Social interactions and school achievement in Argentina *

María Ana Lugo †
University of Oxford

Preliminary draft: February 27, 2009

Abstract

Argentina is a highly unequal society with an increasingly segregated educational system. As economic gaps widen, children have fewer and fewer opportunities to interact with others from different social background. This paper explores the likely impacts of this trend on students' academic achievement in Argentina and the potential effect of policies to redress segregation. It builds a theoretical framework combining two opposing social interaction views: the role models and the competition hypotheses. A direct consequence of this approach is that the allocation of students across schools on economic grounds plays an important role in the aggregate test scores results. Data on test scores of 10-11 year old children in Argentina is used. The identification is based on school fixed effects and the random assignment of students across classes within schools. Positive role models have a significant effect on both rich and poor. However, when social distances are large, 'role model' effects are outweighed by the detrimental impacts of competition on the very poor. These results provide a more nuanced and theoretically grounded picture of the effect of social interactions on student performance, going beyond the estimation of mean peer effects that have become standard in the literature.

JEL codes: D31, I20, I21, O15
Keywords: Economic Segregation, Inequality, Academic Achievement

1 Introduction

Argentina is a highly unequal society with an increasingly segregated educational system. According to the latest official figures, the Gini coefficient of per capita household income is 0.51, and one in four children attend private primary schools –compared to one in six 20 years before (Naradowski & Andrada 2001). This means that while the economic gaps are widening, children have less and less opportunities to interact with others from different social background. In this context, it is important to understand the consequences of this change on students' academic achievement. Among other factors, the socioeconomic status of classmates and neighbours can influence a child's effort

* Preliminary draft, please do not cite without permission. Comments are most welcome. I am most grateful to Tony Atkinson, Stefan Dercon, Jean-Yves Duclos and Justin Sandefur for their comments and suggestions. I am also indebted to Marta Kisilevsky at the Ministry of Education for providing the data, and to Cecilia Veleda for her continuous help elucidating the secrets of the Argentine educational system. All errors remain mine.

† Department of Economics, University of Oxford, Manor Road Building, Manor Road, OX1 3UQ, Oxford, UK. maria.lugo@economics.oxford.ac.uk

at school. Based on alternative views on the externalities of social interactions, I provide a theoretical framework where these different views are integrated and opposing forces are allowed to work simultaneously. Taking proper account of these factors is key to understand the consequences of increasing segregation of children across schools on economic grounds.

The literature on the economics and sociology of education has long studied the determinants of pupils' achievements at school. Whilst most of the variation in performance is explained by children's own inputs –ability, dedication, and family background– (Hanushek 1986, Ermisch & Francesconi 2001, Fertig 2003, Woessmann 2004), and school inputs and technologies –teachers, quality of teaching, incentives structure– (Hanushek 1971, Hanushek, Kain, Markman & Rivkin 2003, Hanushek & Rivkin 2006), factors related to social interaction with others are found to be significant determinants of school performance. Two main directions have been taken. The first line of work relates to how the behaviour and characteristics of classmates –generally referred to as peer-group effects– can benefit or harm children's own performance (Manski 1993, Cutler & Glaeser 1997, Hanushek et al. 2003). Other studies focus on whether neighbourhood characteristics – i.e. outside the school – can influence what happens within the school (Borjas 1995, Cutler & Glaeser 1997, Lazear 1999). According to some of these investigations, particularly 'bad' neighbourhood conditions, such as high unemployment and high levels of poverty and crime, can deter students from performing to the best of their abilities. In both peer-group and neighbourhood effects literature, great attention has been given to the problems of correctly identifying these effects (Manski 2000, especially) and to alternative ways to get around these issues (non-linear functional form estimation in Brock and Durlauf 2001, 2002; laboratory experiments in Falk and Ichino 2003; randomized experiments in Duflo *et al* 2004 and Miguel and Kremer 2004; group changing interventions in Sacerdote 2001, Zimmerman 2003, Kremer and Levy 2003; instrumental variables Case and Katz 1991 and Hanushek *et al* 2003; explicit modelling of selection Ioannides and Zabel 2003 –see Soetevent 2006 for an excellent literature review). Yet, relatively little has been done on providing a tighter link between the sociological and anthropological views on the effects of social interactions in a community and their implication for estimation.

There are two opposing views as to how other people's characteristics affect students' educational aspirations, in turn a key determinant of academic achievement. In the first view, children benefit when their school peers have high socioeconomic status since this provides them with access to alternative role models, extra parental tutorial and networks (Wilson 1997). In the second view, pupils' competitive feelings can drive down their own academic achievements when they perceive themselves to be relatively deprived (Cutler & Glaeser 1997). Testing these ideas requires including non-linear terms previously ignored in the literature. In particular, a child performance at school is thought to depend on the social distance between himself and each classmate while allowing for the effect to differ whether the child is richer or poorer than his peer. Even when this effect is linear for each pair of students, the cumulative effects of all pupils' social position on one child's achievement is best represented non-linearly. The present paper analyses these interactions both within the school and the the neighbourhood and assess whether these two channels (peer-group and neighbourhood effects) can shed some light on the impact of economic inequality on students' performance at school.

In an environment of high disparities between families, the scope for segmentation on economic grounds is larger. In other words, children can be allocated across schools either according to their family socioeconomic background –hence schools are homogeneous within but different between– or irrespectively of it thus heterogeneous within and similar between. When the effects of social interactions are assumed to be linear (as in most studies) the extent of segmentation on economic grounds does not impact the average academic achievement. However, a proper account of social interaction effects requires dropping the linearity restriction. When this assumption is lifted, the impact of economic inequality on children's outcomes will be mediated by the allocation of pupils across schools on economic grounds. Indeed, the extent of economic segregation will determine the *composition of children within each school*, and hence the strength and form of the effects of social interactions within the school, and the socioeconomic *distance across schools* within a city, therefore the extent to which the neighbourhood effect varies according to the relative position of the school. Allowing for neighbourhood and peer-group effects play a role simultaneously can reinforce the importance segregation channel. The ultimate objective is to study whether it matters if the school is the place where social distances are 'erased' and children are integrated irrespective of their background and the level of inequality in their locality, or, on the contrary, the school system is such that precludes the mixing of backgrounds.

The paper uses data from the Argentine test *Operativo de Evaluación Educativa*, a largely unexplored set of data on 10-11 year-old children administered by the federal government. Students are assessed in Mathematics and Spanish on a yearly basis. I combine this information with the reported answers on parental education and capital assets held at home as proxies for socioeconomic status.

The identification of effects is based on two estimation strategies: the year-to-year variations in a given school, and the random assignment of students across classes within a given school. The school fixed effects component allows me to control, to some extent, for the possible biases due to unobserved factors determining the selection of schools that might be also correlated with performance –inasmuch as the parents' selection of school due to, say, reputation is common across all children that attend a particular school. In turn, exploiting the random variation in peer-groups of classes in a given school will permit to control time-varying factors that are correlated with unobserved students characteristics. The main results show that there is support for the idea that positive role models are present for both poor and rich children. However, when the social distances are large enough, the detrimental impacts of competition on the very poor can be rather strong. Specifically, the fact that the average peer-group effect is concave implies that mixing (i.e. desegregation) is preferred. In addition, the impact of mixing is unambiguously positive for children from a poor background. Finally, once all school level effects are accounted for, neighbourhood effects *per se* are not significant in explaining any variations in achievement.

The next section describes the education system in Argentina and data used. Section 3 presents a theoretical framework for analysing socioeconomic segregation through the lens of peer-group and neighbourhood effects. Section 4 presents the results and section 5 concludes.

2 Argentina

"Argentina is a rich and educated country", or so goes the popular refrain. It is certainly true that among other countries in the region, Argentina is one of the wealthiest and is well positioned in terms of educational indicators. Per capita GDP is around US$11,000, the literacy rate is nearly 98%, primary net enrolment is close to 100%, and secondary school enrolment is just above 80%.[1] However, as in much of Latin America, wealth is distributed extremely unequally and poverty is far from eradicated. According to the latest official figures, the Gini index of per capita household income is 0.51, 16% of people live below the two-dollar-a-day poverty line, and 45% below the national poverty line.

While enrolment rates are consistently high for all provinces and socioeconomic groups, the quality of education received is much less equal. Publicly provided primary education is universally available, and the public school system accounts for approximately 78% of the total student population. Public schools are free and, in theory, offer the same instruction across the country. Still, it is argued that schools with a high proportion of poor students might either receive less resources or/and are less efficient in using these resources to produce good results (Fiszbein 1999). In terms of achievement, average performance in test scores of public schools varies from a minimum of 35 to a maximum of 80 (out of 100).[2] These differences are exacerbated if one considers private establishments, wrere the average is twelve points above that of public schools.

In this context, it seems natural to wonder whether indeed economic inequality affects students' performance at schools; and more interestingly, to explore through which mechanism it does so.

2.1 The Primary Education System

Formal education in Argentina is compulsory until the ninth grade (approximately 15 years of age). The provision of primary schooling, its financing, curriculum and administration are independently controlled by each province.[3] Most of the funding of public schools comes from the provincial government. Payment of teachers salaries, large construction work, and school canteens are centralized at the provincial level. The allocation of teachers is determined by the central authorities in the system, and in most cases, local school authorities have no say in the selection of teachers. Public schools also have school cooperatives (*asociación cooperadora*) to which parents contribute a relatively small monthly fee, money which is then used to cover small repair and maintenance work. Finally, in the context of a compensatory programme for specially disadvantage schools and students, state schools might receive restricted resources from the federal government which can only be used to pay for infrastructure work, provision of didactic material, support of special schools initiatives, and training (Fiszbein 1999, p. 11). Private schools, on the other hand, are funded mainly by fees charged to students and by subsidies from the central government. Subsidies are in the form of salaries for teachers, and vary between 20% and 100% of the total salary, the latter mainly in the case of schools run by churches.

Other than monetary resources, human resources may be distributed unequally and biased across school depending on the composition of the student body. In the public sector, teachers can choose where to be posted, with priority given according to a points system. Points are accrued by tenure, training, and evaluations from school directors. In practice, the less experienced teachers are those teaching in the less favourable schools

[1] INDEC, Censo Nacional de Población y Vivienda (2001)

[2] Author's calculation from ONE 2000

[3] The 1993 Federal Law of Education assigned provinces a leading role in matters of financing, pedagogy, and administration of primary and secondary schools (Fiszbein 1999).

(i.e. with a relatively high proportion of poor children). While points data is not released, publicly available teachers' experience data provides a means of computing the quality of instruction across schools. These data, described in detail on the next section show substantial variance between schools. Most of this variation is found within public schools. Average experience in public and private schools does not differ substantially (in both cases it is around 10 years), however the variance is twice as high as that within the private sector.

