que tendremos este problema, pues la literatura indica que los no-observables como la estimulación temprana, tienen efectos en el rendimiento escolar de los niños y niñas, por lo tanto, hay correlación de los no-observables con los regresores.

En el contexto de la primera ecuación de la parte 3², estimamos por efectos fijos y aleatorios, para luego realizar el test de Hausman. Las tablas se encuentran en el Anexo 1

4.3.

Realizamos una estimación por efectos fijos a nivel de individuo. Como mencionamos antes, esto permite limpiar los no-observables y su contaminación de los parámetros de interés. Para esto, las variables que siguen siendo necesarias, que permanecen como control, son aquellas que varían en el tiempo. Con ello, tenemos la estimación por efectos fijos.

```
. xtreg leng gen_alu lny dependencia bullying, fe i(mrun)
```

Fixed-effects (within) regression	Number of obs	=	313839
Group variable: mrun	Number of groups	=	222096
R-sq: within = 0.0051	Obs per group: min =		1
between = 0.0002	avg =		1.4
overall = 0.0000	max =		2
corr(u_i, Xb) = -0.0669	F(4,91739)	=	116.44
	Prob > F	=	0.0000

	leng	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	gen_alu	4.225459	1.629462	2.59	0.010	1.031729 7.419188
	lny	-1.760252	.2605364	-6.76	0.000	-2.270901 -1.249604
	dependencia	.8183496	.2876532	2.84	0.004	.2545522 1.382147
	bullying	-4.453744	.2337264	-19.06	0.000	-4.911846 -3.995643
	_cons	275.9734	3.404631	81.06	0.000	269.3004 282.6464

sigma_u	50.463197
sigma_e	29.193902
rho	.7492411 (fraction of variance due to u_i)

F test that all u_i=0:	F(222095, 91739) =	3.54	Prob > F = 0.0000
------------------------	--------------------	------	-------------------

Figura 6: Análisis por Efectos Fijos a nivel de individuo

Vemos que el bullying sigue teniendo un efecto similar al de la primera estimación de la parte 3. Esto, pues en este caso tenemos menos controles, ya que se asumen fijos ciertos aspectos como el capital cultural.

4.4.

Realizamos el mismo análisis anterior, pero ahora a nivel de establecimientos:

$$^2PtjeLeng_{ij} = \beta_0 + \beta_1 * Sexo_i + \beta_2 * Ln(Ingreso)_i + \beta_3 * dependencia_j + \beta_4 * Bullying_i + \mu_{ij}$$

```
. xtreg leng gen_alu lny dependencia bullying,fe i(rbd)
warning: existing panel variable is not rbd
```

Fixed-effects (within) regression

Number of obs = 313839

Group variable: rbd

Number of groups = 6575

R-sq: within = 0.0131

between = 0.2350

overall = 0.1066

Obs per group: min = 1

avg = 47.7

max = 589

F(4,307260) = 1020.11

Prob > F = 0.0000

corr(u_i, Xb) = 0.2184

	leng	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
gen_alu	9.001158	.1689706	53.27	0.000	8.66998	9.332335
lny	4.811074	.1434984	33.53	0.000	4.529822	5.092327
dependencia	13.53532	6.094428	2.22	0.026	1.590418	25.48023
bullying	-3.40026	.2156068	-15.77	0.000	-3.822843	-2.977676
_cons	168.9008	10.26604	16.45	0.000	148.7797	189.022
sigma_u	24.638097					
sigma_e	43.991809					
rho	.23877286	(fraction of variance due to u_i)				

F test that all u_i=0: F(6574, 307260) = 12.09 Prob > F = 0.0000

Figura 7: Análisis por Efectos Fijos a nivel de establecimiento

El efecto del Bullying sigue siendo significativo y negativo, pero es menor que el caso del análisis a nivel individuo.

4.5.

Realizamos la regresión (1) para el año 2011, cuya tabla de regresión se encuentra en el anexo 2.

Vemos que el efecto del bullying sobre el promedio de lenguaje se duplica respecto del año 2011, posiblemente por una mejora en las políticas enfocadas a prevenirlo o porque los y las estudiantes el año 2013 ya están más maduros y dejan, en parte, de lado estas actitudes.

