

# **Heterogeneidad en el Desempeño Académico de los Estudianties de Argentina: Evidencia a partir de Regresión por Cuantiles**

Gertel, Héctor; Giuliadori, Roberto; Vera, María  
Luz; Bastos, Guadalupe y Costanzo Sonia

2010

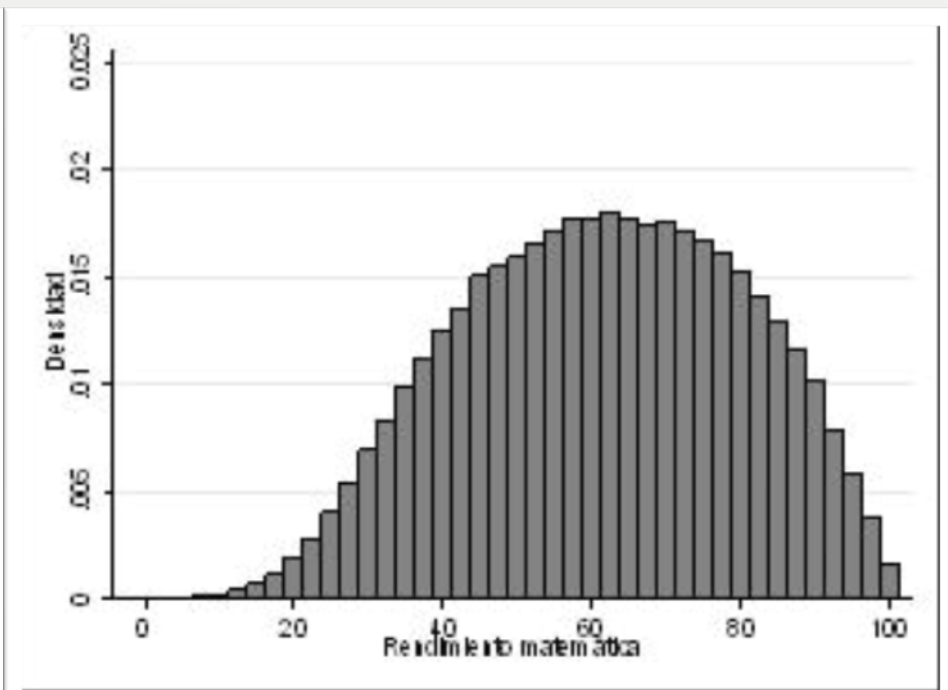
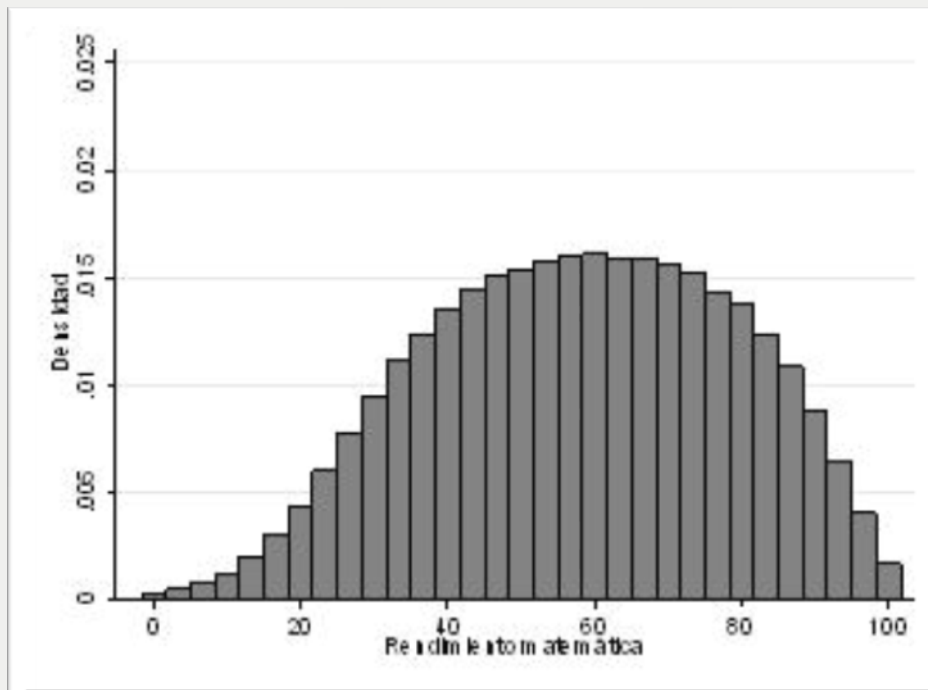
# 1. Introducción

- ▣ Hipótesis: la incidencia de variables de tipo personal, familiar y los efectos fijos no es uniforme a lo largo de la distribución del rendimiento escolar de los alumnos.
- ▣ Metodología: Se emplean regresiones por cuantiles (Koenker y Bassett, 1978) para captar la heterogeneidad de los datos.
- ▣ Datos: la información surge del Operativo Nacional de Evaluación del año 2000 (ONE 2000) tanto a nivel primario como secundario, al último año de cada ciclo. La muestra de trabajo es de 136.587 casos a nivel primario y 84.964 a nivel secundario.

## 2. La heterogeneidad en los resultados de matemática

Nivel Primario  
(total de estudiantes)

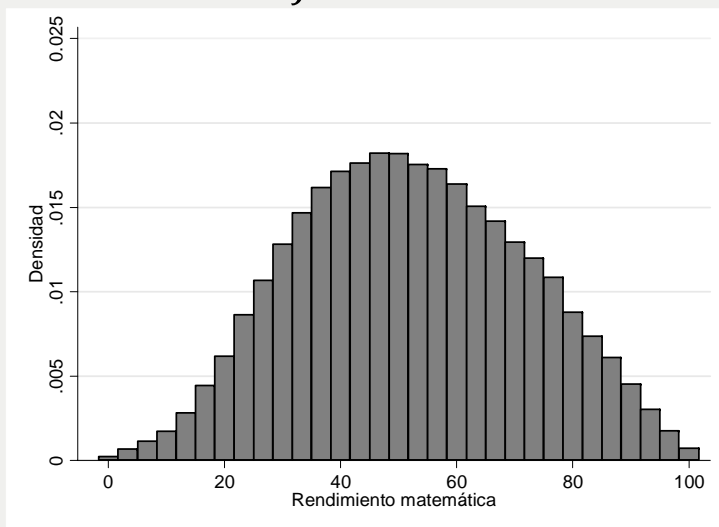
Nivel Secundario  
(total de estudiantes)



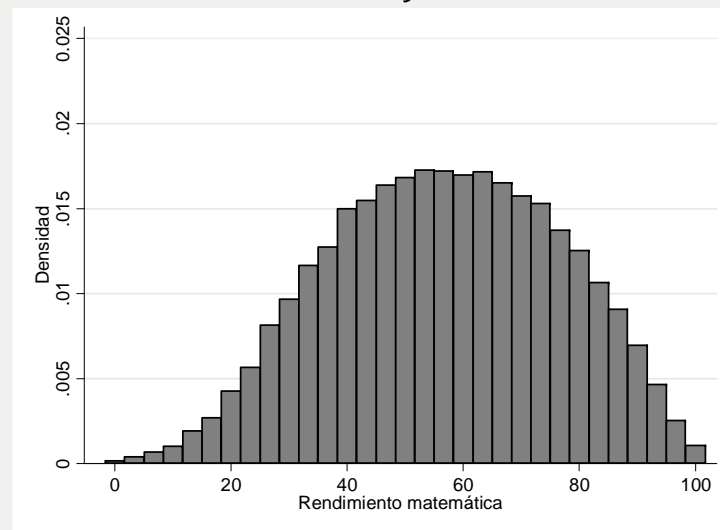
*Fuente: elaboración propia en base a datos del ONE 2000.*