The level of segregation across schools in a neighbourhood is the result of choices made by families and the recipient schools. Parents choose which school to apply to for the education of their children studies, and school authorities choose the pupils they are willing to receive, based on their preferences and certain policy restrictions. These choices may be determined by observed and unobserved characteristics of all participants.[4]. Students are segmented in a first stage when choosing between private and public schools. The ability to pay fees, in a context of scarce scholarships, clearly contributes to the disparity in incomes between public and private schools. As of 2001, in urban Argentina 65% of children in the upper quantile of the income distribution attended private establishments, while only 7% of children in the lowest quantile did so.[5]

Within the pool of public schools, there is still some degree of segregation of students. The 1994 Federal Education Law enshrines the principle of parental school choice, meaning that children are allowed to apply to any school, irrespective of their place of residence. According to the law, the selection of pupils should be on a first-come-first-serve basis, with priority given to children with siblings or parents in the school. The new regulation represent a change from the previous rules whereby children were usually given priority by residence.[6] There was, therefore, in theory and in practice a geographical segmentation given by the existent polarization of the neighbourhoods. Once (and if) the new regulation is in place, any child is free to attend any school irrespective of the distance from the place of residence.[7] In practice, however, parents still are more likely to send their children to the neighbourhood school, and school authorities do manage to apply some level of discretion in the selection of students on non-reasonable basis (Fiszbein 1999).[8] Together, they lead to a particular composition of students according to their socioeconomic status and school segmentation within the state system.

Private schools, on the other hand, have several various ways of selecting students. Family background, psychometric tests, recommendation, or interviews are among the common criteria. Naturally, the ability to pay the fee as a requirement is in itself enough

[4]As I explain in section 4, one should be careful in dealing with the issue of selection of schools and neighbourhoods as, if ignored, they could bias the results on peer-group and neighbourhood effects. The strategy followed in this paper, explained in detailed in 4 involves controlling for observed time-varying family and school characteristics, school fixed effects for time-invariant unobservable components, and making use of the almost random assignment of students across classes within a given school

[5]Data from the *Encuesta de Calidad de Vida* carried out by the National Ministry of Social Development (SIEMPRO) in October 2001. The sample is representative of the whole country urban areas - above 5,000 inhabitants.

[6]Given that Argentine provinces maintain a high degree of autonomy under the Constitution, the 1994 Federal Education Law is not binding for the provincial governments, but rather can be interpreted as a 'recommendation' from the central government. In some provinces the parental choice system was in place before 1994, while other provinces did change the admission system in response to the law, while still even others, notably the City of Buenos Aires, continue to selects pupils according to neighbourhood of residence.

[7]Children wearing state school uniform (*guardapolvos*) are allowed to travel for free in any public transport.

[8]Fiszbein argues that there is evidence of the system being 'far from transparent' and that some families are able to circumvent regulations and either have priority in the preferred school in the district or even be able to enrol the children in schools in other districts

to determined a relatively homogeneous composition of students according to their family background.

In sum, there are three main potential sources of inequality between schools. First, public schools tend to be on average poorer than private schools. Second, provinces may differ in the amount of funding per student given to schools, while teacher quality or voluntary contributions to the school cooperative may differ across schools within and across provinces. Finally, parents may differ in their ability to 'sort' between schools, either through moving to an area with better schools or, where the new 1994 Law is in place, by choosing the school within the neighbourhood.

2.2 The Data: 'Operativo Nacional de Evaluación Educativa'

The Operativo Nacional de Evaluación Educativa (ONE) is an evaluation mechanism set up by the Argentine Ministry of Education in 1993.[9] It administers standardized achievement tests in Mathematics and Spanish at different levels of the educational system, two in the primary school level and two in the secondary school level. Tests are multiple-choice, and build on basic knowledge and capacities previously agreed among all provincial offices. Since it inception, the evaluation has been carried out annually on all students of a randomly selected stratified sample of schools.[10] The sample changes every year but there is significant overlap in the sample between years, and in 2000 all schools and all students were tested. Once the test is completed, students are asked about their personal characteristics (gender, age, educational history) and their family background. In particular, children above a certain age are asked about the family possession of more than a dozen assets, access to basic infrastructure services and the level of education of parents.[11]

I restrict my analysis to children from the sixth grade of primary school (10 to 11 years' old approximately). The rounds used are 1997, 1999, 2000, as these are the only years publicly available. Schools surveyed in more than one year are matched using a unique school identifier; neighbourhoods are defined using a municipality identifier.[12]

In the estimation section I use two different samples of the total data available. Table 1 presents the number of municipalities, schools, and classes selected in each subsample. The first subset, 'full sample', is a cross-section of schools that belong to a neighbourhood where more than one school is surveyed in any one year. Presumably, this will leave me with relatively large municipalities thus my results will be representative of larger areas. The reason for this exclusion is that I need to be able to differentiate

[9] Argentina participated in the PISA project in 2000 and 2006, but given the characteristics of that data, I chose to use the national survey. First, it was impossible to recover the location of schools at a sufficiently high level of disaggregation (necessary for the neighbourhood effects estimation). Second, given the age population targeted by both surveys, I believe it is more appropriate to use the nation survey, covering children in primary education (10 years-old) instead of PISA which assesses achievements for 15 year-old students. In Argentina, on average more than 14% of all children of that age are not enrolled in formal education and in some regions it is a high as 32%, whereas only 1% for 10 years-old children do not attend primary schools.

[10] The sample is stratified at two levels (province, urban/rural, state/private) and conglomerate (section within the school). All pupils in the class present at the day of the evaluation have to do the test

[11] Previous relevant works on using ONE include a series of reports published by the Ministry of Education (MCE 1996, MCE 1997) use the results from the 3rd grade students to show that while poor students tend to have lower test scores, those attending classes with higher concentration of poor students have even lower scores on average. Other studies using ONE, though not related to school segregation, are (Berlinski, Galiani & Gertler 2006, Herrera-Gomez 2005).

[12] Argentina is divided administratively into 24 provinces and more than 4,000 municipalities. Only one province Neuquén is excluded from the sample because it decided not to participate in the program

between school effects and neighbourhood effects. This sample also excludes schools that have only one class surveyed in grade six. This exclusion will allow me to identify peer-group effects in a convincing way. The resulting sample includes more than 700 municipalities, nearly 9,700 schools and, 21,000 classes. The second subset, 'school panel', is an unbalanced panel of schools, thus it excludes schools that were surveyed only in one of the three years. Given that the sampling of schools is random, I do not expect to have any particular bias due to this exclusion. As in the previous sample, I also exclude schools in municipalities where only one school is surveyed in any one year and schools with one class surveyed. Together, these exclusion criteria leave a total of 4,218 classes, in 1,529 schools (11% of the total in the country) and 205 municipalities.

Table 1: Sample and sub-samples. Number of municipalities, schools and classes.

Samples	Municipalities		Schools		Classes	
Full sample						
All available data	3,754	1.00	14,091	1.00	27,011	1.00
- with #school> 1 per neigh.	1,315	0.35	11,531	0.82	23,389	0.87
- & with #classes> 1 per school	716	0.20	9,695	0.69	21,109	0.78
Panel of schools						
Interviewed at least twice ...	1,125	0.30	2,674	0.19	6,203	0.23
- with #school> 1 per neigh.	256	0.07	1,698	0.12	4,340	0.16
- & with #classes> 1 per school	205	0.05	1,529	0.11	4,218	0.16

2.2.1 Test scores and Socioeconomic status

Pupils' achievement is obtained from test scores in Spanish as the percentage of correct answers to multiple choice questions.[13] In the empirical part of the paper, I use year dummy variables to capture variability in the level of difficulty at different years. All students present at the day of the survey in the chosen class were tested.[14] Of tested pupils, all have either results for Math or Spanish, and approximately 7% of the students have missing values in one of the tests.

To measure pupils' socioeconomic background, I construct two variables: the average parental education attainments and an index of capital good held at home. Pupils are asked about the highest level of education that each their parents has achieved, classified in seven categories. When two parents are present, I use the average of both.[15]

[13] I use the raw test scores. Because my regression estimates control for time effects with year dummies, there is no need for any adjustment for grade inflation or differences in difficulty across years. Official sources suggest the use of Item Response Model - Rasch model - but in the context of the present work this is not necessary.

[14] Unfortunately, it is not possible to know the proportion of absentees on that day in each school, so there is a possible bias if, for instance, teachers suggested students of lower ability not to come on the day of the test. In 2003, the proportion of absentees to the tests was on average 87% and 84% for Spanish and Maths, respectively (DINIECE 2003).

[15] Categories are incomplete primary, complete primary, incomplete secondary, complete secondary, incomplete tertiary, complete tertiary, and don't know. Values assigned to each category are between 1 and 6, with 1 representing incomplete primary and 6 complete tertiary.
A note of caution is in order: the proportion of non-response in this variable is approximately 16%.

The capital good index is formed using the pupils' responses to questions about nineteen household assets and access to various services, weighted using relative price indices, as provided by the National Bureau of Statistics (see Table 7 in the Statistical Appendix).

$$CGI_i = \sum_{j}^{K} p_j \cdot C_{ji},$$

where CGI_i is the index of capital goods for individual i, p_j is the price of good j, C_{ji} is an indicator variable of good j for individual i, and K is the number of goods. Capital goods included are boiler, fridge, freezer, gas cooker, fan, microwave oven, boiler or thermotank, washing machine, fan, air conditioner, car, drying machine, computer, audio equipment, car, telephone, colour TV, VHS player, music equipment, video camcorder, cable TV, and computer while services are electricity and water and sanitary. The index ranges from 0 to 43. Price of the cooker is used as the benchmark. Due to data limitations, I use constant relative prices for all years, thus assuming that these have not changed in the period under analysis.

2.2.2 Descriptive statistics

Table 2 presents summary information on test scores, parental education and assets at the individual, school and municipality level, distinguishing between the public and the private sector. The average Spanish test score of the pupils of the sixth grade is 60 with a standard deviation of 19. The average level of parental education is 3.5, that is, between incomplete and complete secondary school (Statistical Appendix, Figure 4). 17% of pupils in the sample attend private institutions. On average these children obtain 10 points higher test scores and come from more educated families (4.43 versus 3.37). As expected, private schools have a more homogenous set of students (lower standard deviation of parental education and asset holdings) and a significantly lower proportion of poor children. Additionally, private schools are located in, on average, richer municipalities with a lower proportion of poor children.[16]

3 Theoretical Background

This section builds a tighter, more micro-founded link between the different theories of social interactions and the empirical research. The empirical literature on social interaction and school performance can be broadly subdivided into two groups: papers that focus on the interaction between students within the school, on the one hand; and those that emphasize the environment where the children live, on the other. Loosely speaking, the first approach concentrates on 'peer-group effects' while the second on 'neighbourhood effects'. For a given level of inequality in the society, the distribution of children across schools on economic grounds (i.e. the level of segregation) will determine both the social composition within the school and the distance to those outside the school. Therefore, economic inequality will affect academic achievement at school through the impact of peers in the classroom and others in the neighbourhood, through the degree of segregation across schools.

