

Efectos de Largo Plazo de Familias en Acción sobre el Capital Humano



Javier E. Báez
Banco Mundial

Adriana Camacho
Univ. de los Andes

Motivación

- Los Programas de Transferencias Condicionadas (PTC) como Familias en Acción buscan:
 - reducir la pobreza,
 - proteger hogares vulnerables contra riesgos.
- En 1998 más de 30 países adoptaron PTC
- Revisión de literatura de 136 evaluaciones, solo 14% de ellas revisan resultados de LP.
- Esta es una evaluación de LP donde evaluamos:
 - Graduación del bachillerato y mejoras en calidad edu
 - A futuro revisaremos ingreso a educación superior, participación laboral
- Este programa ha beneficiado aproximadamente a $\frac{1}{4}$ de la población Colombiana.

Relevancia de Incrementar asistencia, participación y terminación

De los niños que entraron a 1ero,

- 60.7% llega a 5to
- 48.9% llega a 9no
- 34.8% llega a 11

(Sanchez et al, 2010 Res. 166)

Generalidades: Familias en Acción

- Fomenta inversiones en nutrición y educación infantil (Similar a Progresá)
- Programa tiene intenciones sobre **asistencia, no sobre resultados escolares.**

Tratamiento Educación básica y media:

- **Condiciona** en el 80% de asistencia escolar
- 8/15 dólares primaria/bachillerato
- Focaliza : niños 7-18 años de familias pertenecientes
 - Listados de nivel 1 del SISBEN,
 - condición de desplazamiento
 - indígenas.

Cobertura y evolución del programa

- ✓ Piloto 2001, 2002-3 Iniciaron 622 municipios con menos de 100,000 habitantes
- ✓ Expansión 2005, desplazados,
- ✓ Entrada de capitales 2007
- ✓ Para el 2009 existe expansión nacional (1,097 municipios y 2.5 millones de familias)

- **Participación:** 62% de familias elegibles en munis 2002
- **Costo** 0.27% del PIB en 2009.

Evidencia de programas similares en educación

- ↑ **Terminar Bachillerato** (Colombia-Bogotá y Pakistan)
 - 4 - 7 p.p. bachillerato (Alam et al. 2010)
 - 4 p.p. bachillerato y transición de 23 p.p a la educación superior (Barrera et al. 2008)
 - 5-7 p.p. bachillerato dado por PACES (Angrist et al. 2006)
- ↑ **Puntajes en exámenes**
 - preescolar y primaria (Argentina, Nicaragua y Turquía) (Heinrich 2007)
 - NO hay impactos puntaje en bachillerato (México y Camboya). (Berhman et al. 2005, Filmer y Shady 2009)

Evidencia sobre FeA

- ↑ **Matriculas** en 5-7 p.p. (niños de 14-17 años de edad), 2pp en primaria (Attanasio et al., 2005; IFS et al., 2006).
- ↓ **Repetición** en 0.1 años (IFS et al., 2006).
- **NO** es concluyente en **Pruebas SABER** (Garcia & Hill, 2009):
 - ↑ Puntajes alumnos de 5 ° zonas rurales).
 - ↓ Puntajes alumnos 9°
 - No existe evidencia previa en relación con los impactos sobre graduación de bachillerato o resultados del ICFES a nivel nacional.

Impacto de LP en calidad no es claro

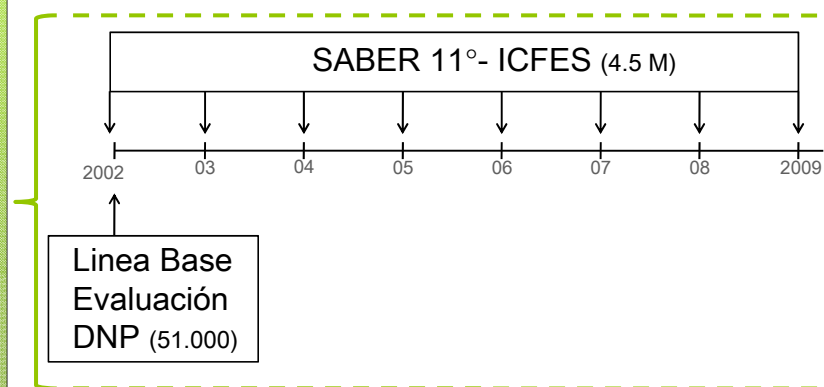
- ✓ Efecto ingreso
- ✓ Condicionalidad en asistencia
- ✗ Congestión dada por subsidios de demanda
- ✗ Selección negativa de los niños que ingresan por programa (motivación, menor nivel socio-económico)
- ✗ Pueden existir subsidios a la demanda pero deficiencias en la oferta que afecten el resultado final. (Benson, 2012)