# a) Nivel Primario

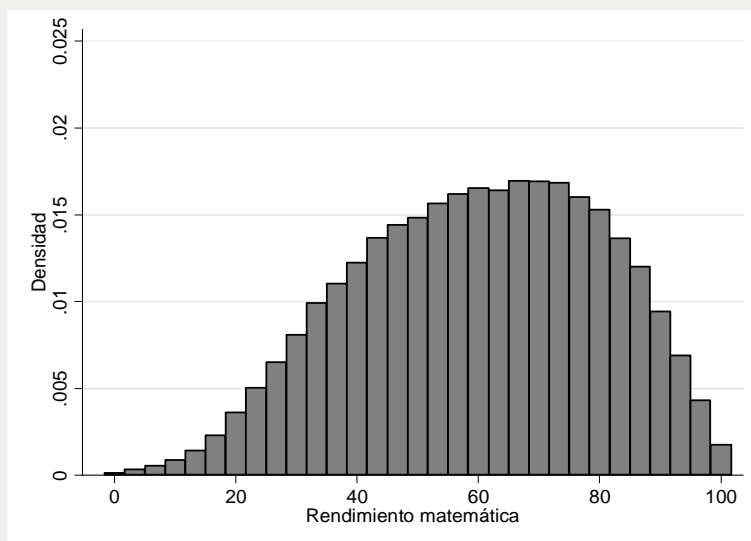
## Bajo NSE



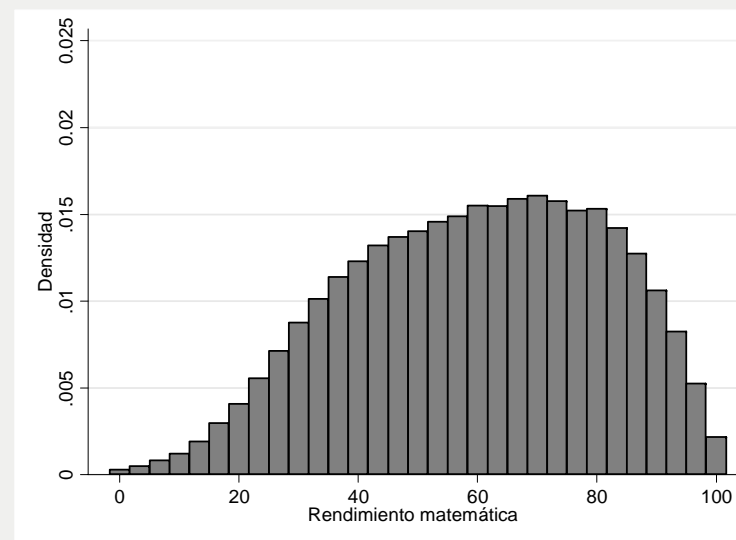
## Medio Bajo NSE



## Medio Alto NSE



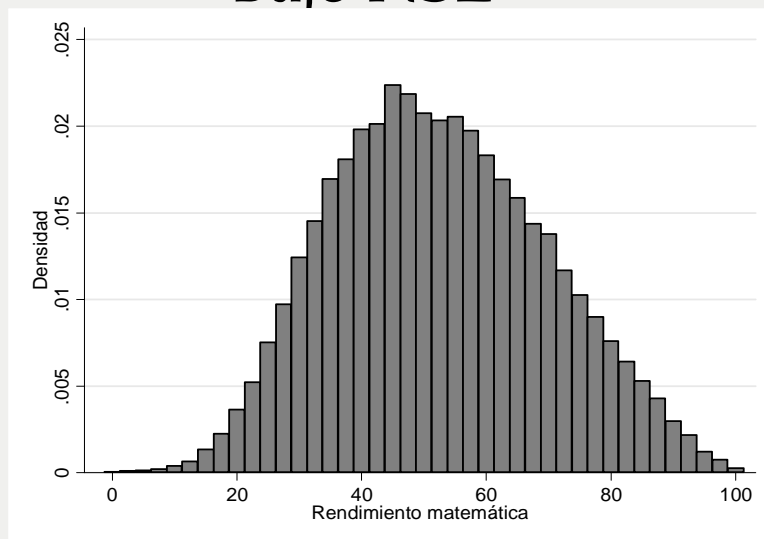
## Alto NSE



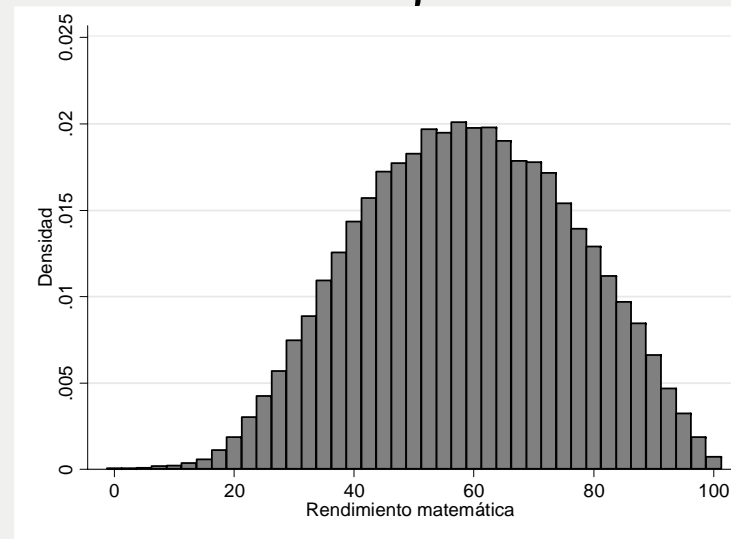
*Fuente: elaboración propia en base a datos del ONE 2000.*