Además, se realiza un análisis estimando la ecuación aplicando primeras diferencias. Esto es:

$$\Delta Leng_{ij} = \gamma_0 + \gamma_1 * \Delta Ln(Ingreso)_i + \gamma_2 * \Delta Bullying_i + \eta_{ij}$$

El efecto del Bullying, como se observa en la tabla del Anexo 3, es de ¡54 puntos en promedio! Acá se obtiene la diferencia para quienes sufrían bullying en 2011 y no lo hacen en 2013, es decir, los 54 puntos son lo que aumenta el rendimiento escolar habiendo sido víctima de bullying y dejar de serlo.

4.6.

Realizamos una regresión de las diferencias como la del ejercicio 4.5, añadiendo como regresor el puntaje del SIMCE 2011. En el anexo 4 se encuentra la tabla disponible. Siendo el nuevo regresor una variable significativa pero muy poco relevante, lo interesante es que provoca un efecto aún mayor de dejar de ser un niño o niña que sufre bullying. El coeficiente de la diferencia Bullying pasa de 54 a 62 puntos.

Finalmente, agregamos controles del tipo género, dependencia, diferencia en el consumo de

drogas y alcohol, así como la diferencia del promedio del establecimiento entre el SIMCE 2011 y 2013. Esta tabla se adjunta en el Anexo 5

5. Referencias

El Mostrador. (2017). <http://www.elmostrador.cl/braga/2017/05/02/educacion-sexista-en-chile-ser-hombre-o-mujer-no-da-lo-mismo/>
 González. P. (2014) en: <http://radio.uchile.cl/2014/11/01/son-los-cuentos-infantiles-un-arma-del-sexismo/>
 Comunidad Mujer. (2016). Género, educación y trabajo - La brecha persistente. Informe GET.
 Paredes, V.(2013). A teacher like me or a student like me? Role model versus teacher bias effect.

6. Anexos Sección 2

6.1. Anexo 2.1.1.

Figura 8

Ing10	Lenguaje	Matemá-s	Libros	Internet	Comput-r	Munici-l	Parasub	Partic-r
1	236.4129	235.8816	82.98718	.3199541	.6348021	.5884661	.4099647	.0015692
2	241.3362	243.5889	77.71296	.4093889	.7403996	.5570522	.4414438	.001504
3	244.7173	250.6553	74.3081	.5103602	.7959124	.5077154	.4907153	.0015692
4	248.2404	256.2705	70.04948	.611774	.8641005	.4562872	.5421435	.0015693
5	250.022	258.6152	69.35011	.6410827	.8759075	.4364742	.5603217	.0032041
6	255.2786	266.5878	69.08525	.7505055	.9243804	.3832222	.6107624	.0060154
7	259.2567	273.3775	68.82027	.8226117	.9490676	.3323743	.6573596	.0102661
8	264.6299	282.2888	71.11027	.88829	.972507	.2640251	.7048516	.0311233
9	273.5737	296.6452	77.34443	.9462459	.9871208	.1796247	.7139214	.1064539
10	292.6732	330.0724	99.60119	.9871336	.9957062	.048192	.356176	.595632
Total	256.7164	269.5833	76.03613	.6991953	.8748793	.3753449	.5487661	.0758889

Puntajes por Decil de Ingreso

6.2. Anexo 2.1.2.

Mate10	Inghogar	Libros	Internet	Comput~r	Munici~l	Parsub	Partic~r
1	298991.1	77.28895	.5634999	.7650335	.5964596	.3949651	.0085753
2	311869.6	75.97439	.5680315	.7910284	.5682194	.4236073	.0081733
3	334656.3	74.42396	.5944725	.8134206	.5247217	.4647478	.0105305
4	365017.5	74.15882	.6249906	.8440762	.4771199	.5076267	.0152534
5	399696.9	72.69059	.6552653	.8622582	.4276115	.5502733	.0221152
6	452489.9	72.12521	.6944855	.889038	.3614651	.6014049	.03713
7	530381.2	72.46129	.7457975	.9156823	.3012084	.6347591	.0640325
8	644464.3	74.73795	.7866276	.9328056	.2430263	.6425848	.1143889
9	800335	78.13733	.8347987	.9559874	.2057289	.5993779	.1948931
10	1068857	87.06037	.8788852	.9715445	.1699353	.4945579	.3355068
Total	529249.4	75.94545	.7024319	.877639	.387565	.5313809	.0810541