Contribución

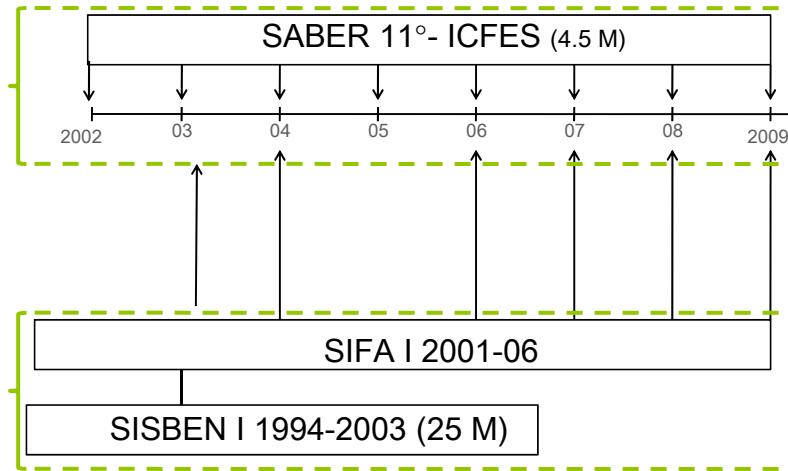
1. Evalúa la dinámica del programa dado que puede llegar a seguir niños con exposición de 9 años al programa.
2. Robustez de análisis con dos fuentes de información
 - Línea de base del panel de hogares utilizada para la evaluación de corto plazo -DNP.
 - Sistemas de Información Administrativos
 - SIFA + SISBEN
3. Evalúa heterogeneidad de efecto por genero y ubicación

Datos:

Encuestas Evaluación DNP + ICFES



Datos: Encuestas SIFA+ SISBEN+ICFES



Combinación de datos

- Se combinaron los datos de 4 formas:
 - Id
 - Apellidos y Id corto (primeros 7 dígitos id)
 - Apellidos y Fecha Nacimiento
 - Nombres y Apellidos (exacto o variaciones fonéticas)
- Corroboramos la combinación de datos, revisando la distancia fonética en los 4 casos.
- Unidad de análisis:** Niños nacidos entre 1975 - 1994 del tratamiento que se espera terminen bachillerato en el período 2003-2009.

Problemas de selección negativa en esta evaluación

- El programa **no** fue asignado aleatoriamente. Focalizado según:
 1. **Características municipio:** número de habitantes, banco
 2. **Socio-demográficas:** Niño que es llevado al colegio debido al programa puede ser más pobre
- Participantes pueden diferir de NO Participantes en características:
 - Observables (edad, género, etc.)
 - No observables (motivación, habilidad)

Diferencias entre Tratamiento y Control en la Línea Base

Variable	Treated		Control		Difference	t-stat	Significance	ND
	Mean	N	Mean	N				
Demographic								
Age (household head)	45.330	2,415	44.940	1,766	0.390	1.076		0.034
Age (spouse)	41.388	2,415	41.002	1,766	0.386	1.115		0.035
Age (child)	12.359	2,420	12.481	1,766	-0.123	-1.746	*	-0.055
Gender (household head)	0.772	2,420	0.840	1,766	-0.068	-5.576	***	-0.173
Gender (child)	0.436	2,420	0.428	1,766	0.008	0.517		0.016
Household structure								
Is household single headed?	0.021	2,420	0.020	1,766	0.002	0.376		0.012
Number of household members	6.072	2,420	6.165	1,766	-0.093	-1.327		-0.042
Number of children	1.337	2,420	1.419	1,766	-0.082	-2.484	**	-0.078
Consumption and assets								
Monthly household consumption	200,000	2,362	212,000	1,697	-12,600	-3.80	***	-0.12
Does the family own the house?	0.667	2,420	0.649	1,766	0.018	1.184		0.037
Education, health and work								
Does household head read?	0.826	2,276	0.831	1,673	-0.004	-0.362		-0.012
Household head completed secondary or more?	0.057	2,267	0.078	1,673	-0.02	-2.493	**	-0.081
Years of schooling (household head)	3.715	2,129	4.037	1,576	-0.322	-2.969	***	-0.099
Did children suffer from diarrhea?	0.112	1,086	0.112	702	0	0.013		0.001
Does household head work?	0.887	2,316	0.885	1,720	0.002	0.182		0.006

Diferencias entre Tratamiento y Control en la Línea Base

Variable	Treated		Control		Difference	t-stat	Significance	ND
	Mean	N	Mean	N				
Dwelling characteristics								
Located in an urban area?	0.562	2,420	0.695	1,766	-0.133	-8.943	***	-0.278
No walls?	0.007	2,414	0.011	1,765	-0.004	-1.41		-0.045
Connected to piped water?	0.685	2,409	0.762	1,766	-0.078	-5.599	***	-0.174
Connected to gas?	0.12	2,390	0.14	1,759	-0.021	-1.955	*	-0.062
Connected to sewage?	0.334	2,417	0.324	1,766	0.01	0.708		0.022
Community								
Altitude	658.161	2,415	567.069	1,766	91.092	3.846	***	0.121
Students per teacher	22.485	2,415	22.678	1,766	-0.192	-1.139		-0.036
Square metres of classroom per student	2.949	2,415	2.502	1,766	0.448	6.969	***	0.215
Number of banks	1.694	2,369	0.909	1,766	0.784	14.215	***	0.446
Number of health centers	1.134	2,369	0.844	1,766	0.29	6.875	***	0.215
Region = East?	0.214	2,415	0.245	1,766	-0.03	-2.281	**	-0.072
Region = Central?	0.289	2,415	0.162	1,766	0.127	9.953	***	0.307
Region = Pacific?	0.114	2,415	0.127	1,766	-0.013	-1.227		-0.039
Affected by violent attacks?	0.025	2,415	0.031	1,766	-0.006	-1.107		-0.035