# b) Nivel Secundario

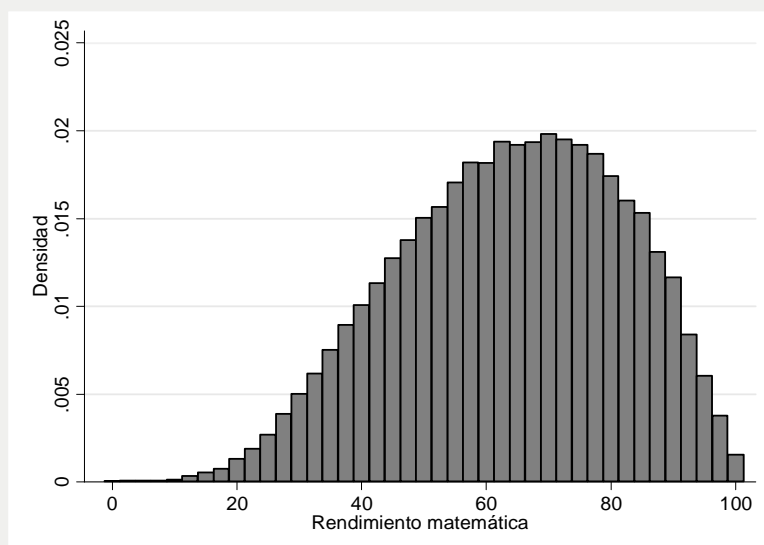
## Bajo NSE



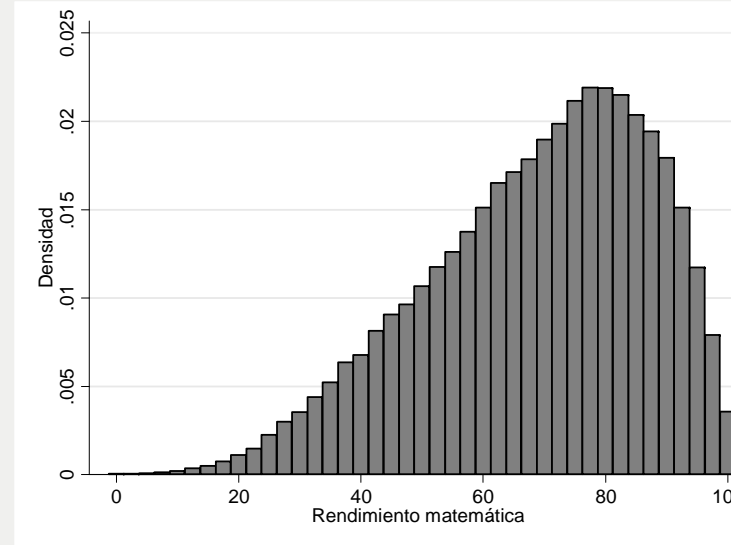
## Medio Bajo NSE



## Medio Alto NSE



## Alto NSE



*Fuente: elaboración propia en base a datos del ONE 2000.*

# Pruebas formales de heterocedasticidad

## Test de Breusch - Pagan

---

---

Test de Heteroscedasticidad de Breusch-Pagan / Cook-Weisberg

Ho: la varianza de  $\ln y_{it}$  es constante

---

Nivel Primario	Nivel Secundario
Chi2(1)= 2432.24	Chi2(1)= 113.30
Prob > chi2 = 0.00	Prob > chi2 = 0.00

---

---

## Test de White

---

---

Test de White

Ho: homocedasticidad contra H1: heteroscedasticidad inestruita

---

Nivel Primario	Nivel Secundario
Chi2(1)= 2432.24	Chi2(1)= 113.30
Prob > chi2 = 0.00	Prob > chi2 = 0.00

---

---

*Fuente: elaboración propia en base a datos del ONE 2000.*

### 3. Aspectos metodológicos

## La regresión por MCO

Para un proceso de regresión del tipo  $y = x\hat{\beta} + u$

El problema es  $\min_{\hat{\beta} \in \square} \left[ (y - x\hat{\beta})^2 \right]$

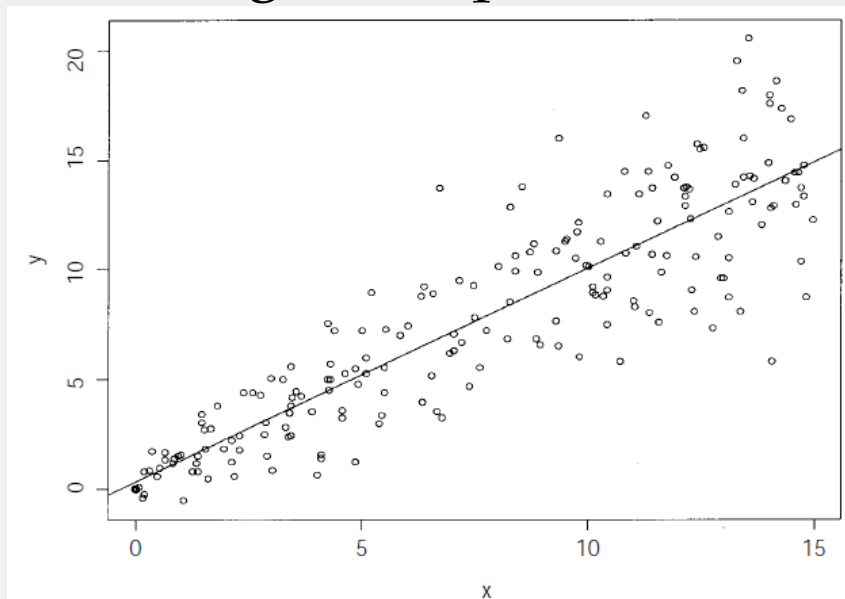
Suponiendo  $E[u^2] = \sigma^2 \quad \forall i = 1, \dots, n$

$$E[u / x] = 0$$

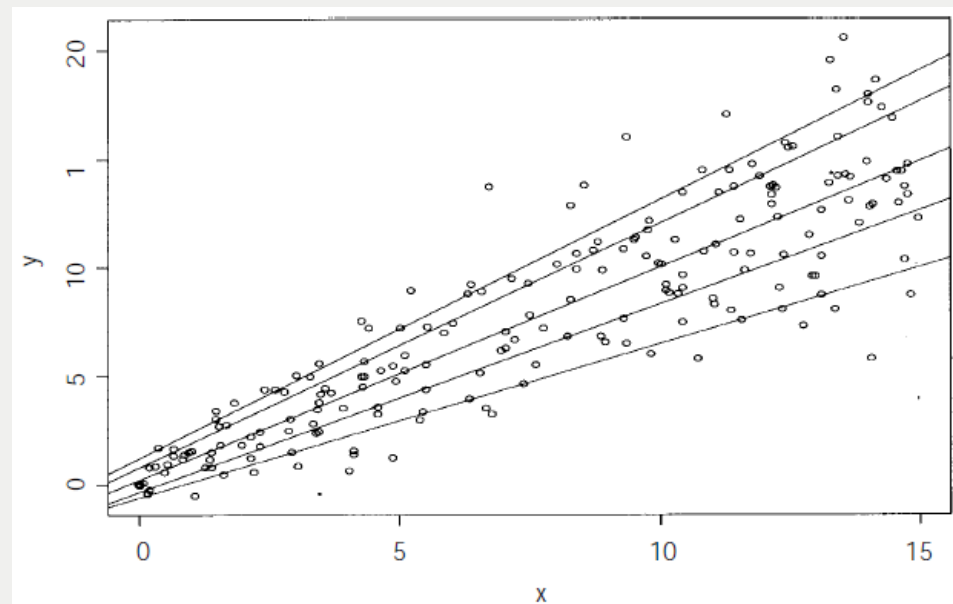
De lo cual surge  $E[y / x] = x' \beta$

$$\frac{\partial E[y / x]}{\partial x} = \beta$$

# Regresión por MCO



# Regresiones por Cuantiles





# Regresiones por cuantiles

Sea el proceso regresivo  $y_i = x_i \hat{\beta} + u_i$

El problema es (Bassett y Koenker, 1978)

$$\min_{\hat{\beta} \in \mathbb{R}^k} \left[ \sum_{i \in \{i: y_i \geq x_i \hat{\beta}_\Phi\}} \Phi |y_i - x_i \hat{\beta}_\Phi| + \sum_{i \in \{i: y_i < x_i \hat{\beta}_\Phi\}} (1 - \Phi) |y_i - x_i \hat{\beta}_\Phi| \right]; \quad 0 < \Phi < 1$$

Se obtienen estimaciones condicionadas de  $\beta$  para distintos  $\Phi$ .

