

Social interactions and school achievements in Argentina

María Ana Lugo

World Bank - LCRCE

NIP/LACEA

Medellin, November 2010

Outline

Introduction

Motivation. Segregation and school achievement

Peer-group effects

Data and estimation strategy

Data

Identification and estimation strategy

Results

Results

Conclusions

Motivation. Segregation and school achievement

Objective: to understand the effect of social segregation across schools on students' achievement - mean and gap between poor and non-poor.

- Argentina since 1980s: ↑ social segregation in schools, ↑ gaps in performance between the poor and non-poor (Llach and Schumacher 2006).
- ↑ achievement gap bt. public and private school (Galiani et al 2008) and ↑ Gini in achievements in poorest provinces (Etchart et al 2004).
- Two channels from segregation to achievements:
 1. Distribution of financial and human resources (Mayer 2002)
 2. Composition of the students within the school - **Peer group effects**

Motivation. Segregation and school achievement

Objective: to understand the effect of social segregation across schools on students' achievement - mean and gap between poor and non-poor.

- Argentina since 1980s: ↑ social segregation in schools, ↑ gaps in performance between the poor and non-poor (Llach and Schumacher 2006).
- ↑ achievement gap bt. public and private school (Galiani et al 2008) and ↑ Gini in achievements in poorest provinces (Etchart et al 2004).
- Two channels from segregation to achievements:
 1. Distribution of financial and human resources (Mayer 2002)
 2. Composition of the students within the school - **Peer group effects**

Motivation. Segregation and school achievement

Objective: to understand the effect of social segregation across schools on students' achievement - mean and gap between poor and non-poor.

- Argentina since 1980s: ↑ social segregation in schools, ↑ gaps in performance between the poor and non-poor (Llach and Schumacher 2006).
- ↑ achievement gap bt. public and private school (Galiani et al 2008) and ↑ Gini in achievements in poorest provinces (Etchart et al 2004).
- Two channels from segregation to achievements:
 1. Distribution of financial and human resources (Mayer 2002)
 2. Composition of the students within the school - **Peer group effects**

Motivation. Segregation and school achievement

- Perpetuation of inequality

If children's performance at school depends on their peers, higher levels of social segregation lead to greater inequality in academic achievement and thence to greater inequality in later-life outcomes. And excessive segregation may threaten present-day social cohesion. In some circumstances, greater social segregation may even reduce average achievement levels"

(Jenkins et al. 2006, p. 1)

Motivation. Segregation and school achievement

- Perpetuation of inequality

*If children's performance at school depends on their peers, higher levels of social segregation lead to **greater inequality in academic achievement** and thence to greater inequality in later-life outcomes. And excessive segregation may threaten present-day social cohesion. In some circumstances, greater social segregation may even reduce average achievement levels"*

(Jenkins et al. 2006, p. 1)

Motivation. Segregation and school achievement

- Perpetuation of inequality

*If children's performance at school depends on their peers, higher levels of social segregation lead to greater inequality in academic achievement and thence to greater inequality in later-life outcomes. And excessive segregation may threaten present-day social cohesion. In some circumstances, greater social segregation may even reduce **average achievement levels**"*

(Jenkins et al. 2006, p. 1)

Motivation. Segregation and school achievement

- Perpetuation of inequality

*If children's performance at school depends on their peers, higher levels of social segregation lead to greater inequality in academic achievement and thence to greater inequality in later-life outcomes. And excessive segregation may threaten present-day **social cohesion**. In some circumstances, greater social segregation may even reduce average achievement levels"*

(Jenkins et al. 2006, p. 1)

Peer-group effects

- General peer-group effects: impact of characteristics of classmates on a student performance (A_i).

$$A_i = f(\mathbf{x}_{-i})$$

where \mathbf{x}_{-i} is a vector of characteristics of classmates, excl. i .

- Manski (1993). Endogenous effects (behaviour) and contextual effects (characteristics).
- Generally, $\mathbf{x}_{-i} = \underline{\text{average}}$ level of the classmates ($\bar{A}_{-i}, \bar{Y}_{-i}$)
- I depart from this standard approach.

Peer-group effects

- General peer-group effects: impact of characteristics of classmates on a student performance (A_i).

$$A_i = f(\mathbf{x}_{-i})$$

where \mathbf{x}_{-i} is a vector of characteristics of classmates, excl. i .

- Manski (1993). Endogenous effects (behaviour) and contextual effects (characteristics).
- Generally, $\mathbf{x}_{-i} = \underline{\text{average}}$ level of the classmates ($\bar{A}_{-i}, \bar{Y}_{-i}$)
- I depart from this standard approach.