Estrategia Empírica: PSM

- Propensity-Score Matching (PSM)
- Estiman 3 tipos de modelos de propensión a participación en el programa:
 1. Variables individuales y de hogar
 2. Variables individuales y de hogar + variables de educación municipales
 3. Variables “Attanasio et al (2005)” excepto variables municipales de salud.
- Se restringe muestra al soporte común y se genera emparejamiento.
- PSM elimina problemas de focalización y selección en la entrada al programa.

Variables de Resultados

- **Finalización de bachillerato:** presentación SABER 11. Proxy con correlación 90% (Banco Mundial, 1993).
 - La persona cumple con las siguientes condiciones
 - es menor a 18 en momento de inscripción en FA,
 - esta debajo de grado 11,
 - le faltan un número de años escolares con los cuales logra llegar a 11 máximo en 2009,
- **Puntajes de las Pruebas:**
 - Puntaje total (sin idioma extranjero) y por áreas
 - Matemáticas, Español
 - Puntaje estandarizado y ponderado

Estadísticas descriptivas

Variable	Obs	Mean	Std. Dev.	Min	Max
Icfes	4,158	0.523	0.500	0.000	1.000
Icfes old	2,731	0.127	0.334	0.000	1.000
Overall score	2,523	-0.569	0.613	-5.092	1.920
Mathematics	2,523	-0.229	0.892	-5.322	4.197
Spanish	2,523	-0.278	0.870	-5.423	2.574

Graduación					
	OLS		Matching		
	(1)	(2)	(1)	(2)	(3)
All sample	0.030* [0.017]	0.050*** [0.018]	0.0401** [0.0187]	0.0840** [0.0220]	0.0696 ** [0.0214]
Observations	3,452	3,452	3,476	3,476	3,888
Boys	0.011 [0.026]	0.036 [0.027]	0.0206 [0.0301]	0.0661* [0.0363]	0.0587* [0.0348]
Observations	1,478	1,478	1,490	1,490	1,687
Girls	0.046** [0.022]	0.065*** [0.023]	0.0523** [0.0245]	0.0856*** [0.0290]	0.0899*** [0.0290]
Observations	1,974	1,974	1,986	1,986	2,198
Urban	-0.008 [0.021]	0.015 [0.022]	-0.0052 [0.0229]	0.0492 [0.0274]	0.0390 [0.0254]
Observations	2,102	2,102	2,120	2,120	2,372
Rural	0.091*** [0.028]	0.115*** [0.029]	0.0868*** [0.0314]	0.1044*** [0.0402]	0.1176*** [0.0398]
Observations	1,350	1,350	1,356	1,356	1,514

Errores estándar calculados usando *Bootstrap* con 200 replicaciones.

Variables de Resultados

- **Finalización de bachillerato:** presentación SABER 11. Proxy con correlación 90% (Banco Mundial, 1993).
 - La persona cumple con las siguientes condiciones
 - es menor a 18 en momento de inscripción en FA,
 - esta debajo de grado 11,
 - le faltan un número de años escolares con los cuales logra llegar a 11 máximo en 2009,
- **Puntajes de las Pruebas:**
 - Puntaje total (sin idioma extranjero) y por áreas
 - Matemáticas, Español
 - Puntaje estandarizado y ponderado

Problema de selección adicional...

- Se observan puntajes para niños que llegan a grado 11, pero el programa tiene efecto sobre la asistencia escolar.
- Estarán sobre-representados los tratados, sesgo de selección negativo, entonces la distribución de puntajes para tratamiento y control no son comparables.
 - Algunos estudiantes tomarían siempre el examen
 - Otros lo toman dado que el programa los llevo hasta 11 grado.
- Los resultados de LP tendrán problemas adicionales de selección negativa que no captura PSM.