Aplicación: regresión para el  
j-ésimo cuantil

$$\text{renmat}_{ij} = \beta_{0j} + \beta_{1j}(\text{gen})_{ij} + \beta_{2j}(\text{act})_{ij} + \beta_{3j}(\text{rep})_{ij} + \beta_{3j}(\text{renotras}_{t-1})_{ij} + \beta_{4j}(\text{renmat}_{t-1})_{ij} \\ + \beta_{5j}(\text{INSE})_{ij} + \beta_{6j}(\text{des\_hnos})_{ij} + \beta_{7j}(\text{libros})_{ij} + \beta_{8j}(\text{gest})_{ij} + \beta_{9j}(\text{reg})_{ij} + u_i$$

## Matriz de correlación y descriptores. Nivel primario (arriba) y secundario (abajo).

	1	2	3	4	5	6	7	8	9	10	11	mean	sd
1) Género (masculino=1)	1											0.499	0.500
2) Actitud hacia la materia	0.010	1										75.811	25.952
3) Repitencia (S=1)	0.063	-0.013	1									0.189	0.391
4) Rendimiento otras materias	-0.131	0.077	-0.314	1								84.299	19.209
5) Rendimiento matemática $t_{-1}$	-0.035	0.200	-0.256	0.626	1							82.116	24.078
6) INSE	0.052	-0.048	-0.256	0.204	0.160	1						38.509	16.929
7) Deserción hermanos (S=1)	0.043	-0.010	0.160	-0.119	-0.091	-0.107	1					0.075	0.263
8) Textos escolares (S=1)	-0.024	0.031	-0.106	0.093	0.066	0.202	-0.053	1				0.777	0.416
9) Gestión (privado=1)	-0.015	-0.075	-0.196	0.117	0.084	0.385	-0.097	0.081	1			0.239	0.426
10) Región (CABA=1)	0.002*	-0.068	-0.052	0.027	0.016	0.187	-0.042	0.012	0.121	1		0.061	0.239
11) Rendimiento matemática $t$	0.026	0.083	-0.241	0.289	0.283	0.285	-0.167	0.092	0.257	0.108	1	57.926	20.841

	1	2	3	4	5	6	7	8	9	10	11	mean	sd
1) Género (masculino=1)	1											0.439	0.496
2) Actitud hacia la materia	-0.065	1										80.602	24.356
3) Repitencia (S=1)	0.086	-0.035	1									0.244	0.429
4) Rendimiento otras materias	-0.163	0.126	-0.246	1								78.221	15.380
5) Rendimiento matemática $t_{-1}$	-0.061	0.084	-0.165	0.440	1							71.422	28.144
6) INSE	0.087	-0.117	-0.202	0.111	0.075	1						40.182	15.918
7) Deserción hermanos (S=1)	0.022	-0.021	0.050	-0.055	-0.030	-0.072	1					0.011	0.106
8) Textos escolares (S=1)	-0.171	0.101	-0.120	0.228	0.133	0.213	-0.062	1				64.307	30.319
9) Gestión (privado=1)	-0.052	-0.024	-0.205	0.052	0.024	0.363	-0.038	0.142	1			0.339	0.473
10) Región (CABA=1)	0.019	-0.054	-0.030	-0.052	-0.019	0.207	-0.007	0.070	0.159	1		0.079	0.269
11) Rendimiento matemática $t$	0.093	0.001*	-0.247	0.231	0.213	0.369	-0.074	0.180	0.286	0.134	1	61.299	18.947

*Fuente: elaboración propia en base a datos del ONE 2000.*

# 4. Resultados

## Coeficientes de regresión y valores p

Variable	Nivel Primario				Nivel Secundario			
	OLS	Q <sub>25</sub>	Q <sub>50</sub>	Q <sub>75</sub>	OLS	Q <sub>25</sub>	Q <sub>50</sub>	Q <sub>75</sub>
Constante	26.343 (0.000)	7.512 (0.000)	20.990 (0.000)	40.153 (0.000)	25.228 (0.000)	9.789 (0.000)	21.225 (0.000)	37.902 (0.000)
Género (masculino=1)	1.731 (0.000)	1.795 (0.000)	2.125 (0.000)	2.011 (0.000)	4.396 (0.000)	3.769 (0.000)	5.249 (0.000)	5.509 (0.000)
Actitud hacia la materia	0.049 (0.000)	0.051 (0.000)	0.057 (0.000)	0.055 (0.000)	0.006 (0.008)	0.014 (0.001)	0.002 (0.574)	-0.004 (0.311)
Repitencia (S=1)	-4.819 (0.000)	-6.012 (0.000)	-5.604 (0.000)	-4.467 (0.000)	-4.917 (0.000)	-4.887 (0.000)	-5.398 (0.000)	-5.363 (0.000)
Rendimiento otras materias	0.158 (0.000)	0.196 (0.000)	0.190 (0.000)	0.154 (0.000)	0.143 (0.000)	0.168 (0.000)	0.165 (0.000)	0.139 (0.000)
Rendimiento matemática <sub>t-1</sub>	0.127 (0.000)	0.141 (0.000)	0.150 (0.000)	0.139 (0.000)	0.089 (0.000)	0.097 (0.000)	0.099 (0.000)	0.090 (0.000)
INSE	0.192 (0.000)	0.240 (0.000)	0.229 (0.000)	0.177 (0.000)	0.298 (0.000)	0.316 (0.000)	0.345 (0.000)	0.311 (0.000)
Deserción hermanos (S=1)	-7.804 (0.000)	-8.941 (0.000)	-9.044 (0.000)	-8.214 (0.000)	-5.947 (0.000)	-6.607 (0.000)	-5.776 (0.000)	-6.251 (0.000)
Textos escolares (S=1)	0.650 (0.000)	0.908 (0.000)	0.668 (0.000)	-0.008 (0.000)	0.041 (0.000)	0.042 (0.000)	0.041 (0.000)	0.037 (0.000)
Gestión (privado=1)	5.351 (0.000)	7.259 (0.000)	5.261 (0.000)	3.266 (0.000)	6.880 (0.000)	8.160 (0.000)	7.618 (0.000)	6.095 (0.000)
Región (CABA=1)	3.034 (0.000)	3.973 (0.000)	2.959 (0.000)	1.763 (0.000)	4.213 (0.000)	5.616 (0.000)	4.821 (0.000)	3.120 (0.000)
R <sup>2</sup> y pseudo-R <sup>2</sup>	0.2192	0.1392	0.1388	0.1120	0.2758	0.1532	0.1792	0.1679
n	136,587	136,587	136,587	136,587	90,530	90,530	90,530	90,530
F	3,835				3,448			

Fuente: elaboración propia en base a datos del ONE 2000.