Leng10	Inghogar	Libros	Internet	Comput~r	Munici~l	Parsub	Partic~r
1	357932.7	76.82729	.6089754	.8009758	.5511073	.4279487	.020944
2	363886.9	75.85226	.6056807	.8067505	.5214669	.455388	.0231451
3	381711.1	75.0701	.6192481	.8245577	.4806334	.4931603	.0262063
4	403617.6	73.60564	.6428103	.842015	.4355493	.5299334	.0345173
5	449439.1	73.84346	.6649455	.8652453	.3984363	.5530998	.0484639
6	491154.4	72.79183	.6974772	.8843497	.3639371	.5736694	.0623935
7	558748.4	73.97769	.7276526	.902525	.3329658	.5813872	.085647
8	626500	75.06481	.759584	.9205717	.2939232	.5904514	.1156255
9	708183.8	77.34555	.8013758	.9425673	.2691594	.5761616	.154679
10	886218.8	84.83607	.8414088	.9590153	.2273115	.5329491	.2397394
Total	528919.2	75.9442	.7020441	.8776107	.3874577	.5314108	.0811315

Figura 9: Ingresos por Deciles de Puntaje

6.3. Anexo 2.1.3.

cod_depe2	mean
Municipal	37.61789
Particular subve	38.29585
Particular pagad	29.17047
Total	37.31045

Figura 10: Estudiantes por sala según tipo de dependencia del establecimiento

cod_depe2	mean
Municipal	37.54056
Particular subve	38.30518
Particular pagad	29.12331
Total	37.32968

Figura 11: Establecimientos Mixtos

cod_depe2	mean
Municipal	39.39907
Particular subve	39.51553
Particular pagad	30.73764
Total	38.29008

Figura 12: Establecimientos sólo de Hombres

cod_depe2	mean
Municipal	37.14024
Particular subve	37.33724
Particular pagad	27.89623
Total	36.31845

Figura 13: Establecimientos sólo de Mujeres

6.4. Anexo 2.1.4.

Ingcur5	Lenguaje	Matemá~s
1	228.5262	225.9161
2	238.8595	242.8198
3	248.1957	257.5933
4	265.7729	282.3939
5	286.7358	319.6844
Total	254.5654	267.2214

Figura 14: Por quintil ingreso

Escpad5	Lenguaje	Matemá~s
1	237.7422	240.7193
2	250.0984	259.1169
3	262.2433	278.0226
4	271.656	293.911
5	287.8944	318.1544
Total	257.0643	270.3121

Figura 15: Por quintil educ padre

Escmad5	Lenguaje	Matemá~s
1	236.6391	238.1095
2	250.5106	259.9572
3	262.3929	278.6661
4	273.0494	295.807
5	289.5535	320.0433
Total	256.8338	269.7912