[16] All these differences are statistically significant at a 1% level

Table 2: Summary Statistics

	Total		Public		Private	
n. students	105,543		87,747		17,796	
%	100		83.1		16.9	
n. schools	3,479		2,871		608	
%	100		82.5		17.5	
	Total		Public		Private	
variable	mean	sd	mean	sd	mean	sd
test score Spanish	60.10	19.35	58.11	19.02	69.91	17.94
parents education	3.55	1.64	3.37	1.60	4.43	1.52
class mean edu	3.55	0.86	3.37	0.75	4.43	0.83
class st.d edu	1.40	0.24	1.43	0.23	1.26	0.28
class cv edu	0.42	0.11	0.44	0.10	0.30	0.12
class poverty rate	0.14	0.13	0.16	0.13	0.05	0.09
neigh mean edu	3.57	0.46	3.53	0.46	3.72	0.47
neigh st.d edu	1.58	0.12	1.58	0.12	1.58	0.11
neigh poverty rate	0.14	0.08	0.14	0.08	0.12	0.06
segregation index	0.29	0.21	0.28	0.20	0.35	0.24
assets	26.43	13.60	25.13	13.51	32.83	12.18
class mean assets	26.43	6.82	25.13	6.27	32.83	5.76
class st.d assets	11.83	1.82	12.06	1.59	10.67	2.35
class cv assets	0.49	0.17	0.51	0.16	0.35	0.15
school poverty rate	0.17	0.15	0.19	0.15	0.08	0.11
neigh mean assets	26.51	4.36	26.29	4.43	27.58	3.79
neigh st.d assets	12.96	0.87	12.95	0.87	12.97	0.86
neigh poverty rate	0.17	0.09	0.17	0.1	0.15	0.07
segregation index	0.21	0.16	0.20	0.15	0.25	0.18

Source: Operativo Nacional Educativo, Ministerio de Education y Cultura, Argentina

This section explains the models of social interactions and builds a framework to study first, those within the school, and second within a neighbourhood. Then, the paper explores the implications that these ideas may have for the analysis of economic segregation within and across schools.

3.1 Models of social interactions inside the school

The term peer-group effects to refer to externalities or spillovers between students within a classroom that affect student performance. In analyzing these effects, Manski (1993) provides a useful distinction between *'endogenous interactions'* and *'contextual interactions'*. The former category refers to the effect of classmates' behavior. Positive interactions can lead to higher quality of discussion in the class, and faster or better development of children's knowledge. Conversely, disruptive behavior may have a negative impact on other students' capacity to learn. The latter category refers to the effect of classmates' pre-existing personal characteristics, rather than their behavior in the classroom *per se*. Finally, Manski (2000) also delineates a third category of, potentially illusory, peer effects. These 'correlated effects' refer to classmates' shared - and possibly unobserved - characteristics that are related to similar outcomes. If not properly accounted for, such correlated factors can lead to overestimation of the effect of peer interactions. The challenge in empirical work, therefore, is to be able to identify endogenous or contextual interactions when correlation effects are also present.

This paper will focus on the impact of **contextual interactions** with respect to students' family background, which *a priori* can be positive or negative, while controlling for endogenous interactions.

Models of contextual interactions generally work through the educational aspirations of children which influence student outcomes such as academic achievement (Stewart, Stewart & Simons 2007). Aspirations, in turn, are largely determined by parents characteristics (such as the level of education achieved, incomes, and professional activities) and to a lesser extent, by the child's environment outside the family. In this paper, the focus is on the impact that others in the school (classmates) and outside the school (in the neighbourhood) may have on a student achievement. There are two main models suggesting how other people's characteristics affect aspirations.

3.1.1 Role model idea

The first model argues that being among children from higher socioeconomic background can have a positive influence on school achievement. If, for instance, one student has access to extra tutorials at home from her parents, other classmates may also directly or indirectly benefit from these advantages. Even if such a direct mechanism is not at work, Wilson (1997) argues that being surrounded by other children from better-educated and richer families might change the student's behaviour because the classmates' parents represent additional role models that can change her own aspirations. [17]

[17] Wilson's work focused on the effects at the level of neighbourhoods, rather than at the school level. Still, it can be argued that the same mechanisms are present and are even stronger within the school than outside. While children might be influenced by what they observe in the streets, most of the social interactions with others indeed occur within the context of the school. In this paper, both within school and within neighbourhood effects will be included. Wilson and others identified five distinct mechanisms through which neighbourhood characteristics affect educational outcomes (Ainsworth 2002). For simplicity, I focus on the main one, known in the sociological literature as collective socialization, that works

In other words, there is a positive externality from being in a class with richer kids, that works through the children's aspirations and expectations of future prospects. Formally, if A_i is the achievement of student i, and (x_i, x_j) represent the socioeconomic status of students i and j that share the classroom, the role model idea suggests that

$$\begin{cases} \gamma_P = \frac{\partial A_i}{\partial x_j} \geq 0 & \text{for } x_i \leq x_j; \\ \gamma_R = \frac{\partial A_i}{\partial x_j} = 0, & \text{for } x_i > x_j. \end{cases}$$

One can also allow for richer students to be affected by the lower socioeconomic level of her peers. In that case, their performance is lower than it would be in the absence of the poorer classmates but, just as for poorer children, that increases in the peers' social position benefits the relatively rich students. However, the magnitude of the effect does not necessarily be the same as for those that fall below the classfellow. In other words, $\gamma_R = \frac{\partial A_i}{\partial x_j} \geq 0$, for $x_i > x_j$, where γ_R need not be the same as γ_P.

3.1.2 Competition idea

In principle, contextual peer-effects may also work in reverse. Cutler and Glaeser (1997) propose an alternative view where children's performance is driven by their feeling of competition with others. A student that is richer than another one has an additional motivation to achieve higher levels than in the absence of the worse off kid. As the worse off kid becomes richer, however, the contextual effects wears off up to the point were both have the same socioeconomic status and there is no extra effect. In other words, the higher the socioeconomic status of a classmate relative to that of the child, the lower the child's performance at school.[18] Let us first assume that the competition works only for those that are richer than the classmates. In that case,

$$\begin{cases} \gamma'_P = \frac{\partial A_i}{\partial x_j} = 0 & \text{for } x_i \leq x_j; \\ \gamma'_R = \frac{\partial A_i}{\partial x_j} \leq 0, & \text{for } x_i > x_j. \end{cases}$$

Figure 1 represents both ideas in the $(A_i, x_j - x_i)$ space, showing how student's i achievement at school varies as the socioeconomic status of classmate j increases. The left-hand-side graph shows the role model idea, while the right-hand-side the competition hypothesis.

In the role model hypothesis, when i is richer than j there is no effect from being in the same classroom. As x_j increases, student i becomes less rich up to a point where she is poorer than her schoolfellow. From then on, there is a positive effect from further increases in j socioeconomic status. On the contrary, according to the competition model as the classmate's social position improves, the benefits from being surrounded is positive but decreasing for richer students, and null for those that fall below x_j.

through the shaping of role models children are exposed to outside their home. The other mechanisms are social control (adult supervision), social networks, perceptions of occupational opportunities, and institutional characteristics (such as school quality), and have a similar positive effect. In Ainsworth (2002) the different mechanisms are tested using US data. In short, the author finds that the role model shaping (educational expectations) and school processes (institutional characteristics) are the most important 'mediators' of neighbourhood characteristics on time spent on homework, while also the others were important in determining test scores -though their effect is smaller.

[18]It is fair to say that in studies of neighbourhood effects there is fewer support to this hypothesis compared to the role model hypothesis. Exceptions include Ginther, Haveman amd Wolfe (2000) and Duncan (1994).

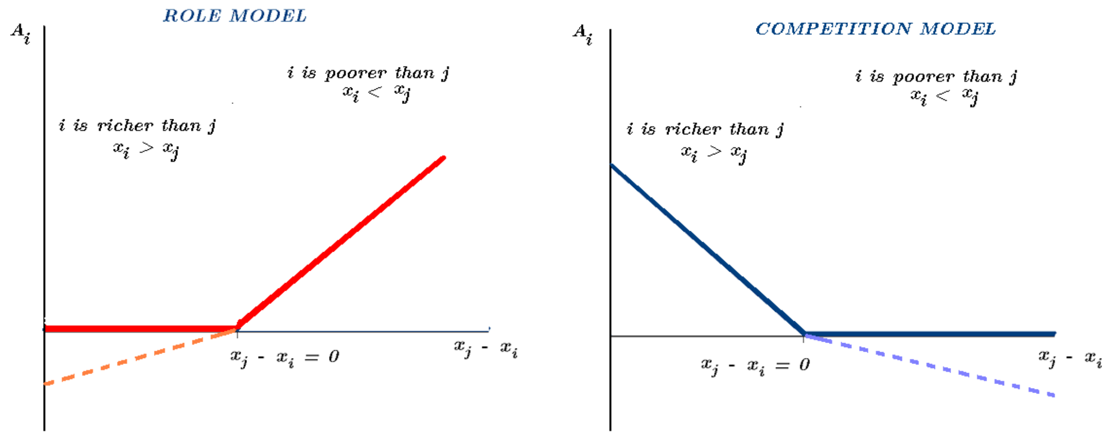


Figure 1: Models of contextual effects

It is also possible to think that the competition hypothesis produces an externality to both those above and those below the schoolfellow's socioeconomic background. Cutler and Glaeser (1997) argue that exposure to more affluent classmates may reinforce social stigmas on the poor which undermine the performance of already disadvantaged students (Ginther, Haveman & Wolfe 2000, see also). In this case, the relationship between the socioeconomic status of peers and the pupil's own performance is negative, for both poor and rich students, though the effect need not be of the same magnitude.

$$\begin{cases} \gamma'_P = \frac{\partial A_i}{\partial x_j} \leq 0 & \text{for } x_i \leq x_j; \\ \gamma'_R = \frac{\partial A_i}{\partial x_j} \leq 0, & \text{for } x_i > x_j. \end{cases}$$

where γ'_P and γ'_R need not to be the same. The competition graph will thus be represented by a everywhere decreasing line, with a possible break at the point where $x_j = x_i$.