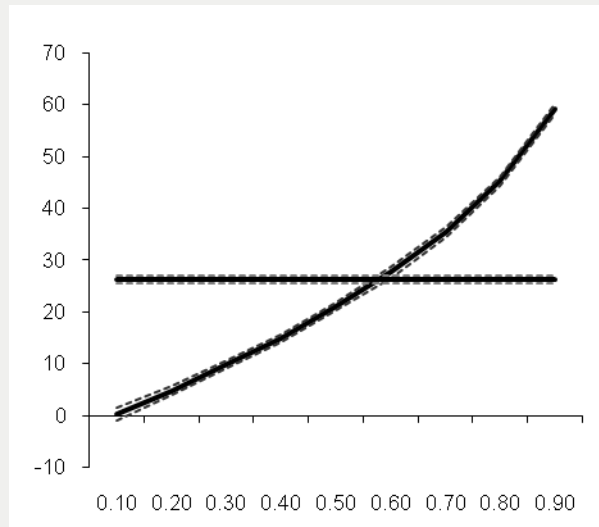
# Prueba de hipótesis de igualdad de coeficientes: valores p

Variable	Nivel Primario			Nivel Secundario		
	q <sub>25</sub> - q <sub>50</sub>	q <sub>25</sub> - q <sub>75</sub>	q <sub>50</sub> - q <sub>75</sub>	q <sub>25</sub> - q <sub>50</sub>	q <sub>25</sub> - q <sub>75</sub>	q <sub>50</sub> - q <sub>75</sub>
Género (masculino=1)	0.020	0.168	0.230	0.000	0.000	0.006
Actitud hacia la materia	0.016	0.150	0.183	0.000	0.000	0.098
Repitencia (S=1)	0.027	0.000	0.000	0.000	0.022	0.863
Rendimiento otras materias	0.325	0.000	0.000	0.617	0.000	0.000
Rendimiento matemática <sub>t-1</sub>	0.062	0.675	0.002	0.427	0.050	0.005
INSE	0.010	0.000	0.000	0.000	0.363	0.000
Deserción hermanos (S=1)	0.752	0.045	0.001	0.318	0.676	0.630
Textos escolares (S=1)	0.183	0.000	0.000	0.709	0.123	0.088
Gestión (privado=1)	0.000	0.000	0.000	0.005	0.000	0.000
Región (CABA=1)	0.000	0.000	0.000	0.011	0.000	0.000

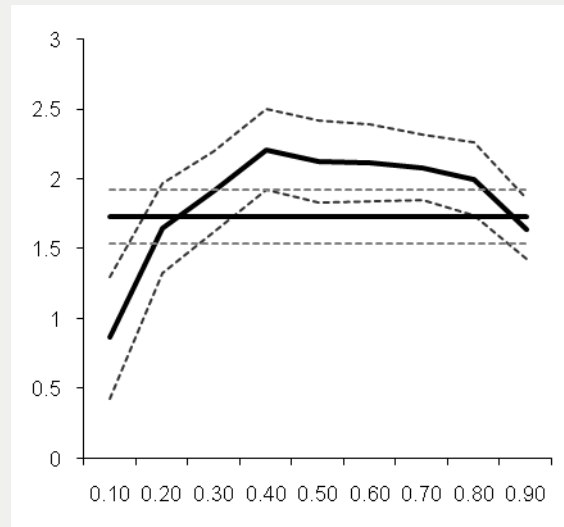
Fuente: elaboración propia en base a datos del ONE 2000.

# Coeficientes estimados por cuantiles: nivel primario

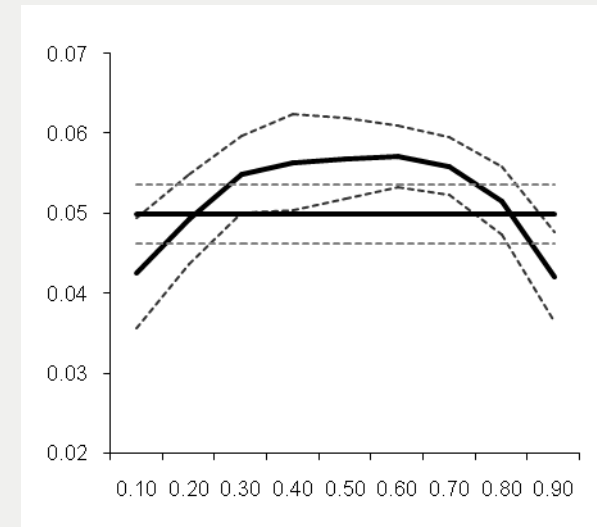
Constante



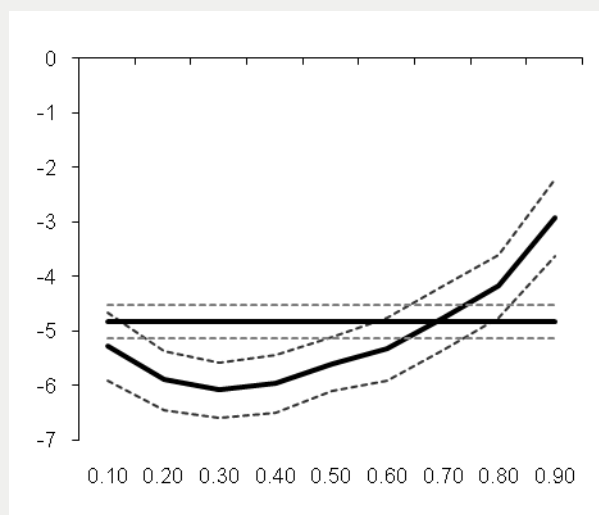
Género



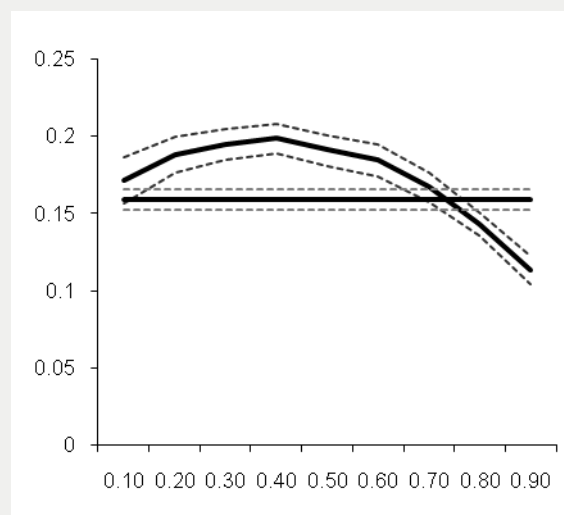
Actitud hacia la materia



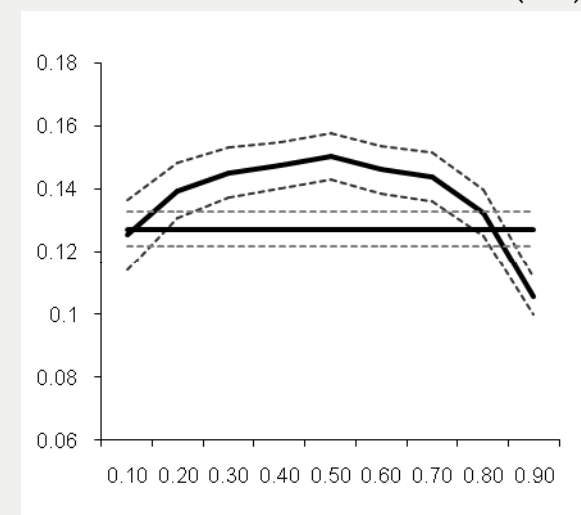
Repitencia



Rendimiento otras materias

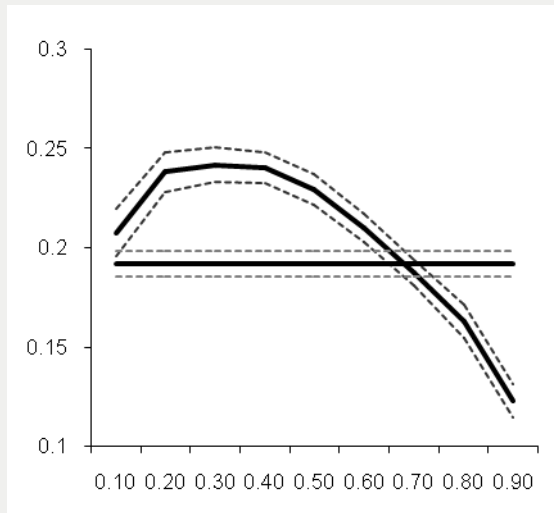


Rendimiento matemática (t-1)