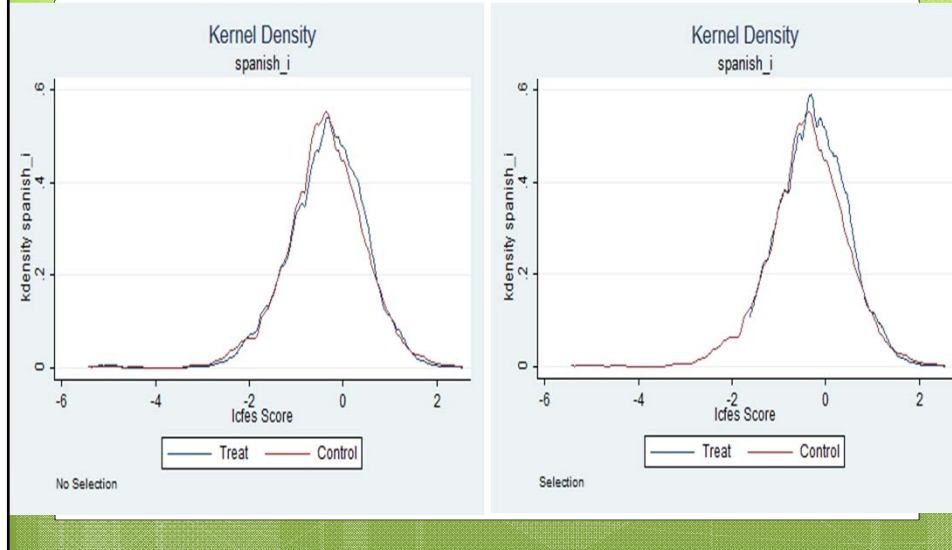
Alternativa: construir límites no paramétricos

- Lee, 2002, Angrist et al. 2004
- Truncamiento simétrico (mismo percentil) para las distribuciones tratamiento y control .
- ¿Cómo?** Seleccionar un percentil θ tal que $q_0(\theta) = 0$ y después eliminar percentiles menores a θ_0 en la distribución Y_1 (encontramos limite superior)

$$E(Y_i|D_i = 1, Y > q_0(\theta)) - E(Y_i|D = 0, Y > q_0(\theta))$$

- Comparaciones contaminadas por selección arrojan un efecto menor que el real de FA sobre los puntajes.

Kernel de puntajes sin y con corrección por selección



Puntajes SABER 11- SIN corrección

Dependent Variable	OLS		Matching estimates		
	(1)	(2)	(1)	(2)	(3)
Mathematics	-0.401	-0.593	-0.0189	-0.0141	-0.0439
	[0.392]	[0.413]	[0.0474]	[0.0545]	[0.0558]
Observations	1,867	1,867	1,867	1,867	2,044
Spanish	0.398	0.079	0.0502	-0.0113	0.0171
	[0.350]	[0.368]	[0.0465]	[0.0902]	[0.0579]
Observations	1,867	1,867	1,867	1,867	2,044
Overall Testscores	0.179	-0.086	0.0292	0.0040	0.0171
	[0.226]	[0.237]	[0.0327]	[0.0340]	[0.0361]
Observations	1,867	1,867	1,867	1,867	2,044

Nota: Errores estándar calculados usando *Bootstrap* con 200 replicaciones.

Puntajes SABER 11- CON corrección

Outcome: Test Score	(1)		(2)		(3)	
	Control ϕ	Upper Bound	Control ϕ	Upper Bound	Control ϕ	Upper Bound
Mathematics	48	0.033 [0.0419]	50	0.0374 [0.0495]	60	0.0179 [0.0464]
Observations		1,850		1,850		2,023
Spanish	48	0.0982** [0.0438]	50	0.0421 [0.0525]	60	0.0689 [0.0567]
Observations		1,847		1,847		2,021
Overall test score	48	0.0698** [0.0328]	60	0.0413 [0.0419]	60	0.0498 [0.0358]
Observations		1,850		1,850		2,022

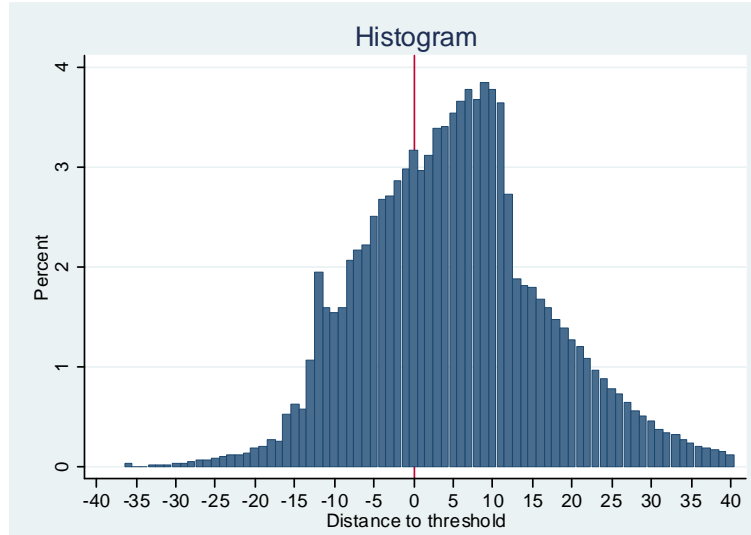
Nota: Errores estándar calculados usando *Bootstrap* con 200 replicaciones.