Both models suggest that a child's performance will be affected, among other things, by how far away her income is from each of the other students, in a non-linear way. Let $g(x_j/x_i)$ be the contextual effect of the j classmate's background (x_j) on ith performance given her on socioeconomic position x_i, then a mathematical approximation of the effect represented in figure 1 can written as follows:

$$g(x_j/x_i) = \gamma_1(x_j - x_i) + \gamma_2(x_j - x_i)^2 \tag{1}$$

Each student shares the classroom with a number (n) other children. Thus, the expected effect of classmates' characteristics on student i test score can be expressed as:

$$\begin{aligned} E\left[g(x_j/x_i)\right] &= E\left[\gamma_1(x_j - x_i) + \gamma_2(x_j - x_i)^2\right] && (2) \\ &= \gamma_1 E\left[(x_j - x_i)\right] + \gamma_2 E\left[(x_j - x_i)^2\right], && (3) \end{aligned}$$

with $E[.]$ being the expected value of a variable. An *increasing and convex* relationship between her test score and the absolute difference of socioeconomic background (that is, $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$) will support the aspiration notion, while an *decreasing and* *convex* ($\gamma_1 \leq 0$ and $\gamma_2 \geq 0$) will be consistent with the relative deprivation model. (If competition is at work for both poor and rich children, the relation need not to be convex -linear and concave are also possible.) Therefore, the testing of the two models boils down to testing the coefficient on the squared differences γ_2.

One could also think that both models are, to some extent, at work for each child. Poorer students benefit from the interaction with richer classmates because it affects their aspirations and expectation of future social status, but at the same time, this effect is reduced (or even canceled out) by the feeling of being relatively deprived. Therefore, the estimated coefficient will represent in truth the net gain or loss from both effects.

Note that with little rearrangement, expression 3 can be written in terms of the mean, variance class characteristics and an interaction term:

$$E\left[g(x_j/x_i)\right] = -\gamma_1 x_i + \gamma_2 x_i^2 + \gamma_1 \bar{x_j} + \gamma_2 \bar{x_j}^2 + \gamma_2 \sigma_j^2 + \gamma_2 \bar{x_j} x_i. \quad (4)$$

where

$\bar{x}_j$ is the school average socioeconomic background of classmates (excluding student i);

σ_j^2 is the school variance of socioeconomic background in the class (excluding student i).

This expression allows us to compare more easily the standard approach to peer-group contextual effect and the one propose in this paper. Traditionally, externalities from social interactions are tested using a school production function which includes the average characteristic of peers, such as wealth, family background, etc. This approach has some econometric difficulties - discussed in the estimation section - an, I argue, a structural difficulty. In particular, they do not generally account for (a) non-linearity of the effect, and (b) different compositions of the classmates that lead to the same average, i.e. different degrees of dispersion.[19] If the relationship between peer-group characteristics and educational outcomes is non-linear –as suggested above– average outcomes will hinge on the distribution of children across schools in a given neighbourhood. The introduction of dispersion of economic backgrounds within a school, on the other hand, will be useful in capturing the idea that the behaviour of the student could depend of his or her distance to each member in the group, rather than simply its average.[20]

3.2 Neighbourhood effects

Over and above the effects from classmates, children performance may be also be influenced by social interactions outside the school. Indeed, certain characteristics of neighbourhoods have been found to affect children's educational outcomes (Ginther et al. 2000). The ideas behind these effects are essentially the same as those presented

[19] Examples of studies that did include non-linearities in their estimations are Zimmerman (2003), Winston and Zimmerman (2003), and Sacerdote (2001). Both works explore the existence of non-linearity of the 'endogenous interaction', rather than the contextual one, in higher education. In both cases, the evidence though weak suggests the existence of non-linearities of the effect of peers' past behaviour on students' current performance.

[20] Akerlof (1997) presents an int eresting model of social interactions based on social distances which predicts both negative and positive externalities from interaction. In this framework, agents make their choices (for instance, demand for education) in order to seek status and distance themselves from others in their social group (status seeking model), or to move closer to the group they belong to (conformist model), or a mixture of the two depending on their social distance from other members of society (social distance model). Which of these alternatives prevails will depend on how 'close' or 'distant' the person is from the different agents in the group. While his model is one on endogeneous effects (instead of contextual), I find it illuminating in terms of how the effect of others might not be constant (linear) across the spectrum and that social distances may lie at the heart of it.

before. In fact, notions of role models and competition are presented supporting neighbourhood effects, rather than within school contextual effects. On the one hand, choices and values of neighbours can influence the aspirations, motivations, and attitudes of the youth through their socialization and the building of role models (Ginther et al. 2000).[21] The prediction is thus that 'bad' neighbourhoods (high poverty, unemployment and delinquency) tend to lead to lower educational outcomes for youth. An alternative approach argues that in effect, good neighbourhoods can have a detrimental impact on children, especially those who are poor or belong to a minority group. In this view, competition with the rest (non minority non-poor) can lead to social pressure, hostility, and perceptions of failure associated with the minority group, which discourage children from worse-off backgrounds. These approaches are usually referred to as *competition* (Cutler & Glaeser 1997) or *relative deprivation* theories whereby the comparison with more advantaged schoolmates leads to worse performance.

The validity of these alternative hypotheses have been tested using most commonly variables such as proportion of low (high) income families, composition of neighbourhood in terms of race or ethnic groupings, unemployment rate, violence and crime rates, and levels of high school dropouts. According to Ginther *et al* (2000), most studies find evidence consistent with the first line of argument, though competition or relative deprivation hypotheses are also found significant. Because of the specific focus of the present paper on the effect of distribution of socioeconomic status on educational outcomes, neighbourhood average and dispersion of family background and assets are used in the model.[22] Therefore, the effect of neighbourhood l on student's i achievement in school k is expressed as:

$$h(x_l) = \delta_1 \bar{x}_l + \delta_2 \sigma_l \tag{5}$$

where

$\bar{x}_l$ is the average socioeconomic background in neighbourhood l ;
σ_l is the dispersion of socioeconomic background in neighbourhood l.

In sum, the level of economic inequality can affect children's educational outcomes through two distinct channels. First, through the allocation of children across schools, inequality determines both the average and the dispersion within a school, and hence the characteristics of the peer-group effect; second, directly, inequality can be seen directly as one characteristic of the neighbourhood determining the role models, motivation and aspirations of the youth. Therefore, it is economic segregation which lies behind the relation between inequality and school achievement

[21] Wilson (1997, 1996) and others have described five interrelated mechanisms through which neighborhood characteristics affect educational achievement: collective socialization (role models), social control (monitoring and sanctioning deviant behaviour), social capital, differential occupational opportunity, and institutional (i.e., school) characteristics (Ainsworth 2002). See also Ginther *et al.* (2000) for an excellent literature review.

[22] Caution with results is required. When estimating neighbourhood effects, the problem of identification appears once more. Most importantly, if unobserved family characteristics are related to both choice of neighbourhood and children's educational outcomes, the estimation of neighbourhood effects could be over or under-estimated. I deal with this problem as best as possible using a school-fixed effect estimation, and assuming that the unobservable variables that determine the selection of school work in a similar way when selecting the neighbourhood, over and above the effect of individual and family observable factors included in the model.

3.3 Economic Segregation

"At a general level, segregation is the degree to which two or more groups are separated from each other" (Allen & Vignoles 2006, p. 3). Much of the literature on school segregation refers to the distribution of minority pupils across schools, where 'minorities' are defined in terms of a few number of categories. Examples of minority groups considered are those defined according to race, ethnic or religious affiliation (Allen & Vignoles 2006, Cutler & Glaeser 1997, Gorard & Taylor 2002, **?**, Taylor, Gorard & Fitz 2000).

In this paper I use the concept of segregation in terms of social classes determined by socioeconomic background, that is, *economic segregation*. My focus on economic segregation is motivated by the particular social context of Argentina. In societies where race or religion are not distinctive elements of individual identity (either because the vast majority of the population people share the same language, race and faith - such as in Argentina - or because there is a high degree of integration on those grounds), social classes can emerge as a salient element that generates a sense of belonging that divides the society.

In this context, a high degree of economic segregation across schools implies that poor children do not share the classroom with children from a relatively better-off background. If segregation is low, students are mixed within the schools. Therefore, the differences in socioeconomic status between schools are small relative relative to the differences within the school. A rather natural index of the degree of segregation is defined as the ratio of between-school dispersion to within-school dispersion:[23]

$$S_l = \frac{\sum_{k=1}^{K}(\bar{x}_{kl} - \bar{x}_l)^2}{\sum_{i=1}^{N}(\bar{x}_{ikl} - \bar{x}_{kl})^2},$$

where subscripts i, k, l refer to individuals, schools, and neighbourhood, respectively, x represents the socioeconomic status, and $\bar{x}$ is the average. In Argentina, the level of school segregation estimated using the variance as the underlying measure of dispersion, ranges from values close to 0 to over 2. At the lowest, most of the dispersion happens within the schools; at the highest, two-thirds of the total variation is explained by the differences between school socioeconomic characteristics, while only a third corresponds to the differences across children within schools. On average, the segregation index is between 0.20 and 0.36, representing that the between-schools variation is a third of the within-school component (Table 11 and figure 5 in the statistical appendix).

The following section explains schematically how the degree of segregation is expected to affect the composition of peers and the distances to children in other schools in the neighbourhood.

3.3.1 A tale of two cities: *Mixlermo* and *Segrano*

Let us assume there are two municipalities – $l = \{Mixlermo, Segrano\}$– which have exactly the same distribution of socioeconomic characteristics (x) among families but

[23]Other measures of segregation often used include the Gorard segregation index (Gorard & Taylor 2002), the dissimilarity index, and the square root index, all of which are variations of Lorenz-consistent measures of inequality. The segregation index used in the paper is taken from the polarization literature (Zhang & Kanbur 2001). This idea of segregation has been use elsewhere to describe residential segregation by class (Jargowsky 1997), and a similar index has been proposed. For an interesting discussion on the UK literature, see Allen and Vignoles (2006).

differ in way their children are allocated across schools. In each of these neighbourhood there are N students which have on average $\bar{x}_l$ background characteristics and a standard deviation of σ_k. There are only two schools per municipality, A and B, each with $N/2$ students.[24]

Figure 1 depicts the situation in the two schools (rows) in each of the two municipalities (columns). The light gray distribution represents the underlying distribution in the neighbourhood – which is exactly the same in both cases – while the shaded area represents the distribution of students within a given school.