Figura 16: Por quintil educ madre

6.5. Anexo 2.2.1.

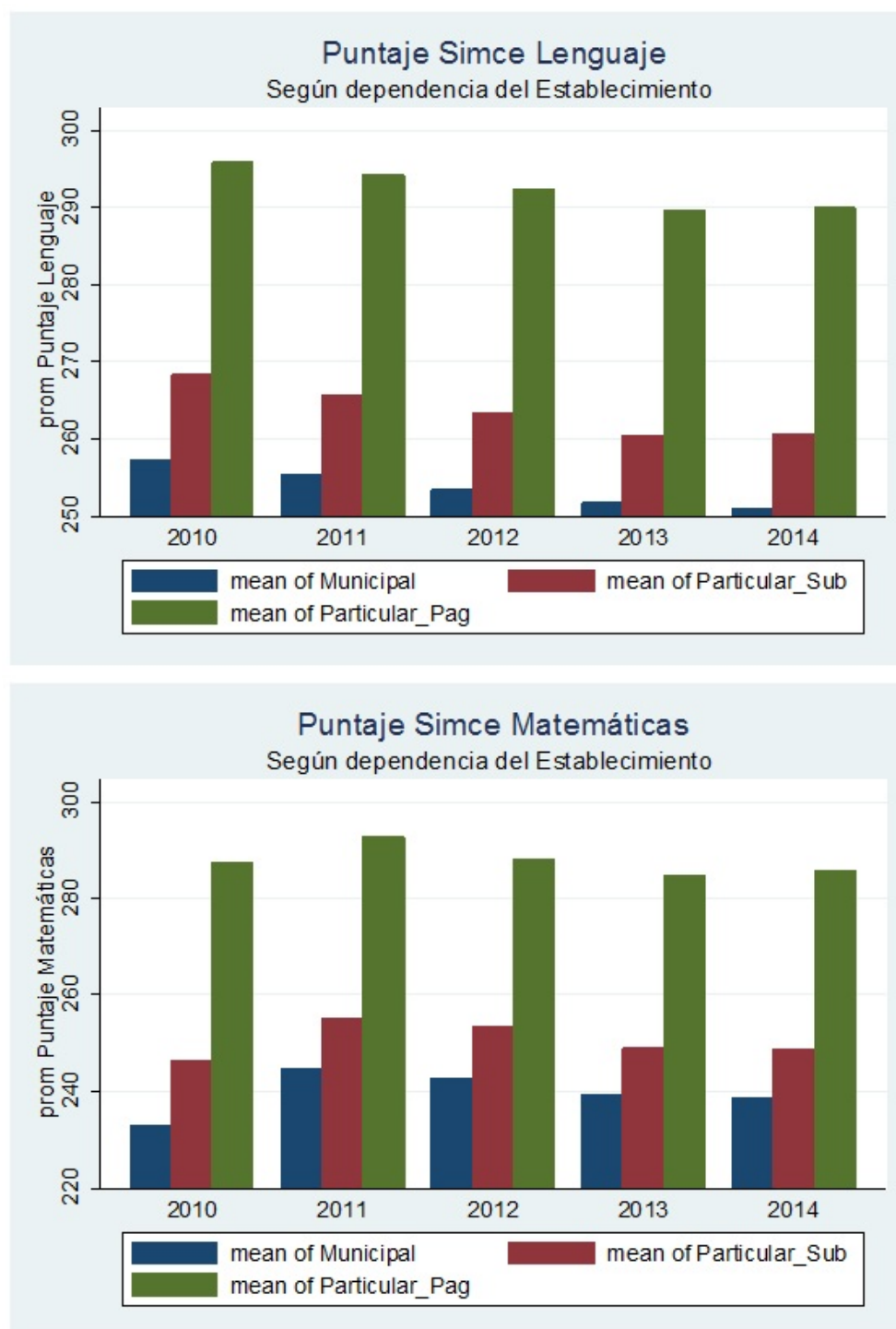


Figura 17

6.6. Anexo 2.2.2.

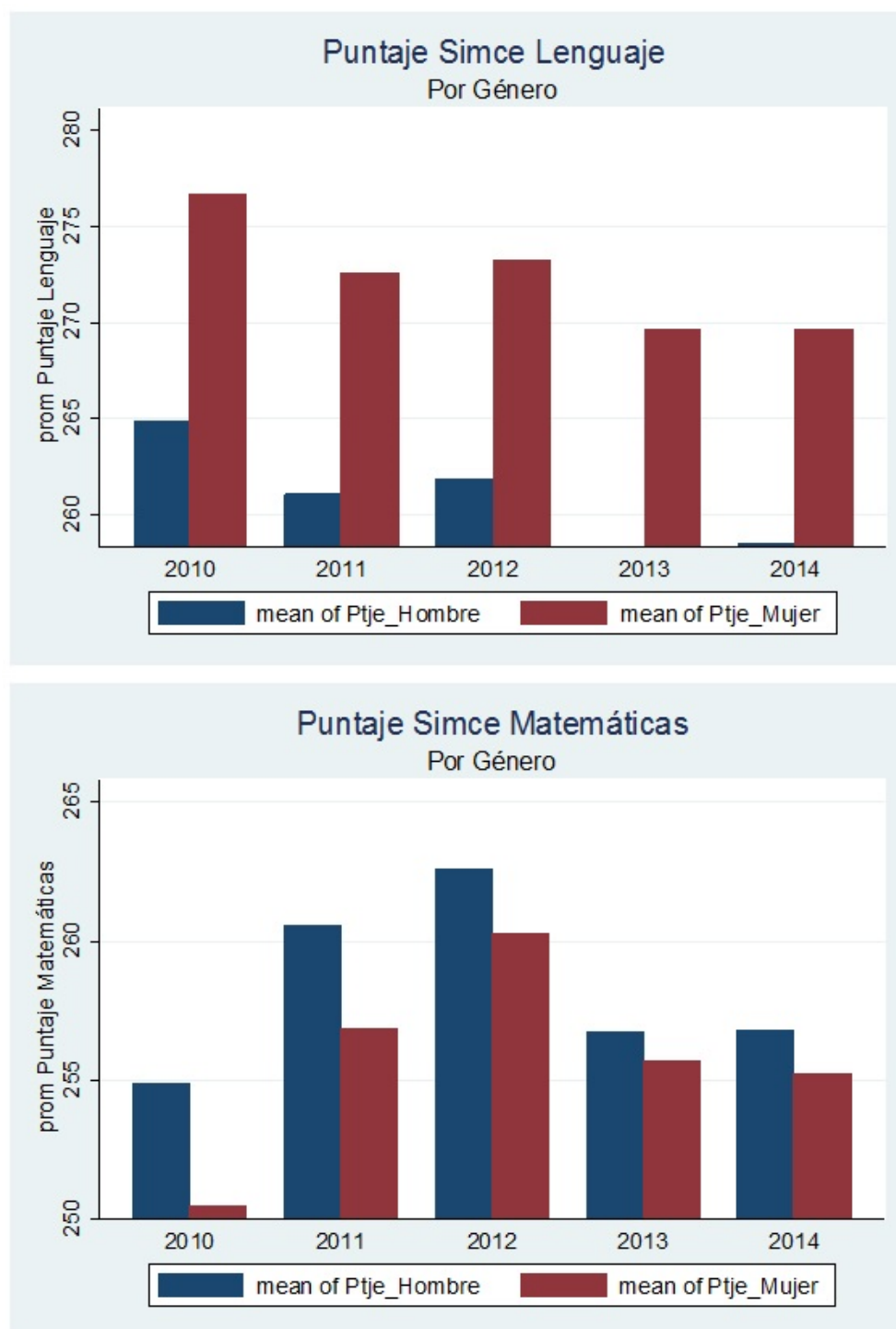


Figura 18

6.7. Anexo 2.2.3.

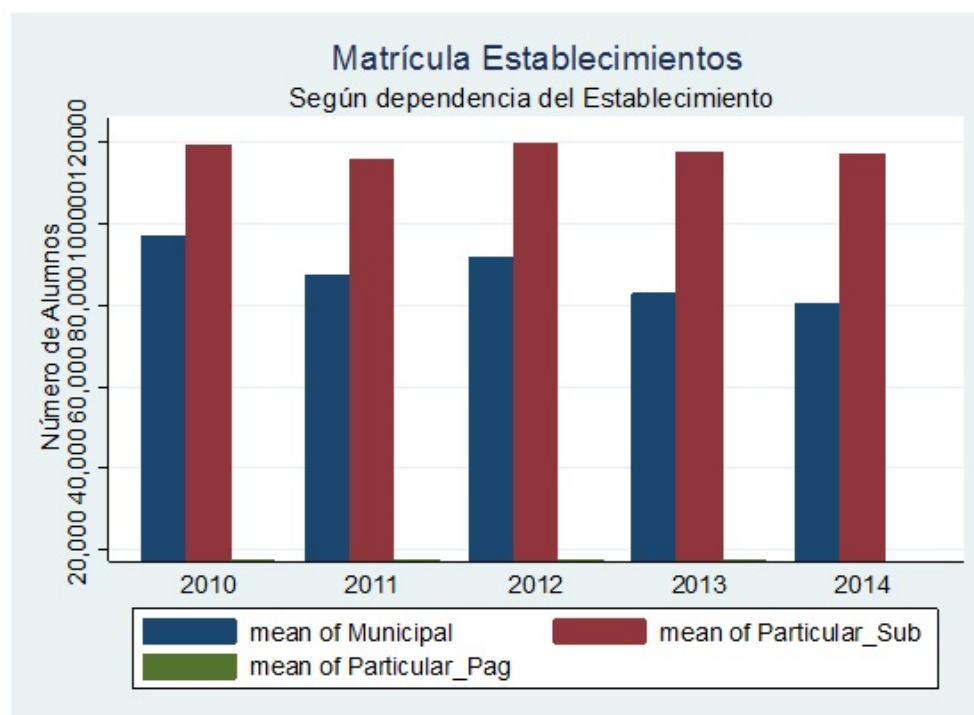


Figura 19

6.8. Anexo 2.2.4.

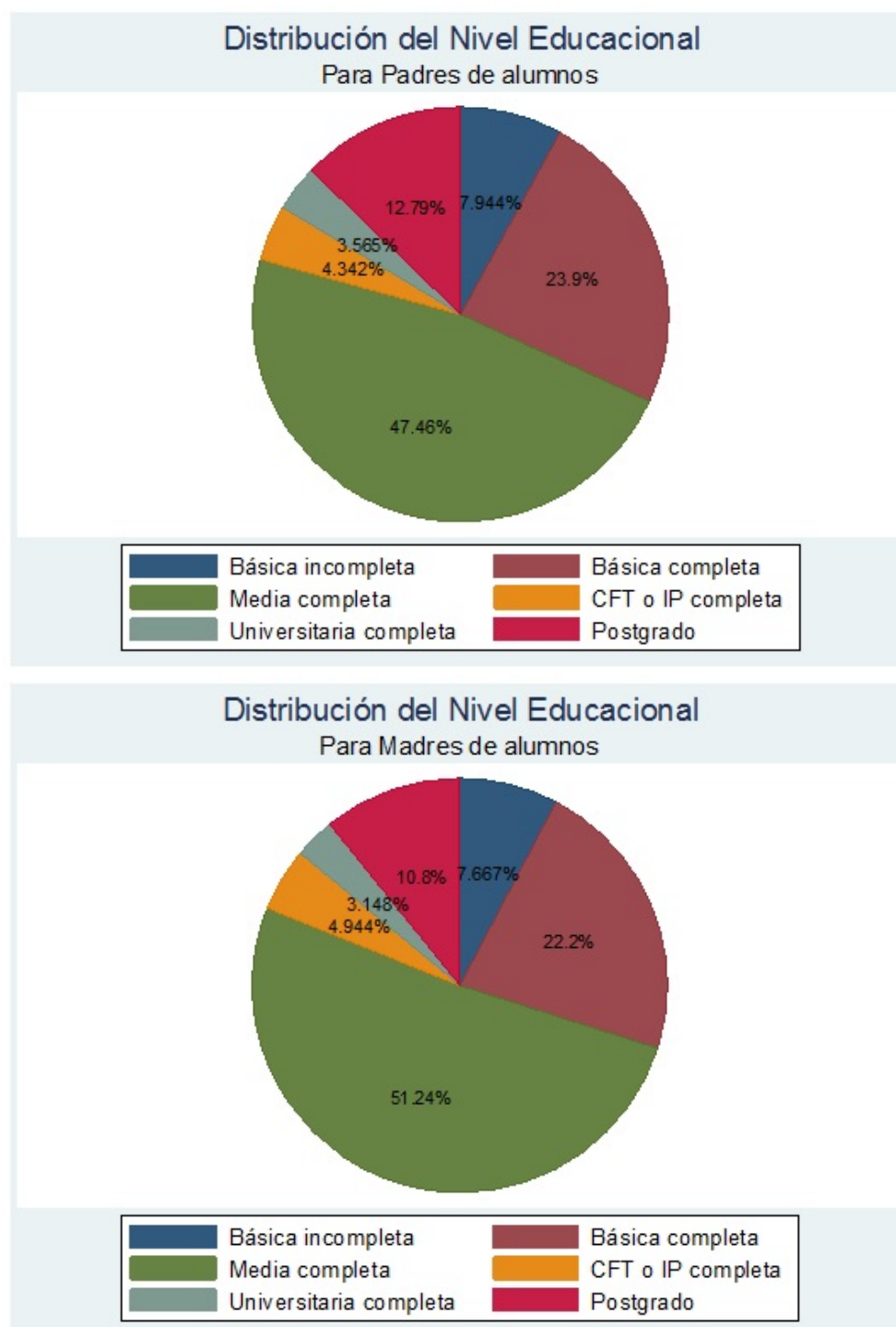


Figura 20

6.9. Anexo 2.2.5.

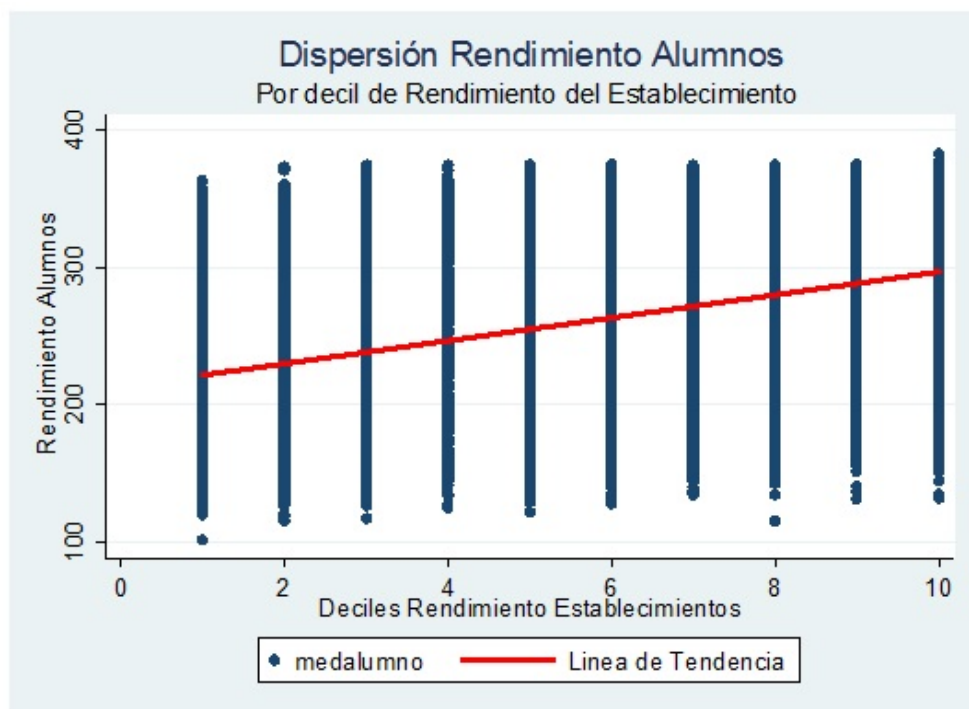


Figura 21

Anexo 1

Fixed-effects (within) regression	Number of obs	=	313839
Group variable: mrun	Number of groups	=	222096