Estrategia Empírica: RDD

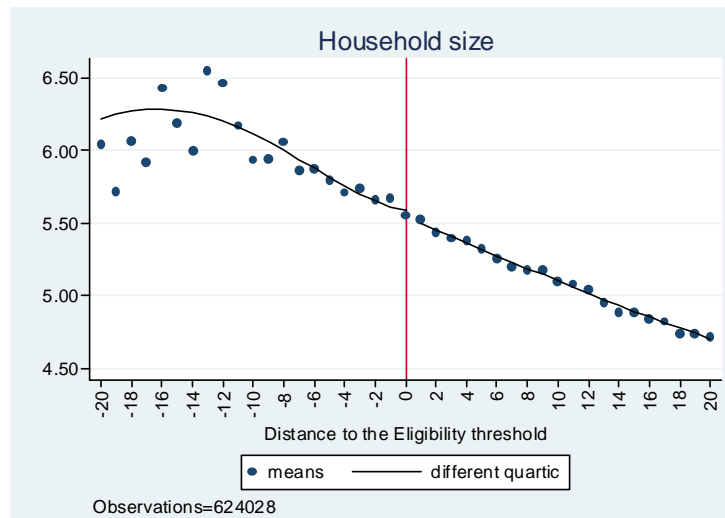
Regresión Discontinua (RD)

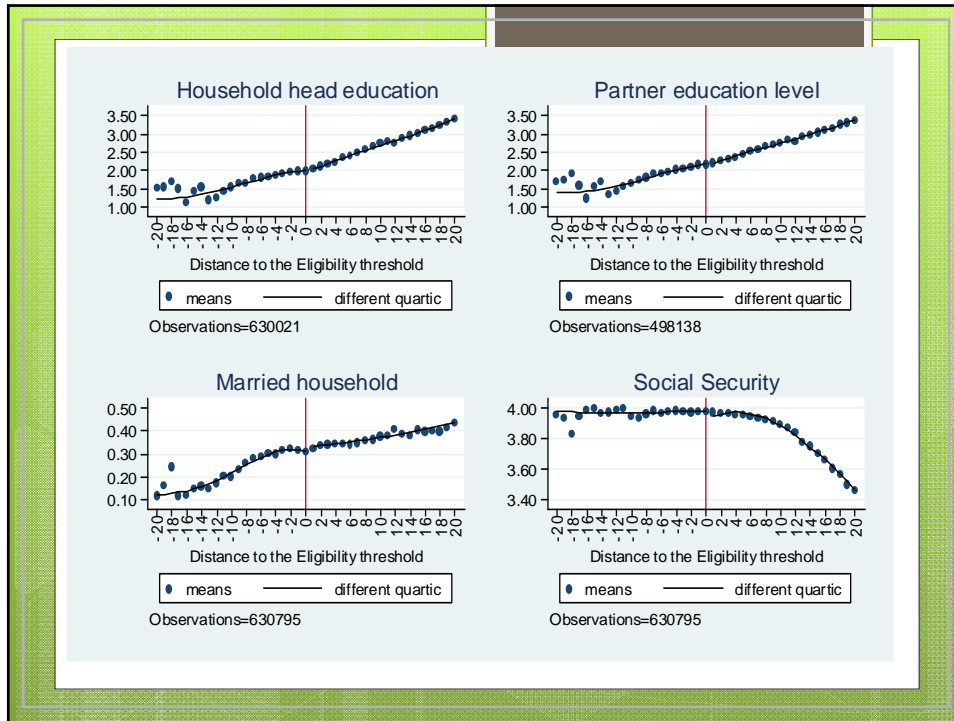
- Explora la discontinuidad del tratamiento de subsidios en el criterio de elegibilidad
 - Nivel 1 de SISBEN: Áreas urbanas/rurales puntaje 36/18
- Focalización SISBEN "ideal" para aplicar RDD, se deben cumplir :
 1. Asignación exógena justo **alrededor del punto de corte**
 - Revisar la distribución del puntaje
 - Otras variables, diferentes a las del programa, no deben saltar alrededor del punto de corte.
- Así nos aseguramos de comparar personas muy similares alrededor del punto de corte...

Distribución puntaje Sisben



...No existe salto en otras variables...





Estrategia Empírica : RDD

- Conocemos la regla de selección para la asignación del tratamiento.

$$elig_i = 1(\text{puntaje}_i < \overline{\text{puntaje}})$$

→ El indicador de elegibilidad (punto de corte) debe ser un fuerte determinante de afiliación a Familias en Acción.

... Solo de esta forma podemos asumir que el cambio en **resultados escolares** se debe a un **efecto del tratamiento**

- Efectos reportados son locales y por lo tanto no pueden ser generalizados para toda la población

- Dado que el tratamiento no se asignó para todos los elegibles... instrumentamos el tratamiento usando la probabilidad de recibir tratamiento

Primera Etapa

$$FA_i = \delta_0 + \delta_{elig_i} + f_1(punt_i | punt_i \leq \overline{punt}) + f_2(punt_i | punt_i > \overline{punt}) + \gamma_t + \gamma_m + v_i$$

- Forma funcional: cúbica, cuadrática, no paramétrica
- Esperamos: $\delta \neq 0$