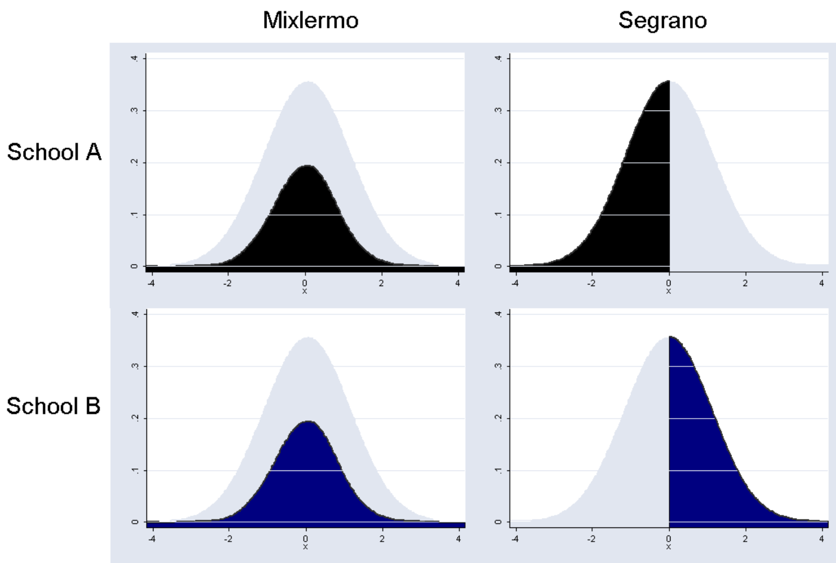


Figure 2: Tales of two Cities. Distribution of socioeconomic economic background
Note: The lighter distributions represent the distribution in the neighbourhoods, while the darker ones are those of each school

In Mixlermo there is no economic segregation across schools. Both schools receive the same proportion of poor to non-poor children, and the classrooms have a perfect mixing of children coming from different socioeconomic background. In School A_{MIX} and School B_{MIX} the average and dispersion of characteristics of the pupils are the same, and equal to the neighbourhood.

In contrast, Segrano allocates pupils across schools strictly based on their economic background. School A_{SEG} charges high fees and therefore only children coming from better-off background can access it, and the rest of the students are allocated in school $_{SEG}$. In consequence, the average characteristics of school A_{SEG} are higher than in the neighbourhood, which in turn are higher than in school B_{SEG}. On the other hand, the dispersion in both schools is lower than in the neighbourhood. The question is which of the two municipalities has higher average academic achievement.

Expression 4 presented above show how the average characteristics and the composition of classmates can affect a child's academic achievement. There are, of course, other factors that determines the performance of the student at school. These typically include other individual characteristics (such as age and gender), and school-level characteristics (including teachers quality, class size, classmates' performance, among others).

[24]I refer to 'neighbourhood' and 'municipality' interchangeably. The first is the common way to refer to effects of the environment, the second is the aggregation unit I use in the empirical section.

Hanushek (1971) was the first to formalize this into a school production function. Adding the contextual effect terms $g(x_j/x_i)$ to Hanushek's specification, the school achievement A_{ikl} of student i in school k in neighbourhood l can be represented as:

$$A_{ikl} = \alpha + \beta_1 x_{ikl} + \beta_2 Z_{ikl} + \beta_3 W_{kl} + \gamma_1 \bar{x}_{kl} + \gamma_2 \bar{x}_{kl}^2 + \gamma_3 \sigma_{kl}, \tag{6}$$

where
$i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, J$, $k = 1, 2, \ldots, K$ and, $l = 1, 2, \ldots, L$;

x_{ikl} is the individual i socioeconomic background who attends school k in neighbourhood l;

Z_{ikl} is the vector of other personal characteristics of individual i;

W_{kl} is the vector of school-level variables in neighbourhood k;

$\bar{x}_{kl}$ is the school average socioeconomic background of students;

σ_{kl} is the school dispersion of socioeconomic background of students.

To keep the argument simple, the school characteristics of all four schools are assumed to be equal ($W_{k,MIX} = W_{k,SEG}$), including materials, pedagogic technology, and quality of the teachers. Thus, the average academic performance $E[A_{ijkl}]$ will depend, over and above children's personal characteristics (family and individual), on the existence of any effect from the composition of the classes, i.e. the average peer-group effect (γ_1, γ_2) and the dispersion peer-group effect (γ_3).

In the case of the former, the shape of the relationship between the mean $\bar{x}_{kl}$ and individual school achievement A_{ijkl} will determine which of the two neighbourhoods will perform better on average (Figure 2). If the relationship is linear, school segregation does not matter, i.e. $E[A_{MIX}] \equiv E[A_{SEG}]$. If the relationship is concave, the lower the segregation the higher the average, i.e. $E[A_{MIX}] \geq E[A_{SEG}]$. On the contrary, if it is convex the higher the segregation, the higher the average performance, $E[A_{MIX}] \leq E[A_{SEG}]$. The implication is that only when the relation between peer-group variables and individual achievement is linear will the allocation of children across schools not affect average performance. In all other cases, economic segregation will play a role in the final outcome. Most previous students by ignoring the possible curvature of this relationship ignore this specific impact of economic segregation across schools.

The second peer-group effect is the dispersion of socioeconomic background within the school. There are reasons to believe that the behaviour of a child when in a group will depend on his or her (social) distance from the other members of the group, or from those that he or she takes a reference - this could be the majority or a subgroup that shares some other characteristic such as religion or race. The degree of dispersion will represent an aggregate measure of distance, in a similar spirit as Akerlof's model of social distance when distance is considered in a quadratic form. If higher heterogeneity in the classroom leads to better performance, the mixed school system will deliver a higher average academic achievement; while the opposite will be true if the relationship is negative.

The overall effect of segregation will depend on the presence of both average and dispersion peer-group effects, their signs and relative magnitudes. In particular, lower segregation will lead to higher average outcomes if and only if the following condition holds:

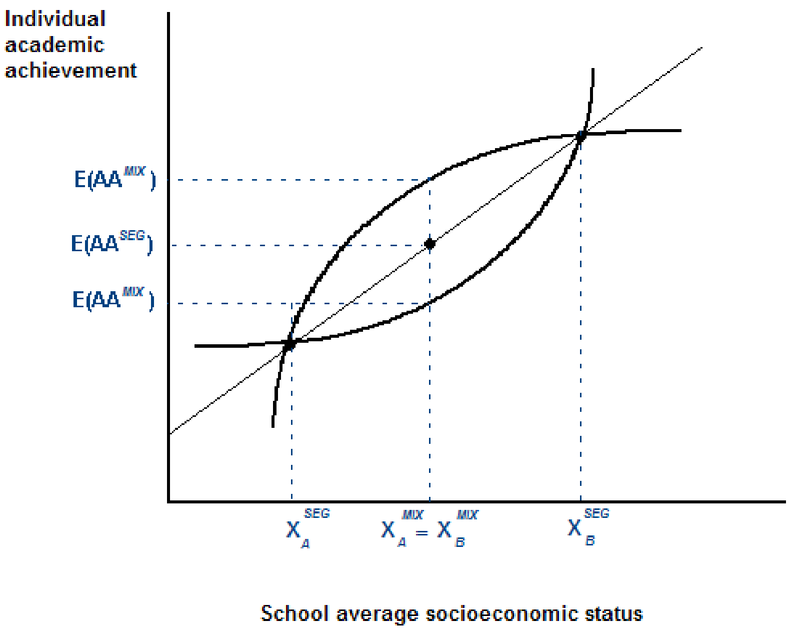


Figure 3: Average socioeconomic level in the school on individual school achievement (average peer-group effect)

Condition I

$$E[A_{ijAMIX}] \geq E[A_{ijASEG}] \Leftrightarrow \frac{\gamma_2}{\gamma_3} \geq -\frac{[\sigma_{MIX} - \frac{1}{2}(\sigma_{A,SEG} + \sigma_{B,SEG})]}{[\bar{x}^2_{MIX} - \frac{1}{2}(\bar{x}^2_{A,SEG} + \bar{x}^2_{B,SEG})]}. \quad (7)$$

Table 3 summarises the results from each of the alternative situations presented above.

Table 3: Average and dispersion peer-group effects. Condition I

		Dispersion peer-group effect $\gamma_3 > 0$	$\gamma_3 < 0$
Mean peer-group effect	concave $\gamma_1 > 0$ $\gamma_2 < 0$	$MIX \succ SEG$	$SEG \succ MIX \Leftrightarrow \frac{\gamma_2}{\gamma_3} < k$
	linear $\gamma_1 > 0$ $\gamma_2 = 0$	$MIX \sim SEG$	$SEG \sim MIX$
	convex $\gamma_1 > 0$ $\gamma_2 > 0$	$MIX \succ SEG \Leftrightarrow \frac{\gamma_2}{\gamma_3} > k$	$SEG \succ MIX$

where $k = -\frac{[\sigma_{MIX} - \frac{1}{2}(\sigma_{A,SEG} + \sigma_{B,SEG})]}{[\bar{x}^2_{MIX} - \frac{1}{2}(\bar{x}^2_{A,SEG} + \bar{x}^2_{B,SEG})]}$.

25

In four of the six possible combinations the conclusion is unambiguous, in the other two whether mixing or segregation is preferred to achieve higher average neighbourhood achievement will depend on the satisfaction of condition I which depends on the actual levels of dispersion and concavity in the two neighbourhoods (proof in the Appendix). Notice that assuming linearity of the peer-group effects, as in the standard analysis, leads invariably to unambiguous conclusions whereby, whenever the effect is positive, the ranking of neighbourhoods on average achievement grounds, will depend on how the dispersion effect plays.

3.3.2 And a third one: *Richleta*

Let us assume there is another municipality – *Richleta*– which, on average, is richer than Mixlermo but school A has similar students than in Mixlermo, while school B has richer students, i.e. a shift of the distribution. Figure 3 compares the distribution of socioeconomic status in both neighbourhoods.