R-sq: within = 0.0051	Obs per group: min	=	1
between = 0.0002	avg	=	1.4
overall = 0.0000	max	=	2
	F(4,91739)	=	116.44
corr(u_i, Xb) = -0.0669	Prob > F	=	0.0000

leng	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
gen_alu	4.225459	1.629462	2.59	0.010	1.031729	7.419188
lny	-1.760252	.2605364	-6.76	0.000	-2.270901	-1.249604
dependencia	.8183496	.2876532	2.84	0.004	.2545522	1.382147
bullying	-4.453744	.2337264	-19.06	0.000	-4.911846	-3.995643
_cons	275.9734	3.404631	81.06	0.000	269.3004	282.6464
sigma_u	50.463197					
sigma_e	29.193902					
rho	.7492411	(fraction of variance due to u_i)				

F test that all u_i=0: F(222095, 91739) = 3.54 Prob > F = 0.0000

Figura 22: Análisis por Efectos Fijos

Random-effects GLS regression	Number of obs	=	313839
Group variable: mrun	Number of groups	=	222096
R-sq: within = 0.0001	Obs per group: min	=	1
between = 0.1280	avg	=	1.4
overall = 0.1145	max	=	2
	Wald chi2(4)	=	27999.33
corr(u i, X) = 0 (assumed)	Prob > chi2	=	0.0000

leng	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
gen_alu	10.67358	.1989429	53.65	0.000	10.28366	11.0635
lny	11.65709	.1238205	94.15	0.000	11.41441	11.89978
dependencia	11.29331	.1653928	68.28	0.000	10.96915	11.61748
bullying	-6.263585	.1863759	-33.61	0.000	-6.628875	-5.898295
_cons	83.64799	1.472716	56.80	0.000	80.76152	86.53446
sigma_u	39.146476					
sigma_e	29.193902					
rho	.64260787	(fraction of variance due to u_i)				

```
. hausman fixed ., sig
```