Beyond the standard approach to peer-group

- $x_{-i} = \bar{A}_{-i}, \bar{Y}_{-i}$ assumes that
 1. **Constant returns** to the mean (linear).
 2. **Equal weight** to all students (irresp. recipient or composition)
- Increase segregation has NO effect on average achievements. Only increases gap between poor and non-poor (iff peer-effect is positive)
- In previous work, I tested the first assumption
 - Concave relationship bt. \bar{Y}_{-i} and A_i . i.e. integration increases mean achievements.
 - (first attempt at ass. 2) Standard deviation matters, then segregation is good for average but bad for the poor.

Beyond the standard approach to peer-group

- $x_{-i} = \bar{A}_{-i}, \bar{Y}_{-i}$ assumes that
 1. **Constant returns** to the mean (linear).
 2. **Equal weight** to all students (irresp. recipient or composition)
- Increase segregation has **NO** effect on average achievements. Only increases gap between poor and non-poor (iff peer-effect is positive)
- In previous work, I tested the first assumption
 - Concave relationship bt. \bar{Y}_{-i} and A_i . i.e. **integration** increases mean achievements.
 - (first attempt at ass. 2) **Standard deviation** matters, then **segregation** is good for average but bad for the poor.

Beyond the standard approach to peer-group

- $x_{-i} = \bar{A}_{-i}, \bar{Y}_{-i}$ assumes that
 1. **Constant returns** to the mean (linear).
 2. **Equal weight** to all students (irresp. recipient or composition)
- Increase segregation has NO effect on average achievements. Only increases gap between poor and non-poor (iff peer-effect is positive)
- In previous work, I tested the first assumption
 - Concave relationship bt. \bar{Y}_{-i} and A_i . i.e. **integration** increases mean achievements.
 - (first attempt at ass. 2) Standard deviation matters, then **segregation** is good for average but bad for the poor.

Relaxing assumption 2 in two ways

1. **Composition effects**, classmates are weighted differently depending on their position in the distribution.

$$\{i; 3, 3\} \quad \textit{versus} \quad \{i; 1, 5\}$$

$$A_i(3, 3) \quad \neq \quad A_i(1, 5)$$

2. **Heterogenous effects**, depending on the socioeconomic level (SES) of recipient student i . How receptive students are?

$$\{1; 3, 3\} \quad \textit{versus} \quad \{5; 3, 3\}$$

$$A_1(3, 3) \quad \neq \quad A_5(3, 3)$$

Ref: C Hoxby 2000 (race)

Relaxing assumption 2 in two ways

1. **Composition effects**, classmates are weighted differently depending on their position in the distribution.

$$\{i; 3, 3\} \quad \textit{versus} \quad \{i; 1, 5\}$$

$$A_i(3, 3) \quad \neq \quad A_i(1, 5)$$

2. **Heterogenous effects**, depending on the socioeconomic level (SES) of recipient student i . How receptive students are?

$$\{1; 3, 3\} \quad \textit{versus} \quad \{5; 3, 3\}$$

$$A_1(3, 3) \quad \neq \quad A_5(3, 3)$$

Ref: C Hoxby 2000 (race)

Preview of results

- Classmates with parents with secondary complete education have the largest effects (w. caution: $\{3, 3\} \succ \{1, 5\}$). \Rightarrow mixing with the mean
- Children from more educated families are able to capture more benefits from the peers' parental education. \Rightarrow The poor loose out again.
- When combining both effects, larger effects from those close to me, but stronger among the rich.

Outline

Introduction

Motivation. Segregation and school achievement

Peer-group effects

Data and estimation strategy

Data

Identification and estimation strategy

Results

Results

Conclusions

Outline

Introduction

Motivation. Segregation and school achievement

Peer-group effects

Data and estimation strategy

Data

Identification and estimation strategy

Results

Results

Conclusions

Data

Data is taken from the 2000 National Evaluation Assessment, Ministry of Education.

- Census of all schools (private/public), all children in the sixth grade (10 years-old approx). Over 600,000 students.
- **Student-level dataset**, with also information about the school and principals.
- Students are assessed in **Spanish**, Maths and Sciences (0 to 100)
- After the test they are asked about parental education (**average**) and durable goods at home

Data

Students characteristics

	Students groups of parental education				
	<complete 1	Complete 1	Complete 2	Complete 3	Total
cases	39,274	107,913	89,513	59,215	295,915
%	0.13	0.36	0.30	0.20	1.00
Parental education (1-6)	1.20	2.60	4.51	6.00	3.67
std dev	0.25	0.57	0.55	0.00	1.66
min	1.0	2.0	4.0	6.0	1.0
max	1.5	3.6	5.5	6.0	6.0
Spanish test score	57.24	62.27	66.01	69.39	64.16
Assets	19.97	24.57	29.04	32.90	26.98
st dev assets	12.86	13.12	13.01	12.16	13.50
Male	0.45	0.47	0.52	0.51	0.49
Repeat	0.31	0.19	0.13	0.08	0.16
Private school	0.12	0.20	0.33	0.48	0.29
Class size	25.07	26.28	26.95	27.21	26.51

Table 2.1: Summary statistics, students' parental education.