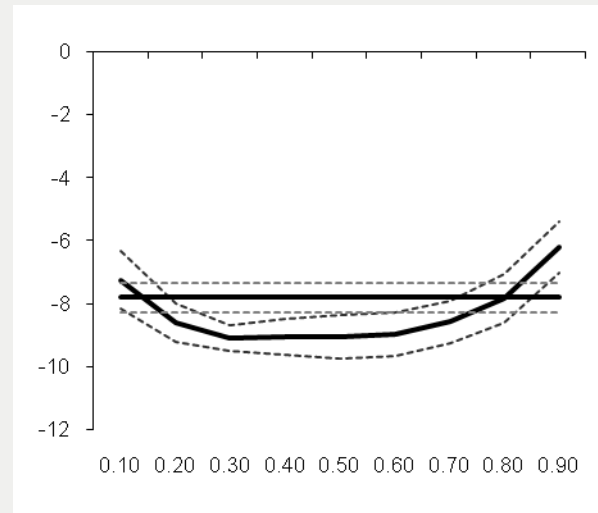


# Coeficientes estimados por cuantiles: nivel primario

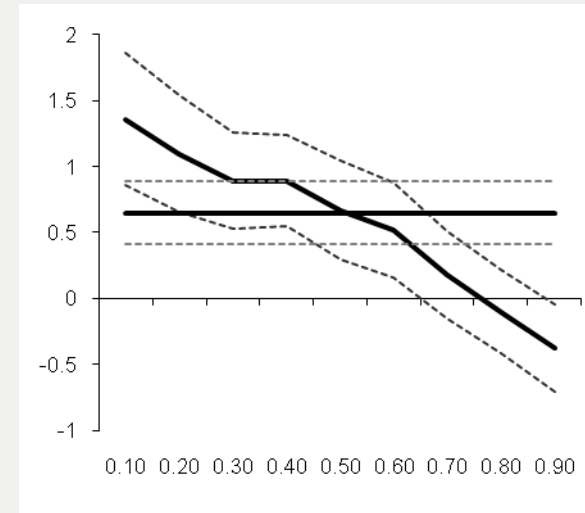
INSE



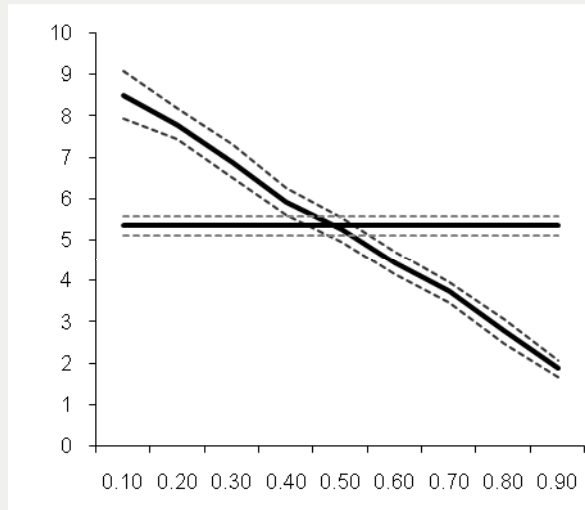
Deserción hermanos



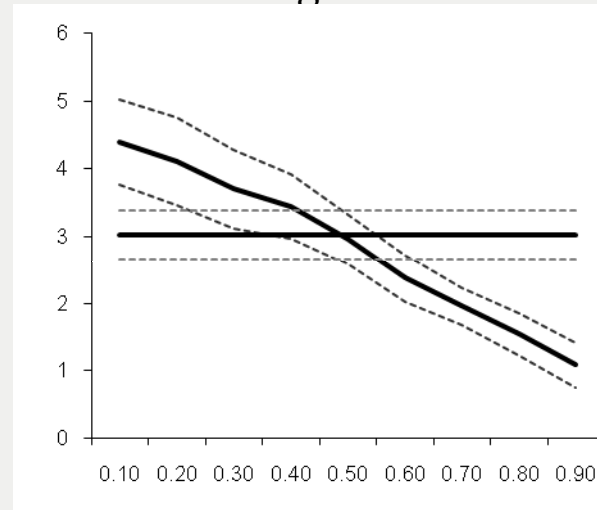
Textos escolares



Gestión

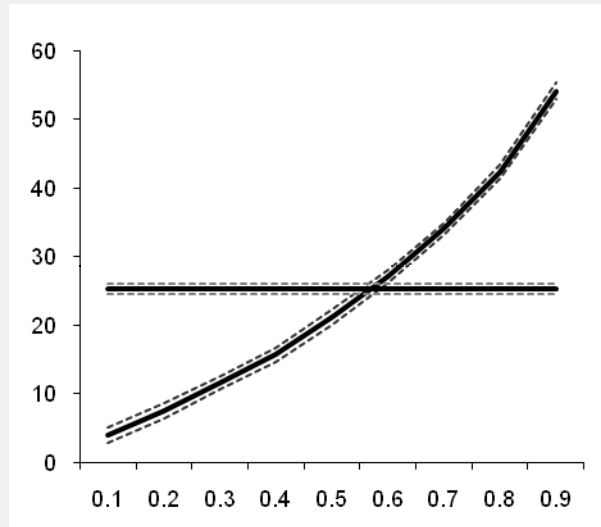


Región

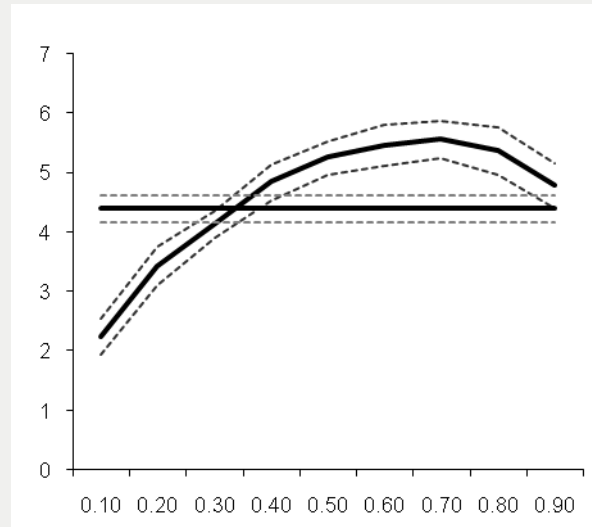