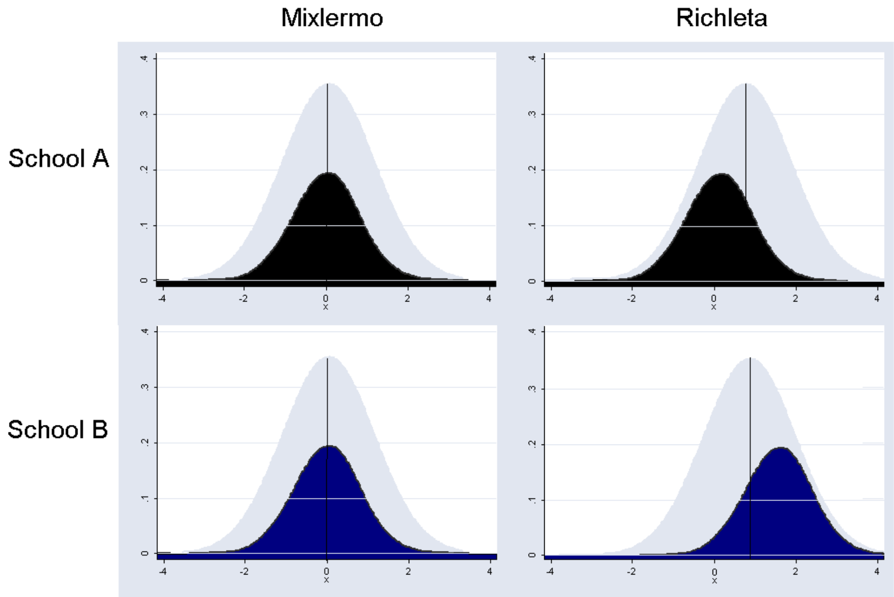


Figure 4: Mixlermo and Richleta. Distribution of socioeconomic economic background
Note: The lighter distributions represent the distribution in the neighbourhoods, while the darker ones are those of each school

Considering only the effects within the school, other things being equal, Richleta will perform on average better than Mixlermo because the mean in school B in the first neighbourhood in higher while the rest of the parameters are the same. The interesting question is though whether children in school A perform better in Richleta or in Mixlermo. In other words, is there any effect from the distribution outside the school?

According to the literature the neighbourhood effect can take two possible directions. On the one hand, the richer and less unequal the neighbourhood, the better the

[25] In the case of this example, where the neighbourhood distribution is normal, one could compute the value of k for a standardized normal; $k = 0.7$

social conditions, the roles models and aspirations of children, the higher their performance at school. Thus, school A in Richleta, being in a richer neighbourhood will perform on average better than school A in Mixlermo. On the other hand, as people compare themselves to others around them (i.e. neighbourhood) students in school A in Richleta –on average poorer than the rest– might feel relatively deprived, hence reducing their expectations on returns to education and in consequence perform worse than otherwise.

Including the neighbourhood effect term $h(.)$ of expression 5, the individual's school production function has the following form:

$$\begin{aligned} A_{ikl} &= \alpha + \beta_1 \bar{x}_{ikl} + \beta_2 Z_{ikl} + \beta_3 W_{kl} + \underbrace{\gamma_1 \bar{x}_{kl} + \gamma_2 \bar{x}_{kl}^2 + \gamma_3 \sigma_{kl}}_{\text{peer-group effect}} \\ &+ \underbrace{\delta_1 \bar{x}_l + \delta_2 \sigma_l}_{\text{neighbourhood effect}} \quad , \end{aligned} \tag{8}$$

where
$\bar{x}_l$ is the average socioeconomic background in neighbourhood l;
σ_l is the dispersion of socioeconomic background in neighbourhood l.

If schools A in both neighbourhoods share characteristics, including those of their students, the ranking of situations based on average achievements depend on the sign of the neighbourhood level parameters, i.e. δ_1 and δ_2. In the case described above, neighbourhood inequality is equal in both municipalities ($\sigma_M = \sigma_R$). Therefore, if richer neighbourhoods have a positive impact on the achievement of students ($\delta_1 \geq 0$), Richleta would be preferred. On the other hand, if the relative deprivation hypothesis is validated, students in school A will perform worse because they are worse-off relative to others outside the school, $\delta_1 \leq 0$, Mixlermo would be preferred. In cases where both neighbourhood mean and dispersion differ, the condition can be expressed as follows:

Condition II

$$E[A_{ijAMIX}] \geq E[A_{ijARICH}] \Leftrightarrow \frac{\delta_1}{\delta_2} \leq \frac{\sigma_{RICH} - \sigma_{MIX}}{\bar{x}_{MIX} - \bar{x}_{RICH}}. \tag{9}$$

Assume Mixlermo is poorer but less unequal, so that the right-hand side of Condition II less than zero. The question is, in terms of pupils' achievements in school A, is it better to be in a neighbourhood that is poorer but more equal or richer but unequal? Table 4 presents the different alternatives:

4 Estimation

4.1 Estimation Strategy

In this section I present a reduced form equation to test whether and how socioeconomic inequality impacts students' performance at school. The estimation model includes a direct mechanism linking inequality to student (the standard deviation of parental education within the neighbourhood, or 'neighbourhood effect') and an indirect mechanism operating via the composition of schools in the neighbourhood (the average and standard deviation of socioeconomic characteristics within the school, or 'contextual peer-

Table 4: Average and dispersion neighbourhood effects. Condition II

		Dispersion neighbourhood effect		
		$\delta_2 < 0$	$\delta_2 = 0$	$\delta_2 > 0$
Mean neighbourhood effect	$\delta_1 < 0$	$MIX \succ RICH$	$MIX \succ RICH$	$RICH \succ MIX \Leftrightarrow \frac{\delta_1}{\delta_2} \leq \frac{\sigma_R - \sigma_M}{\bar{x}_M - \bar{x}_R}$
	$\delta_1 = 0$	$MIX \succ RICH$	$MIX \sim RICH$	$RICH \succ MIX$
	$\delta_1 > 0$	$MIX \succ RICH \Leftrightarrow \frac{\delta_1}{\delta_2} \leq \frac{\sigma_R - \sigma_M}{\bar{x}_M - \bar{x}_R}$	$RICH \succ MIX$	$RICH \succ MIX$

group effect'). The latter is in itself an immediate consequence of the level of economic segregation across schools. The reduced form specification comes from equation (5.3):

$$\begin{aligned} A_{ijklt} &= \alpha + \beta_1 x_{ijklt} + \beta_2 Z_{ijklt} + \beta_3 A_{(-i)jklt} + \gamma_1 \bar{x}_{(-i)jklt} + \gamma_2 \bar{x}^2_{(-i)jklt} + \gamma_3 \sigma_{jkt} \\ &\quad + \delta_1 \bar{x}_{(-k)lt} + \delta_2 \sigma_{lt} + \eta_{ijklt}, \end{aligned} \tag{10}$$

where the subscript 't' refers to time. Z_{ijkt} includes gender, and whether the child has repeated any grade, while $A_{(-i)jklt}$ is the average class academic achievement excluding student i and is included to control for endogenous interaction effects. Variable $\bar{x}$ represents the average socioeconomic status in the reference group, proxied by the level of parental education. $\bar{x}_{(-i)jktl}$ is the average economic level in class j, excluding the economic background of student i and $\bar{x}_{(-k)lt}$ is the neighbourhood l average economic level excluding the value for the school k the student attends to. Finally, η_{ijklt} is an individual specific error term. Standard errors are clustered at the school level to adjust for intra-school correlation and are robust to heteroscedasticity.

Given the relative parsimony of the empirical model, an obvious concern is that estimates of the parameters in (5.5) will suffer from omitted variable bias due to the failure to control for numerous student, school and neighborhood characteristics. To clarify the possible sources of endogeneity, consider the decomposition of η_{ijklt} into four elements

$$\eta_{ijkt} = a_{ijkl} + b_{kl} + c_{klt} + u_{ijklt}, \tag{11}$$

where a_{ijkl} and b_{kl} represent time-invariant unobserved individual and school characteristics, respectively, while c_{klt} corresponds to school time-varying effect, and u_{ijklt} is a random error capturing individual shocks that vary over time.

Examples of unobserved time-invariant factors that will potentially bias the estimates of the peer parameters include individual innate ability and school location (captured in the a_{ijkl} and b_{kl} terms respectively). The problem stems from the possibility that parents choices of school and neighborhood are endogenous to the quality of education provided. A standard concern in the literature is that peer effects may simply reflect a tendency for eager parents to send brighter children to high-achieving schools with, say, wealthier peers and better principals.

To address such difficulties, this paper uses a standard approach, which involves using panel data fixed effects estimation, removing the time-invariant components (Hanushek et al. 2003). Due to the characteristics of the data (there is no unique student ID that permits tracking individual students across years) I am only able to control for school fixed effects. Thus, strictly speaking, I control for the b_{kl} term and not the individual a_{ijkl} effects. However, in as much as the parents' selection of school due to reputation is common across all children that attend a particular school, the school fixed effect component is able to address the most significant part of the problem. It should be clear that with the inclusion of school fixed effects, econometric identification of peer effects relies on the existence of perturbations in the school composition of students from year to year. As expected, these variations are small relative to the overall variation across schools, which makes the identification more demanding, but also more reliable. Given that the sample is perennially limited to children in the sixth grade, the families will most likely differ from year to year, except for those pupils that have to retake the grade. Therefore, the socioeconomic composition of the school or class may vary from over time for a number of reasons. For instance, a rise in the mean socioeconomic status of the school might reflect an economic shock to the community, or a transitory shift in the demographic representation of rich and poor families.[26]

However, school fixed effects cannot address all sources of bias. Apparent peer effects may still be driven by endogenous changes in the composition of students over time – for instance, as scholastically minded parents flock to an upper-class school that has recently performed well – or unmeasured changes in the quality of instruction or school management. In both cases, the estimate of peer effects will be biased due to the correlation between peer characteristics and c_{jkt}.

To address these concerns, I exploit random variation in peer-groups between sixth-grade classrooms in a given school at a given point in time to obtain robust estimates of peer effects which do not suffer from the aforementioned problem related to time-varying school characteristics. Two-thirds of the schools in Argentina have more than one class in each grade. The assignation of pupils to a class is determined at the beginning of primary school (at 6-7 years of age) and remains unchanged throughout the whole seven years of education. This assignation is done with no clear systematic procedure and, for our purpose, is completely random. My identification strategy is thus based on the differences across classes within schools, using a school/year fixed effects estimation procedure.[27] Because all c_{jkt} effects are controlled for, this procedure will produced unbiased estimates of peer group effects, even in the presence of endogenous sorting of pupils between schools or unobserved variation in school quality over time. (Unfortunately, because all classes within a school have the same value for neighbourhood-level variables, this strategy cannot be used to estimate neighbourhood-level parameters.)

[26]There are number of other ways to address the problem of selection. For some the most convincing way of dealing is to include a model of selection of schools (Ginther et al. 2000). Kingdon (1996, 2006) uses a multinomial logit model of choice of school-type for India; Ginther (2000) implements a similar procedure using US data. A second option involves using differences-in-differences estimations. In particular, if information is available for each student at multiple points in time (a panel of students) then one can use student fixed-effect estimation where the observed differences in achievement are related to the observed changes in the variables of interest (rather than levels). None of these options are feasible with the ONE data.