Data

Students characteristics

	Students groups of parental education				
	<complete 1	Complete 1	Complete 2	Complete 3	Total
cases	39,274	107,913	89,513	59,215	295,915
%	0.13	0.36	0.30	0.20	1.00
Parental education (1-6)	1.20	2.60	4.51	6.00	3.67
std dev	0.25	0.57	0.55	0.00	1.66
min	1.0	2.0	4.0	6.0	1.0
max	1.5	3.6	5.5	6.0	6.0
Spanish test score	57.24	62.27	66.01	69.39	64.16
Assets	19.97	24.57	29.04	32.90	26.98
st dev assets	12.86	13.12	13.01	12.16	13.50
Male	0.45	0.47	0.52	0.51	0.49
Repeat	0.31	0.19	0.13	0.08	0.16
Private school	0.12	0.20	0.33	0.48	0.29
Class size	25.07	26.28	26.95	27.21	26.51

Table 2.1: Summary statistics, students' parental education.

Data

Students characteristics

	Students groups of parental education				
	<complete 1	Complete 1	Complete 2	Complete 3	Total
cases	39,274	107,913	89,513	59,215	295,915
%	0.13	0.36	0.30	0.20	1.00
Parental education (1-6)	1.20	2.60	4.51	6.00	3.67
std dev	0.25	0.57	0.55	0.00	1.66
min	1.0	2.0	4.0	6.0	1.0
max	1.5	3.6	5.5	6.0	6.0
Spanish test score	57.24	62.27	66.01	69.39	64.16
Assets	19.97	24.57	29.04	32.90	26.98
st dev assets	12.86	13.12	13.01	12.16	13.50
Male	0.45	0.47	0.52	0.51	0.49
Repeat	0.31	0.19	0.13	0.08	0.16
Private school	0.12	0.20	0.33	0.48	0.29
Class size	25.07	26.28	26.95	27.21	26.51

Table 2.1: Summary statistics, students' parental education.

Empirical specification

- Class composition effect:

$$A_i = \alpha + \sum_{g=2}^G \delta_g P_g + \beta(Y_i, Z) + \epsilon_i$$

where

- $P_g, g = 2, \dots, G$ is the % of classmates in group g (baseline group 1)
 - Y_i and Z are vectors of individual and school characteristics
 - ϵ_i is the unobserved individual term
- Heterogenous effects:

$$A_i = \alpha + \sum_{g=1}^G \gamma_g Q_{gi} * \text{mean}(Y)_{-i} + \beta(Y_i, Z) + \epsilon_i$$

where

- $Q_{gi}, g = 1, \dots, G$ are dummy vles on the SES group of student i
- $\text{mean}(Y)_{-i}$ is the average classmates characteristics, excl. student i

Empirical specification

- Class composition effect:

$$A_i = \alpha + \sum_{g=2}^G \delta_g P_g + \beta(Y_i, Z) + \epsilon_i$$

where

- $P_g, g = 2, \dots, G$ is the % of classmates in group g (baseline group 1)
 - Y_i and Z are vectors of individual and school characteristics
 - ϵ_i is the unobserved individual term
- Heterogenous effects:

$$A_i = \alpha + \sum_{g=1}^G \gamma_g Q_{gi} * \text{mean}(Y)_{-i} + \beta(Y_i, Z) + \epsilon_i$$

where

- $Q_{gi}, g = 1, \dots, G$ are dummy vles on the SES group of student i
- $\text{mean}(Y)_{-i}$ is the average classmates characteristics, excl. student i

Identification and estimation strategy

3 commonly mentioned problems

1. **Reflection problem**: identification of endogenous vs. contextual effects → **General peer effects**

(Manski 1993, Brock and Durlauf 2001/02, Katz 2001, Moffitt 2005)

2. **Self-selection problem**: unobserved characteristics determining school choice and achievement (e.g. parents' ambition) → school fixed effects

(Sacerdote 2001, McEwan 2003; Evans et al. 1992, Rivkin 2001; Ioannides 2003)