# Coeficientes estimados por cuantiles: nivel secundario

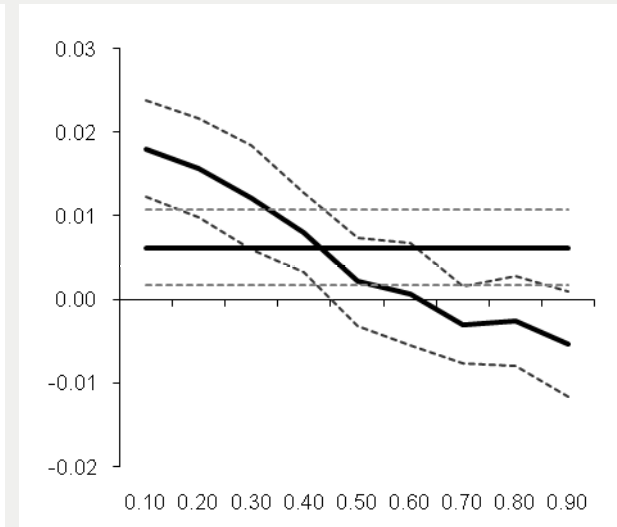
Constante



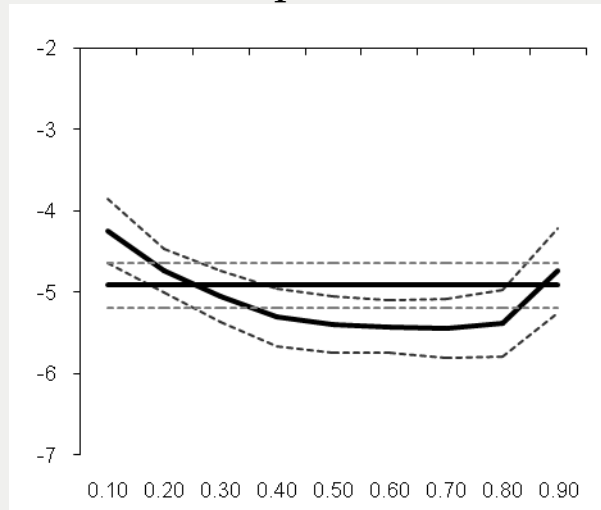
Género



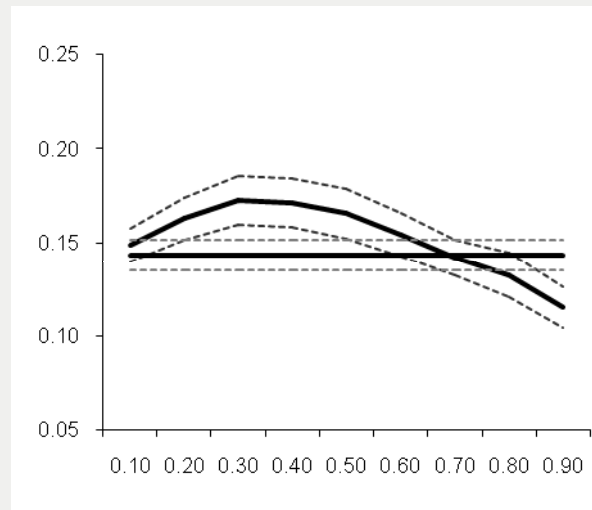
Actitud hacia la materia



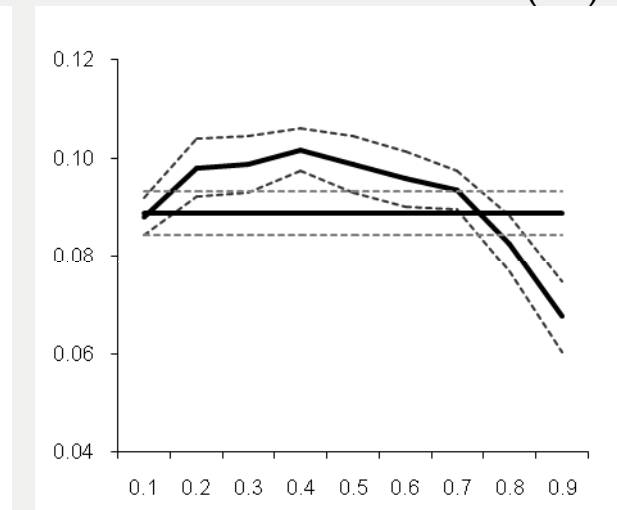
Repitencia



Rendimiento otras materias

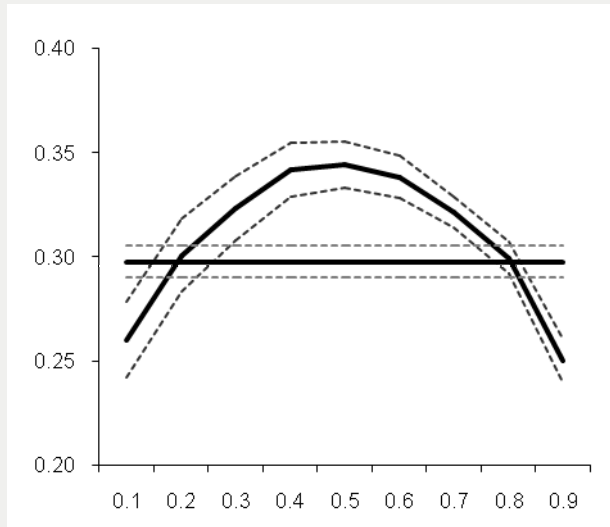


Rendimiento matemática (t-1)