Forma Reducida

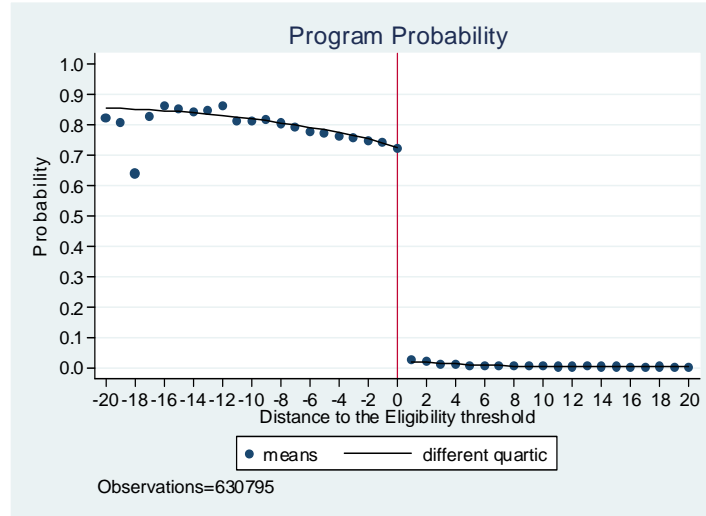
$$r_esc_i = \pi_0 + \pi_{eleg_i} + f_1(punt_i | punt_i \leq \overline{punt}) + f_2(punt_i | punt_i > \overline{punt}) + \gamma_t + \gamma_m + \eta_i$$

- El coeficiente de interés es: $\theta = \frac{\pi_1}{\delta_1}$
- El resultado escolar lo ajustamos por la probabilidad de tratamiento (FA)

Estadísticas descriptivas

Variable	Obs	Mean	Std. Dev.	Min	Max
Eligibility	624,028	0.347	0.476	0.000	1.000
FEA	624,028	0.274	0.446	0.000	1.000
Icfes	624,028	0.243	0.429	0.000	1.000
Overall score	131,744	-0.410	0.650	-6.613	3.241
Mathematics	131,744	-0.090	0.897	-7.245	6.866
Spanish	131,744	-0.087	0.889	-7.257	5.613

Primera etapa ...Salto en el punto de Corte...

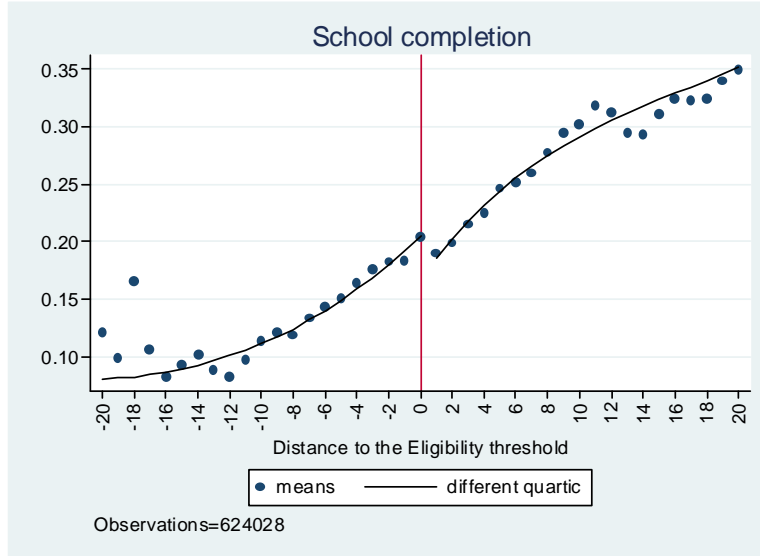


Primera etapa ...Salto en el punto de Corte...