Identification and estimation strategy

3 commonly mentioned problems

1. **Reflection problem**: identification of endogenous vs. contextual effects → General peer effects

(Manski 1993, Brock and Durlauf 2001/02, Katz 2001, Moffitt 2005)

2. **Self-selection problem**: unobserved characteristics determining school choice and achievement (e.g. parents' ambition) → **school fixed effects**

(Sacerdote 2001, McEwan 2003; Evans et al. 1992, Rivkin 2001; Ioannides 2003)

Estimation strategy

- 3. Correlated effects problem:** unobserved factors correlated with characteristics of peers and the achievement of the student (teachers' quality). → **Random assignment** of students across classes (w. school FE)

(Ammermueller 2009, McEwan 2003, Hoxby 2000, Sacerdote 2001; Graham 2005; Brock and Durlauf 2002; Duflo and Saez 2003, Duflo et al. 2004, Miguel and Kremer 2004)

The effect is identified by the variation across classes within a given school. I restrict the analysis to 'big schools' -approx half of schools, almost 4,000 schools, 9,700 classes and 300,000 students

Estimation strategy

3. **Correlated effects problem:** unobserved factors correlated with characteristics of peers and the achievement of the student (teachers' quality). → **Random assignment** of students across classes (w. school FE)

(Ammermueller 2009, McEwan 2003, Hoxby 2000, Sacerdote 2001; Graham 2005; Brock and Durlauf 2002; Duflo and Saez 2003, Duflo et al. 2004, Miguel and Kremer 2004)

The effect is identified by the variation across classes within a given school. I restrict the analysis to 'big schools' -approx half of schools, almost 4,000 schools, 9,700 classes and 300,000 students

Outline

Introduction

Motivation. Segregation and school achievement

Peer-group effects

Data and estimation strategy

Data

Identification and estimation strategy

Results

Results

Conclusions

Class composition effect

Differences according to the distribution of classmates -for the average pupil

Dependent vble: Test scores in Spanish	
Prop class with C1 (baseline = I1)	.016 (.007)**
Prop class with C2 (baseline = I1)	.032 (.007)***
Prop class with C3 (baseline = I1)	.025 (.008)***
edu parents	.195 (.024)***
male	-5.305 (.072)***
repeat	-7.991 (.097)***
class size	.233 (.022)***
Const.	59.118 1 (.787)***
school fixed effect	yes
Obs.	295915
e(N-g)	9704
R ²	.062

Note: Standard errors are robust and clustered by

school

- ↑ the % of children with more educated parents, ↑ the achievement of the average pupil.
- The largest effect is by completed secondary school parents (and not tertiary!). BUT the difference is not significant.
- The implication is that mixing poor and rich with middle class though not among themselves is preferable. (w caution)

Heterogenous effects

Differences according to student's background

	Dependent vble: Test scores in Spanish	
	(1)	(2)
mean(educ)(-i)	.527 (.134)***	
l1i*mean(educ)(-i)		-.076 (.142)
C1i*mean(educ)(-i)		.484 (.135)***
C2i*mean(educ)(-i)		.742 (.135)***
C3i*mean(educ)(-i)		.800 (.137)***
educ parents	.195 (.024)***	-.303 (.057)***
male	-5.305 (.072)***	-5.304 (.072)***
repeat	-7.993 (.097)***	-7.944 (.097)***
class size	.233 (.022)***	.231 (.022)***
Const.	59.249 (.729)***	60.932 (.753)***
school fixed effect	yes	yes
Obs.	295915	295915
e(N-g)	9704	9704
R ²	.062	.063

Note: Standard errors are robust and clustered by school

- There is a **positive peer effect**. Increasing one std deviation the mean of classmates parental background increases the test score by almost one point - 2 1/2 times the effect of increasing her own parental background.
- The better educated the parents, the more the pupil is able to capture benefits from her peers' parental background. This support the idea of heterogenous effects.

Heterogenous effects

Differences according to student's background

	Dependent vble: Test scores in Spanish	
	(1)	(2)
mean(educ)(-i)	.527 (.134)***	
l1i*mean(educ)(-i)		-.076 (.142)
C1i*mean(educ)(-i)		.484 (.135)***
C2i*mean(educ)(-i)		.742 (.135)***
C3i*mean(educ)(-i)		.800 (.137)***
educ parents	.195 (.024)***	-.303 (.057)***
male	-5.305 (.072)***	-5.304 (.072)***
repeat	-7.993 (.097)***	-7.944 (.097)***
class size	.233 (.022)***	.231 (.022)***
Const.	59.249 (.729)***	60.932 (.753)***
school fixed effect	yes	yes
Obs.	295915	295915
e(N-g)	9704	9704
R ²	.062	.063

Note: Standard errors are robust and clustered by school

- There is a positive peer effect. Increasing one std deviation the mean of classmates parental background increases the test score by almost one point - 2 1/2 times the effect of increasing her own parental background.
- The better educated the parents, the more the pupil is able to capture benefits from her peers' parental background. This support the idea of **heterogenous effects**.