[27]Because the identification of effects is based on within school differences, I do not need to use a time-series panel. This allows me to employ a much larger sample, three-times the original one. Results are presented for both the panel of schools sample and the full sample. Given that in 2000 all schools were tested, a great proportion of the cases belong to that year, rendering the full sample much larger than the panel of schools.

4.2 Results

4.2.1 Average student achievement

Table 5 presents the estimates of equation (5.5) from the estimation using parental education as the measure of socioeconomic background.[28] Primary variables of interest are the average peer-group effect, the dispersion peer-group and, the neighbourhood effect, which are captured in the first five variables. Note that all regressions control for the average test score in the class. As discussed in previous sections, this is intended to control for endogenous interactions, restricting the focus to peer-effects due to background characteristics. Other controls are student's parental education, sex, whether the student repeated years previously, class size, and year dummies.

Columns 1 to 5 of Table 5 present results for a panel of schools –that is, the sample is restricted to schools which took the test in more than one year. This allows me to include school fixed effects in the regressions to control for unobserved quality differences between schools which might otherwise bias the estimates of peer effects. The school fixed effects equations (columns 1 to 3) rely on two sources of variation in socioeconomic characteristics to identify the effects of peers on student achievement: variation over time within the school and variation across classrooms. A remaining concern however, is that time series variation will be endogenous if good students endogenously sort into good schools. Columns 4 and 5 overcome this objection by including school-year fixed effects; i.e., a separate dummy variable that is unique to each school in each year. In this case, time series variation in socioeconomic characteristics is no longer used to estimate the peer effects. Rather, I rely exclusively on the quasi-random assignment of students to a given classroom (each of which with a somewhat different socioeconomic composition) to estimate the peer-group effects. Finally, columns 6 and 7 are replications of 3 and 4, respectively, but without restricting the sample to schools with repeat observations.

The main results can be summarized as follows. First, the average peer-group effect is concave with a turning point between 3 and 4, that is between incomplete and complete secondary, depending on the identification strategy. This result suggests that mixing is better than segregating students across schools. Based on column 3, classmates' parents education increases predicted pupil achievement by at least 1 point. Adding unobserved school-year effects more than doubles this effect. Parents with secondary incomplete education contribute through the peer-effect to 5 points in test scores of the mean child (column 5). This is presumably, because the child can benefit from sharing tutorials or materials with classmates, or building role models based on his classmates parents. Because this effect is positive but at a decreasing rate (concave relationship) the result on achievement is on average highest when the level of economic segregation across schools is minimum.

Second, the dispersion peer-group effect (columns 2, 5, and 7) is significantly negative, suggesting that segregation is to be preferred. In other words, a more heterogeneous class results in a lower average achievement of the school. One standard deviation increase in the dispersion within a classroom, reduces the academic achievement of the pupil by 0.6. This result works in the opposite direction to the average peer-group effect. I thus obtain an ambiguous result on the preferred arrangement of students across schools within a neighbourhood. On the one hand, because individual performance is a concave function of classmates' socioeconomic status, mixing students would maximize their av-

[28] OLS results and estimation using the asset index instead of parental education to define socioeconomic status are available upon request .

Table 5: Students' achievement and parental education.

	Outcome variable: Students' test scores in Spanish						
	Panel of schools					Full sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
class mean edu	-.427	1.108	1.109	.908	2.739	1.621	3.655
	(.456)	(.539)**	(.539)**	(.852)	(.930)***	(.432)***	(.487)***
(class mean edu)^2	.018	-.187	-.186	-.126	-.371	-.193	-.463
	(.063)	(.074)**	(.074)**	(.117)	(.126)***	(.059)***	(.066)***
class st.d edu		-1.168	-1.122		-1.542		-1.691
		(.200)***	(.201)***		(.376)***		(.182)***
neigh mean edu			-.080				
			(.218)				
neigh st.d edu			-.818				
			(.687)				
class mean test score	.636	.633	.633	.345	.341	.371	.367
	(.011)***	(.011)***	(.011)***	(.024)***	(.024)***	(.011)***	(.011)***
edu	2.194	2.115	2.116	2.193	2.140	1.625	1.541
	(.166)***	(.168)***	(.168)***	(.163)***	(.163)***	(.098)***	(.098)***
(edu)^2	-.252	-.240	-.240	-.250	-.241	-.194	-.181
	(.023)***	(.023)***	(.023)***				
male	-5.103	-5.101	-5.102	-5.132	-5.130	-5.344	-5.341
	(.126)***	(.126)***	(.126)***	(.124)***	(.124)***	(.070)***	(.070)***
repe	-5.420	-5.415	-5.416	-5.546	-5.540	-7.600	-7.590
	(.147)***	(.147)***	(.147)***	(.154)***	(.154)***	(.095)***	(.095)***
class size	.034	.035	.034	.124	.123	.122	.120
	(.013)***	(.013)***	(.013)***	(.029)***	(.029)***	(.014)***	(.014)***
year dummies	yes	yes	yes	-	-	-	-
school fixed effect	yes	yes	yes	-	-	-	-
school-year fixed effect	-	-	-	yes	yes	yes	yes
Obs.	105523	105523	105523	105523	105523	307285	307285
R^2	.133	.133	.133	.064	.064	.078	.078

erage achievement. On the other hand, because more heterogenous classrooms have on average lower performance is best to segment children on economic grounds. Which of these two effects prevails will depend on the relation between estimated parameters and levels of the specific school-level variables. From Table 3, the determining condition can be expressed as follows:

$$SEGREGATION \succ MIXING \Leftrightarrow \frac{\gamma_2}{\gamma_3} < -\frac{[\sigma_{MIX} - \frac{1}{2}(\sigma_{A,SEG} + \sigma_{B,SEG})]}{[\bar{x}^2_{MIX} - \frac{1}{2}(\bar{x}^2_{A,SEG} + \bar{x}^2_{B,SEG})]}. \quad (12)$$

I compute this condition using estimated values for γ_2 and γ_3 from columns 5 and 7 and $\bar{x}^2_{jk}$ and σ_{jk} using information of the percentile distribution of socioeconomic characteristics in the data from Table 13 in the Statistical Appendix. I use the 10^{th}, 50^{th}, and 90^{th} percentiles as measures of low, average, and high values of $\bar{x}^2_{jk}$ and σ_{jk}. Inserting theses percentile values into expression (12) yields:

$$\text{for column (5) } 0.24 < 0.28,$$
$$\text{for column (7) } 0.27 < 0.28,$$

which suggests that the net effect of the average and dispersion peer effects is such that segregation is to preferred. However, the different between the two side of the condition is not statistically different from zero. Therefore, there is no statistical evidence that economic segregation does indeed increases the academic achievement of the mean student.

Finally, the neighbourhood terms on their own (column 3) do not have any impact on individual achievement. This implies that, in Argentina, the evidence does not support either story of the neighbourhood characteristics affecting children's performance at school.

In sum, all the effects from interacting with others happen only within the school, and there is no unambiguous conclusion as to whether higher or lower economic segregation of children across schools is to be preferred unless specific values for school-level variables are imposed. Even for some reasonable values suggested here, there is statistically no unambiguous effect of segregation on the academic performance of the mean pupil. All these results are robust to log-transformation of the dependent variable and to using index of assets holding as indicator of socioeconomic status (see Tables 15 and 17 in the statistical appendix).

4.2.2 Effects on the poor

The results from the previous regressions, though suggestive, were not able to give us an unambiguous conclusion with respect to the effect of economic segregation across schools. The effects of different types of peer-group characteristics may well balance out and have no impact on the academic performance of the mean student. But, will these results hold for the worse-off children? To answer this question I include a 'poor' dummy variable, interacting with the school- and neighbourhood-level variables. I define a child to be poor if his family is among the 20% with the lowest parental education - in this case, less than complete primary education.[29] Because the poverty line is the same for all students, I define the poor using an absolute poverty.[30] I reserve the term *relative* deprivation for within-neighbourhood comparisons.

[29]For the assets' index, a poor child is one with an index below 15 in 1997 or 2000 or below 10 in 1999.

[30]It is 'relative' in that the line is defined according to the empirical distribution

Table 6: Students' achievement and parental education for poor and non-poor

	Outcome variable: Students' test scores in Spanish		
	Panel of schools		
	(1)	(2)	(3)
class mean edu	.215	1.983	1.939
	(.504)	(.585)***	(.585)***
(class mean edu)^2	-.022	-.264	-.256
	(.070)	(.080)***	(.080)***
class st.d edu		-1.539	-1.440
		(.231)***	(.232)***
neigh mean edu			-.006
			(.223)
neigh st.d edu			-1.387
			(.757)*
edu	1.861	1.810	1.813
	(.258)***	(.259)***	(.259)***
(edu)^2	-.217	-.206	-.206
	(.032)***	(.032)***	(.032)***
class mean test score	.605	.600	.600
	(.012)***	(.012)***	(.012)***
poor	-9.628	-7.257	-9.787
	(2.086)***	(2.100)***	(2.950)***
(poor)class mean edu	1.967	-3.404	-3.139
	(1.266)	(1.450)**	(1.463)**
(poor)(class mean edu)^2	-.692	-.011	-.044
	(.201)***	(.220)	(.221)
(poor)class st.d edu		4.508	4.221
		(.628)***	(.641)***
(poor)neigh mean edu			-.317
			(.375)
(poor)neigh st.d edu			2.254
			(1.451)
(poor)class mean test score	.179	.195	.195
	(.013)***	(.013)***	(.013)***
year dummies	yes	yes	yes
school fixed effect	yes	yes	yes
Obs.	105,523	105,523	105,523
R^2	.135	.136	.136

Control variables not included in the table: male, repeat, class size

Table 6 presents the results including terms interacting the relevant variables with a poor dummy ('non poor' is used as the baseline category). Estimated coefficients on control variables are not presented in the table for the sake of clarity, but are included in the estimation. Interestingly, the effect of most school-level variables is significantly different between poor and non-poor students. The positive mean peer effect found for the non-poor (and the average) is completely eliminated for poor pupils. At the same time, the dispersion peer-group effect while negative for non-poor has a significantly positive impact on the academic performance of poor students. In other words, worse-off children do indeed benefit from being in a classroom with a fairly heterogenous composition of pupils. A one-standard deviation increase in the level of heterogeneity, increases the achievement of poor by approximately 1 point. Although not large in magnitude, from a policy perspective, this result implies that low economic segregation across schools is beneficial both because all students benefit from higher average socioeconomic characteristics of classmates and because poor students, in particular, perform better at school when there is more relative variation across children. These results are corroborated by class fixed-effects estimation (Table 7) and are robust to transformation of the dependant variable and using assets index as the socioeconomic variable.[31]

5 Conclusion

This paper showed how socioeconomic inequality can impact children's performance at school, both directly and through the level of economic segregation across schools. The application to Argentine data led to the following conclusions. First, ignoring curvature of peer-group effect ignores an important mechanism through which economic segregation affects individual academic achievements. If the peer-group effect is concave, as in the case of Argentina, lower segregation is preferred. Second, classroom composition in terms of degree of heterogeneity of economic backgrounds within the school does matter for pupils' performance. Poor children benefit from the mixing, while non-poor pupils do not. This result, together with the previous supports the idea that lower economic segregation of children across schools is to be preferred, especially if the underlying social welfare function gives more weight to students in the lower part of the socioeconomic distribution. Finally, the evidence suggests that neighbourhood characteristics - in terms of average and deviation of socioeconomic status of its population - do not have an impact on achievement.