First Stage Results

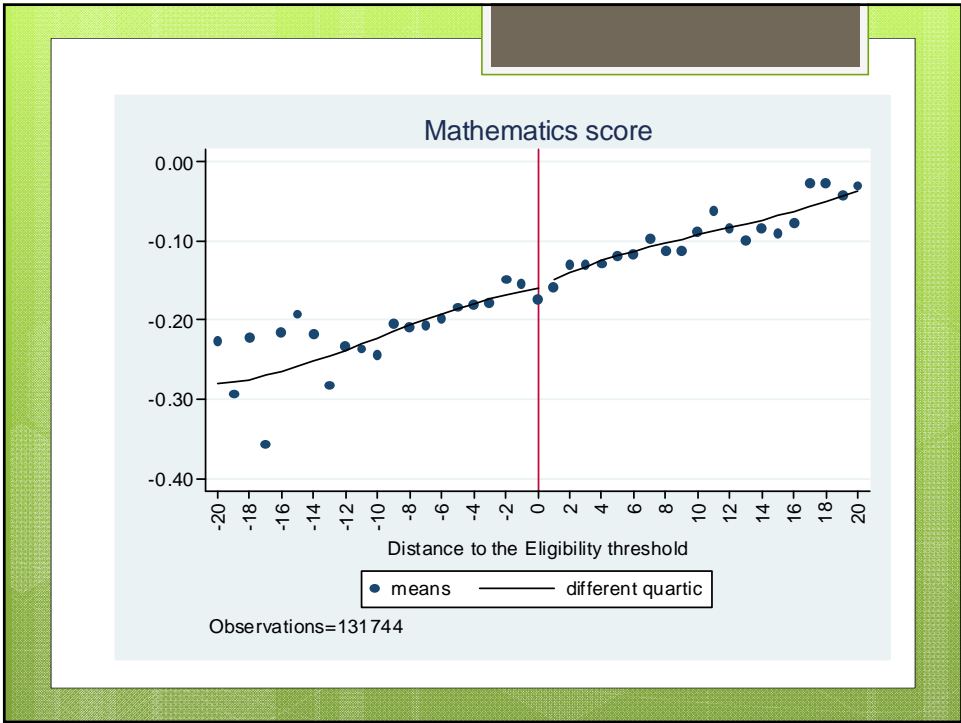
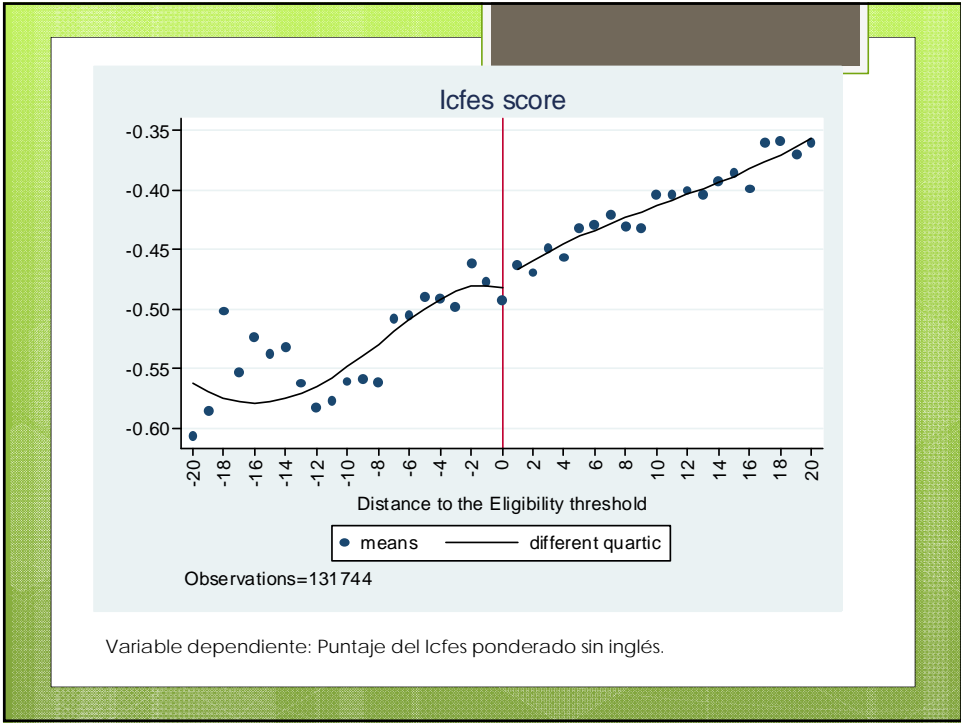
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Same functional form above and below threshold							
Eligibility	0.732*** [0.002]	0.726*** [0.001]	0.731*** [0.001]	0.726*** [0.001]	0.705*** [0.002]	0.714*** [0.002]	
Observations	624,028	624,028	624,028	624,028	624,028	624,028	
R2	0.688	0.741	0.688	0.741	0.689	0.742	
Panel B: Different functional form above and below threshold							
Eligibility	0.702*** [0.002]	0.711*** [0.002]	0.706*** [0.002]	0.708*** [0.002]	0.696*** [0.003]	0.702*** [0.002]	0.703*** [0.008]
Observations	624,028	624,028	624,028	624,028	624,028	624,028	14,647
R2	0.689	0.742	0.689	0.742	0.690	0.742	
Quadratic	Yes	Yes					
Cubic			Yes	Yes			
Quartic					Yes	Yes	
Municipality fixed effects		Yes		Yes		Yes	
Imbens Optimal Bandwidth							0.7573