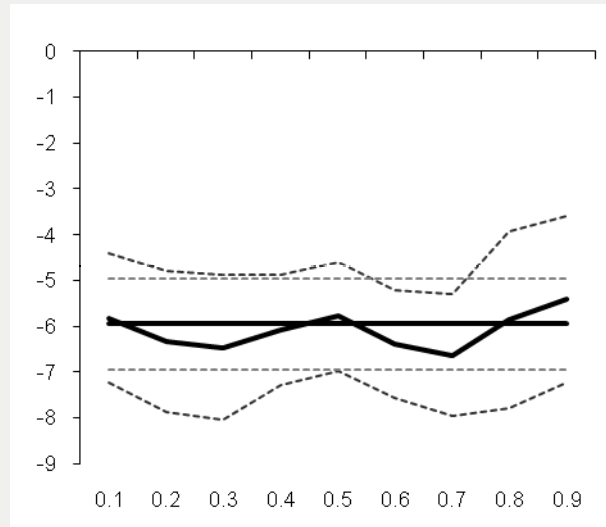


# Coeficientes estimados por cuantiles: nivel secundario

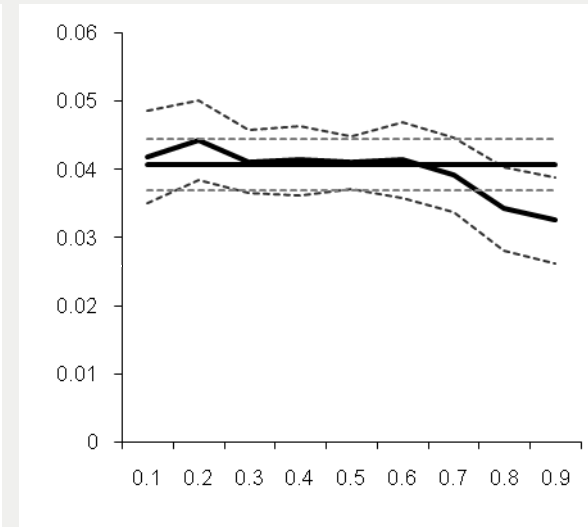
INSE



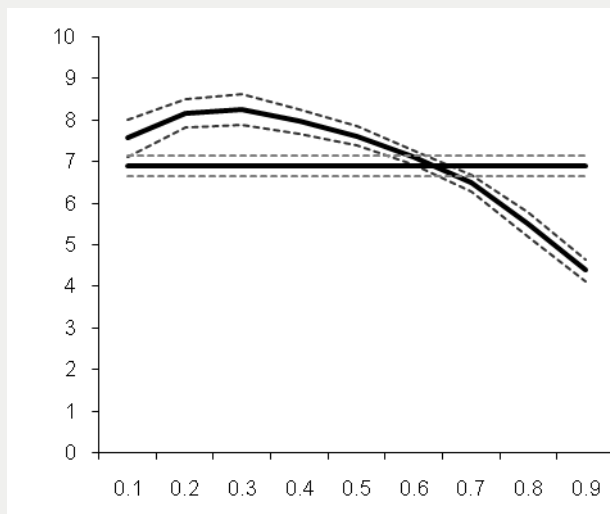
Deserción hermanos



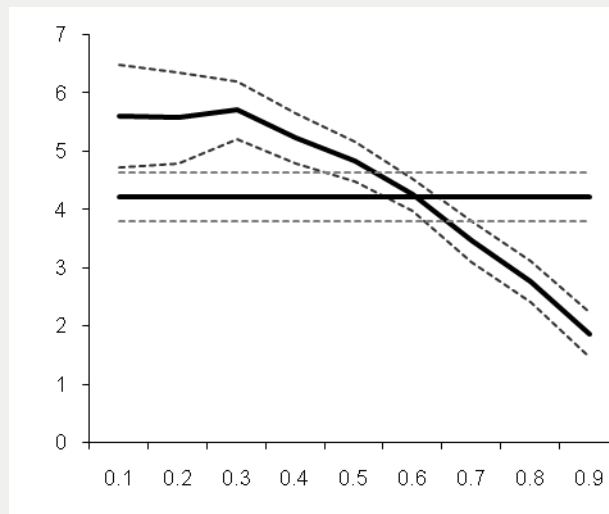
Textos escolares



Gestión



Región





## 5. Conclusiones

- Tanto en el nivel primario como en el secundario, el efecto de los factores intervinientes en el modelo se expresó con un diferente nivel de fuerza en cada uno de los cuantiles.
- Este tipo de análisis proporciona elementos para una mayor focalización de las políticas compensatorias, que podrían entonces diferenciar los tratamientos para el aprendizaje a aplicar sobre los alumnos según sea que éstos se ubiquen más hacia la superior, estén en el centro, o en la parte inferior de la distribución de las notas.

## 6. Referencias Bibliográficas

Alejo, Javier (2006). "Desigualdad Salarial En El Gran Buenos Aires. Una Aplicación De Regresión Por Cuantiles En Microdescomposiciones." Anales XLI Reunión Anual de la Asociación Argentina de Economía Política (AAEP). Disponible en [www.aaep.org.ar/anales/download/2006/programa2006.pdf](http://www.aaep.org.ar/anales/download/2006/programa2006.pdf)

Arias, Omar, Hallock, Kevin and Sosa Escudero, Walter (2000). "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data." Disponible en [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=238898](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=238898).

Bassett, Gilbert and Koenker, Roger (1978). "Regression Quantiles." *Econométrica*, 46(1) 33-50.

Bassett, Gilbert and Koenker, Roger (1982). "An Empirical Quantile Function for Linear Models with iid Errors." *Journal of the American Statistical Association* 77(378), pp. 407-15.

Bassett, Gilbert and Koenker, Roger (1986). "Strong Consistency of Regression Quantiles and Related Empirical Processes." *Econometric Theory* 2, pp. 191-201.

Birch, E. R., and Miller P. W. (2006). Student Outcomes at University in Australia: A Quantile Regression Approach. *Australian Economic Papers*, March, 1-17

Buchinsky, Moshe (1962). "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression." *Econométrica* 62(2), pp. 405-58.

Buchinsky, Moshe (1998). "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach." *Econométrica* 13(1), pp. 1-30.

Cameron Colin A. y Trivedi P.K. (2010). *Microeconometrics using STATA*, Texas. Stata Press

Cervini, Rubén (2003) "Diferencias de resultados cognitivos y no-cognitivos entre estudiantes de escuelas públicas y privadas en la educación secundaria de Argentina: Un análisis multinivel." *Education Policy Analysis Archives*, 11(5).

Cervini, R. (2002) "Desigualdades en el Logro Académico y Reproducción Cultural en la Educación Primaria de Argentina - Un modelo de tres niveles." *Revista Mexicana de Investigación Educativa*, 7(16): 445-500 México: Consejo Mexicano de Investigación Educativa. <http://www.comie.org.mx/revista/Pdfs/Carpeta16/16investTem2.pdf>

Gasparini, L. and Sosa Escudero, W. (2003) "Implicit Rents from Own-housing and Income Distribution: Econometric Estimates for Greater Buenos Aires" *Journal of Income Distribution*. Volume 12, Number 1-2 Spring-Summer.

Gertel Héctor, Giuliadori Roberto, Herrero Verónica and, Fresoli Diego (2007) "Los factores determinantes del rendimiento escolar al término de la educación básica en Argentina. Una aplicación de técnicas de análisis jerárquico de datos". XVI Jornadas de la Asociación de Economía de la Educación. Gran Canaria. España. July 12-13, 2007.

Giuliodori, Roberto F., Gertel, Héctor R., Casini, Rosanna and Gonzalez, Mariana (2009). "Análisis de desempeño académico mediante regresión por cuantiles aplicado a estudiantes de dos facultades de la Universidad Nacional de Córdoba. XXXVII Coloquio Argentino de Estadística. Catamarca. Octubre 7-9.

Hallock, Kevin F. and Koenker, Roger (2001) "Quantile Regression" *The Journal of Economic Perspectives* 15(4), pp. 143-56.

Herrero V., Palacios P., Ruíz Díaz, F. (2005) "Un índice de nivel socio-económico de los hogares para aplicar a un modelo explicativo del rendimiento escolar". XXXIII Coloquio de la Sociedad Argentina de Estadística. Villa Giardino. Córdoba. Argentina.

Kenney-Benson G. et al (2006) "Sex Differences in Math Performance: The Role of Children's Approach to Schoolwork" *Developmental Psychology*, Vol 42 (1), pp. 11-26

Machado, José A. F. and Koenker, Roger (1999). "Goodness of Fit and Related Inference Processes for Quantile Regression." *American Statistical Association* 94(448), pp. 1296-310.

Pereira, Pedro and Martins, Pedro (2004) "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries" *Labour Economics* 11, pp. 355- 71.

Zhu, Li-Xing and He, Xuming (2003). "A Lack-of-Fit Test for Quantile Regression." *Journal of the American Statistical Association* 98(464), pp. 1013-22.