	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	fixed	.	Difference	S.E.
gen_alu	4.225459	10.67358	-6.448124	1.629654
lny	-1.760252	11.65709	-13.41735	.2314637
dependencia	.8183496	11.29331	-10.47496	.2379968
bullying	-4.453744	-6.263585	1.809841	.1439417

b = consistent under H_0 and H_a ; obtained from xtreg
B = inconsistent under H_a , efficient under H_0 ; obtained from xtreg

Test: H_0 : difference in coefficients not systematic

```
chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B)
        = 5012.49
Prob>chi2 = 0.0000
```

Figura 24: Test de Hausman

Anexo 2

```
. reg leng gen_alu lny dependencia bullying if agno==2011
```

Source	SS	df	MS	Number of obs = 183438		
Model	52960968.3	4	13240242.1	F(4,183433) = 6072.61		
Residual	399943141183433	2180	32274	Prob > F = 0.0000		
				R-squared = 0.1169		
				Adj R-squared = 0.1169		
Total	452904110183437	2468	98995	Root MSE = 46.694		

leng	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
gen_alu	10.37788	.2182831	47.54	0.000	9.950054	10.80571
lny	11.87353	.1558884	76.17	0.000	11.56799	12.17907
dependencia	13.5616	.21513	63.04	0.000	13.13995	13.98325
bullying	-9.842007	.4823981	-20.40	0.000	-10.7875	-8.896518
_cons	79.1601	1.807159	43.80	0.000	75.61811	82.70209

Figura 25: Análisis de regresión

Anexo 3

```
. reg dleng dingfam dbullying, r
```

Linear regression

Number of obs = 487945
F(2,487942) =28686.00
Prob > F = 0.0000
R-squared = 0.1291
Root MSE = 126.08

dleng	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
dingfam	.0000759	4.41e-07	172.03	0.000	.000075	.0000768
dbullying	54.54923	.3475208	156.97	0.000	53.8681	55.23036
_cons	-30.80865	.202613	-152.06	0.000	-31.20577	-30.41154

Figura 26: Análisis de primeras diferencias

Anexo 4

```
. reg dleng dingfam dbullying lenguaje2011, r
```

Linear regression

Number of obs = 487945
F(3,487941) =76311.79
Prob > F = 0.0000
R-squared = 0.3589
Root MSE = 108.17

dleng	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
dingfam	.0000564	3.65e-07	154.62	0.000	.0000557	.0000571
dbullying	62.51787	.2716309	230.16	0.000	61.98549	63.05026
lenguaje2011	-.596304	.001464	-407.30	0.000	-.5991735	-.5934345
_cons	91.66122	.3560496	257.44	0.000	90.96337	92.35907

Figura 27: Análisis de primeras diferencias con control 2011

Anexo 5

```
. reg dleng dingfam dbullying lenguaje2011 gen_alu dependencia dalcdro dpromest, r
```

Linear regression

Number of obs = 487943
F(7,487935) = .
Prob > F = 0.0000
R-squared = 0.8974
Root MSE = 43.272

dleng	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
dingfam	9.23e-07	1.38e-07	6.67	0.000	6.52e-07	1.19e-06
dbullying	7.501356	.1510828	49.65	0.000	7.205238	7.797473
lenguaje2011	-.1254704	.0008238	-152.30	0.000	-.1270851	-.1238557
gen_alu	2.407597	.1252177	19.23	0.000	2.162175	2.65302
dependencia	1.8972	.1099089	17.26	0.000	1.681782	2.112618
dalcdro	5.663231	.2335851	24.24	0.000	5.205412	6.121051
dpromest	.8737804	.0009475	922.16	0.000	.8719233	.8756376
_cons	12.71318	.1953518	65.08	0.000	12.33029	13.09606

Figura 28: Análisis de primeras diferencias con controles