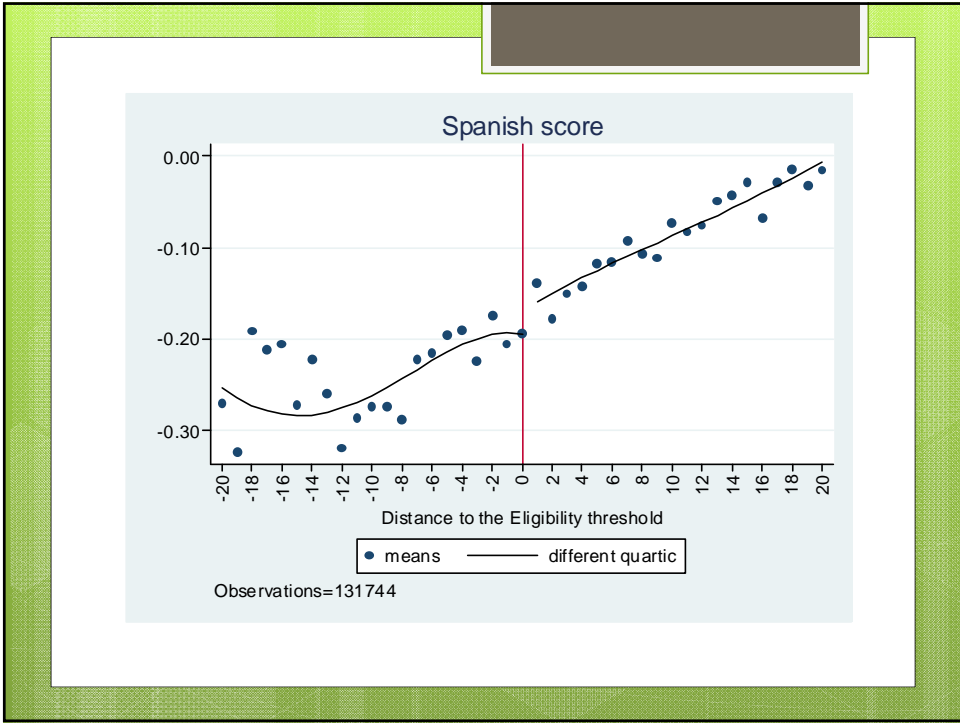
Graduación



Graduación

	All sample		Boys		Girls		Urban		Rural	
FA	0.036***	0.024**	0.033***	0.023	0.045***	0.018	0.044***	0.042**	0.038***	0.021*
	[0.006]	[0.011]	[0.006]	[0.015]	[0.007]	[0.017]	[0.009]	[0.019]	[0.006]	[0.013]
Observations	208,366	25,249	308,345	11,374	315,544	11,191	257,689	10,202	359,952	16,078
bandwidth		1.310		1.204		1.143		1.414		1.342
Municipality fixed effects	Yes		Yes		Yes		Yes		Yes	
Functional form	Cuartic		Cuartic		Cuartic		Cuartic		Cuartic	





Outcome: Test scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mathematics	-0.024	-0.022	-0.034	-0.034	-0.002	-0.015	-0.048***
	[0.018]	[0.018]	[0.023]	[0.022]	[0.027]	[0.027]	[0.015]
Observations	131,744	131,744	131,744	131,744	131,744	131,744	11,689
Imbens Optimal Bandwidth							0.60
Spanish	-0.044**	-0.039**	-0.037	-0.037	-0.039	-0.048*	-0.100***
	[0.018]	[0.018]	[0.023]	[0.023]	[0.028]	[0.027]	[0.010]
Observations	131,744	131,744	131,744	131,744	131,744	131,744	24,724
Imbens Optimal Bandwidth							1.283
Overall test score	-0.020	-0.020	-0.025	-0.028*	-0.009	-0.025	-0.057***
	[0.013]	[0.013]	[0.016]	[0.016]	[0.020]	[0.019]	[0.009]
Observations	131,744	131,744	131,744	131,744	131,744	131,744	17,031
Imbens Optimal Bandwidth							0.886
Quadratic	Yes	Yes					
Cubic			Yes	Yes			
Quartic					Yes	Yes	
School fixed effects		Yes		Yes		Yes	

Effect of program participation	Outcome: Test scores					
	Mathematics		Spanish		Overall test score	
Boys	0.025	0.036**	-0.057	-0.111***	-0.015	-0.064***
	[0.041]	[0.015]	[0.041]	[0.012]	[0.030]	[0.010]
Observations	57,637	9,319	57,637	12,828	57,637	11,619
Imbens Optimal Bandwidth		0.990		1.355		1.230
Girls	-0.045	-0.132***	-0.034	-0.073***	-0.028	-0.068***
	[0.035]	[0.021]	[0.036]	[0.019]	[0.025]	[0.011]
Observations	74,087	6,992	74,087	8,119	74,087	12,881
Imbens Optimal Bandwidth		0.708		0.827		1.315
Urban	-0.011	-0.063***	0.024	0.018	0.018	-0.017
	[0.042]	[0.019]	[0.041]	[0.025]	[0.030]	[0.014]
Observations	64,036	9,076	64,036	4,443	64,036	8,413
Imbens Optimal Bandwidth		1.264		0.639		1.176
Rural	-0.064*	-0.017	-0.102***	-0.154***	-0.084***	-0.065***
	[0.039]	[0.012]	[0.039]	[0.012]	[0.028]	[0.011]
Observations	65,523	8,925	65,523	14,116	65,523	9,223
Imbens Optimal Bandwidth		0.725		1.220		0.752
Functional form	Quartic		Quartic		Quartic	
Year fixed effects	Yes		Yes		Yes	
School fixed effects	Yes		Yes		Yes	

Resultado y Discusión

- **Finalización de bachillerato**
 - Un incremento de 4-8 p.p. (9-20% desv. estándar)
 - Jalonado por efecto niñas y áreas rurales.
- **Puntajes en las Pruebas**
 - En PSM no se encuentra impacto si no se corrige por el problema de selección, luego aparece positivo en algunas áreas
 - En RD el Impacto parece ser negativo
- Importancia de comparar este tipo de programas con programas vouchers (Angrist et al. 2006), subsidios oferta, condicionalidades del programa.