Class compositions effects by student's background

Dependent variable:		Students groups of parental education			
Spanish test scores		Incomplete 1	Complete 1	Complete 2	Complete 3
Proportion of classmates in each quantile (average) excl. <i>i</i>	I1	0.019		-0.051	-0.119
		(0.006)***		(0.006)***	(0.008)***
	C1	-0.008	0.016	-0.026	-0.058
		(0.004)**	(0.008)**	(0.003)***	(0.005)***
	C2	-0.018	0.013	0.009	-0.001
	(0.005)***	(0.008)	(0.004)**	(0.005)	
	C3	-0.020	-0.016	0.026	0.035
		(0.008)***	(0.009)*	(0.005)***	(0.005)***
	Obs.	295,915			
	N groups	9,704			
	R ²	0.065			
controls: mean(edu), edu _{<i>i</i>} , gender, repeat, class size					

Robustness

- Public schools only
- Math test scores
- Assets' index
- Combined effects controlling for mean characteristics

Outline

Introduction

Motivation. Segregation and school achievement

Peer-group effects

Data and estimation strategy

Data

Identification and estimation strategy

Results

Results

Conclusions

Conclusions

- Need to go beyond the standard peer-group effect approach to understand consequences of increase segregation.
- Some evidence of class composition effects. Mixing with this middle group may increase overall performance.
- The more educated the parents, the more the child is able to benefit from her classmates. Peer-group effect is another source of increasing gap between social groups.
- Combined effects: larger benefits from socially closer peers.
- Consequences of integrations/segregation only view in terms of achievements. But Jenkins et al. (2006) 'social cohesion'

School fixed effects

Table 5.1: Regression of Spanish tests scores on parental education.
School fixed effects

	Dep vble: Spanish test scores		
	OLS	OLS	FE
parental education; class mean(educ) $(-i)$	1.899 (.037)***	0.585 (.022)***	0.195 (.024)***
	—	5.586 (.080)***	0.527 (.134)***

controls: male, repeat, class size

Random assignment

Testing random assignment of students across classes within a given school. Steps:

1. Run regression of variable of interest (parental education) using school fixed effects regression

$$Edu_{ijk} = cons + u_k + \epsilon_{ijk}$$

2. Run regression of the residuals from the previous regression on class fixed effects

$$\epsilon_{ijk} = cons + v_j + \omega_{ijk}$$

3. Test whether the class fixed effects (v_j) are jointly equal to zero (F-test)

◀ Back

Random assignment

```

Fixed-effects (within) regression
Group variable: idsec

R-sq:  within = 0.0000
       between = 0.0000
       overall = 0.0000

corr(u_i, Xb) = 0.0000

Number of obs      = 295915
Number of groups   = 21107

Obs per group:  min = 2
                avg  = 14.0
                max  = 40

F(0,274808)       = 0.00
Prob > F          = .

-----+-----
      erroredu |          Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----+-----
      _cons    |  2.08e-10      .0026245     0.00   1.000    - .005144   .005144
-----+-----
      sigma_u  |  .3827878
      sigma_e  |  1.4276788
      rho      |  .06706653   (fraction of variance due to u_i)
-----+-----
F test that all u_i=0:   F(21106, 274808) = 0.84      Prob > F = 1.0000
  
```

Table 5.2: Test of random assignment of students across classes within a school.

Class compositions effects by student's background, controlling for mean(edu)

Dependent vle: Spanish test scores		Students groups of parental education			
		Incomplete 1	Complete 1	Complete 2	Complete 3
Proportion of classmates in each quantile (average) excl. i	I1	0.019 (0.006)***		-0.052 (0.006)***	-0.120 (0.008)***
	C1	-0.008 (0.004)**	0.008 (0.010)	-0.026 (0.003)***	-0.059 (0.005)***
	C2	-0.018 (0.005)***	-0.008 (0.017)	0.009 (0.004)**	-0.001 (0.005)
	Controlling for mean(edu)	C3	-0.020 (0.008)**	-0.046 (0.024)*	0.026 (0.005)***
	Obs.	295,915			
	N groups	9,704			
	R ²	0.065			
controls: mean(assets), assets _{<i>i</i>} , gender, repeat, class size					