[31] Tables upon request.

Table 7: Students' achievement and parental education for poor and non-poor (classes). School-year fixed-effects estimation

	Outcome variable: Students' test scores in Spanish			
	Panel of schools		Full sample	
	(1)	(2)	(3)	(4)
edu	2.202	2.183	1.442	1.401
	(.255)***	(.254)***	(.149)***	(.149)***
$(\text{edu})^2$	-.254	-.247	-.176	-.166
	(.031)***	(.031)***	(.018)***	(.018)***
class mean test score	.321	.316	.346	.339
	(.024)***	(.024)***	(.012)***	(.012)***
class mean edu	1.392	3.372	1.943	4.174
	(.853)	(.935)***	(.443)***	(.498)***
$(\text{class mean edu})^2$	-.142	-.413	-.192	-.496
	(.117)	(.128)***	(.060)***	(.067)***
class st.d edu		-1.818		-2.010
		(.392)***		(.192)***
poor	-7.718	-5.579	-8.537	-6.421
	(2.218)***	(2.243)**	(1.452)***	(1.462)***
(poor)class mean test score	.155	.171	.162	.183
	(.014)***	(.014)***	(.008)***	(.008)***
(poor)class mean edu	1.849	-3.216	1.433	-4.069
	(1.360)	(1.544)**	(.851)*	(.961)***
$(\text{poor})(\text{class mean edu})^2$	-.664	-.022	-.551	.141
	(.213)***	(.232)	(.129)***	(.140)
(poor)class st.d edu		4.324		4.593
		(.637)***		(.370)***
male	-5.117	-5.112	-5.333	-5.329
	(.124)***	(.123)***	(.070)***	(.070)***
repeat	-5.538	-5.539	-7.585	-7.568
	(.154)***	(.154)***	(.095)***	(.095)***
class size	.122	.125	.120	.121
	(.029)***	(.029)***	(.014)***	(.014)***
school-year fixed effect	yes	yes	yes	yes
Obs.	105,523	105,523	304,541	304,541
R^2	.066	.066	.08	.08

Control variables not included in the table: male, repeat, class size

Appendix: Proof

Proposition 1. ***Condition I:*** *If students in two neighbourhoods (M and S) share mean and standard deviation of their socioeconomic characteristics and schools share their observable characteristics (such as size, quality of teachers, equipments), then the average achievement in M is higher than in S,* $E[A_M] \geq E[A_S]$*, if and only if,*

1. $\gamma_1 >$ *and* $\gamma_3 <$ *or,*
2. $sg(\gamma_2) = sg(\gamma_3)$ *and* $\frac{\gamma_2}{\gamma_3} > -\frac{[\sigma_M - \frac{1}{2}(\sigma_{AS}+\sigma_{BS})]}{[\bar{x}^2_M - \frac{1}{2}(\bar{x}^2_{AS}+\bar{x}^2_{BS})]} > 0$

Proof.

$$E[A_M] \geq E[A_S] \Leftrightarrow$$

$$\begin{aligned} \alpha + \beta_1 E[x_{ijM}] + \beta_2 E[Z_{ijM}] + \beta_3 E[W_{jM}] + & \quad \alpha + \beta_1 E[x_{ijS}] + \beta_2 E[Z_{ijS}] + \beta_3 E[W_{jS}] + \\ +\gamma_1 E[\bar{x}_{jM}] + \gamma_2 E[\bar{x}^2_{jM}] + \gamma_3 E[\sigma_{jM}] \geq & \quad +\gamma_1 E[\bar{x}_{jS}] + \gamma_2 E[\bar{x}^2_{jS}] + \gamma_3 E[\sigma_{jS}] \end{aligned}$$

$$\gamma_2 E[\bar{x}^2_{jM}] + \gamma_3 E[\sigma_{jM}] \geq \gamma_2 E[\bar{x}^2_{jS}] + \gamma_3 E[\sigma_{jS}],$$

given that the underlying characteristics of the population in both neighbourhoods are the same, $E[x_{ijS}] = E[x_{ijS}]$, $E[Z_{ijM}] = E[Z_{ijM}]$ and $E[W_{jM}] = \beta_3 E[W_{jM}]$ and schools share characteristics $E[\bar{x}_{jS}] = E[\bar{x}_{jS}]$.

Define

$$\begin{aligned} E[\bar{x}^2_{jM}] = \bar{x}^2_{jM} &= \frac{N}{2}\bar{x}^2_{AM} + \frac{N}{2}\bar{x}^2_{BM}, \\ E[\bar{x}^2_{jS}] = \bar{x}^2_{jS} &= \frac{N}{2}\bar{x}^2_{AS} + \frac{N}{2}\bar{x}^2_{BS}, \\ E[\sigma_{jM}] = \sigma_{jM} &= \frac{N}{2}\sigma_{AM} + \frac{N}{2}\sigma_{BM}, \\ E[\sigma_{jS}] = \sigma_{jS} &= \frac{N}{2}\sigma_{AS} + \frac{N}{2}\sigma_{BS}, \end{aligned}$$

then,

$$\gamma_2(\bar{x}^2_{AM} + \bar{x}^2_{BM}) + \gamma_3(\sigma_{AM} + \sigma_{BM}) \geq \gamma_2(\bar{x}^2_{AS} + \bar{x}^2_{BS}) + \gamma_3(\sigma_{AS} + \sigma_{BS}),$$

$$\gamma_2(\bar{x}^2_{AM} + \bar{x}^2_{BM} - \bar{x}^2_{AS} - \bar{x}^2_{BS}) \geq -\gamma_3(\sigma_{AM} + \sigma_{BM} - \sigma_{AS} - \sigma_{BS}),$$

$$\frac{\gamma_2}{\gamma_3} \geq -\frac{(\sigma_{AM} + \sigma_{BM} - \sigma_{AS} - \sigma_{BS})}{(\bar{x}^2_{AM} + \bar{x}^2_{BM} - \bar{x}^2_{AS} - \bar{x}^2_{BS})} = k.$$

□

Statistical Appendix

Table 8: Price Index of capital and household goods (cook = 1)

good	weight	good	weight
electricity	16.4	water and sanitary services	7.0
cook (natural gas)	1.0	fridge	0.4
fridge with freezer	1.7	microwave	0.3
boiler	0.3	thermotank	0.2
washing machine	1.7	fan	0.2
air conditioner	0.2	car	21.3
telephone	1.3	colour TV	1.7
VHS player	0.3	video camcorder	0.5
music equipment	1.2	computer	1.7
cable TV	9.3		

Table 9: Average assets index by parental education

Parental Education	CGI2
Incomplete primary	11.5
Complete primary	13.3
Incomplete secondary	14.8
Complete secondary	16.0
Incomplete tertiary	15.9
Complete tertiary	17.8

Table 10: Correlation between socioeconomic variables

(obs=105537)	educ	assets	school mean edu	school mean assets	school st.d edu	school st.d assets	neigh st.d edu	neigh st.d assets
parents education	1.0000							
assets	0.3201	1.0000						
school mean edu	0.5239	0.3577	1.0000					
school mean assets	0.3737	0.5015	0.7133	1.0000				
school st.d edu	-0.0642	-0.1104	-0.1226	-0.2202	1.0000			
school st.d assets	-0.1321	-0.1231	-0.2522	-0.2456	0.2702	1.0000		
neigh st.d edu	0.0740	-0.0231	0.1412	-0.0460	0.3036	0.0103	1.0000	
neigh st.d assets	0.0010	-0.1235	0.0019	-0.2464	0.0128	0.2701	0.3239	1.0000

Table 11: Summary Statistics for full sample

	Total		Public		Private	
n. students	304,541		218,065		86,476	
%	100		71.60		28.40	
n. schools	3,479		2,871		608	
%	100		82.5		17.5	
	Total		Public		Private	
variable	mean	sd	mean	sd	mean	sd
test score Spanish	63.96	19.24	60.31	18.97	73.15	16.67
parents education	3.67	1.65	3.39	1.61	4.38	1.54
class mean edu	3.67	0.92	3.39	0.77	4.38	0.89
class st.d edu	1.39	0.33	1.45	0.30	1.24	0.35
class cv edu	0.41	0.13	0.45	0.11	0.31	0.13
class poverty rate	0.13	0.14	0.16	0.15	0.05	0.09
neigh mean edu	3.67	0.47	3.61	0.44	3.82	0.50
neigh st.d edu	1.59	0.09	1.59	0.09	1.57	0.10
neigh poverty rate	0.13	0.07	0.14	0.07	0.11	0.06
assets	63.96	19.24	60.31	18.97	73.15	16.67
class mean assets	27.06	13.50	24.82	13.38	32.69	12.06
class st.d assets	27.06	7.20	24.82	6.47	32.69	5.74
class cv assets	11.59	2.46	11.99	2.17	10.58	2.83
school poverty rate	0.17	0.17	0.21	0.17	0.09	0.11
neigh mean assets	27.06	3.86	26.72	3.96	27.91	3.46
neigh st.d assets	12.93	0.65	12.94	0.67	12.91	0.60
neigh poverty rate	0.17	0.09	0.18	0.09	0.16	0.08

Source: Operativo Nacional Educativo, Ministerio de Education y Cultura, Argentina

Table 12: Summary Statistics for panel of schools sample, per year

Variables	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
	year = 1997 (N = 23,464)				year = 1999 (N = 19,812)				year = 2000 (N = 62,267)			
test score in Spanish	58.05	19.68	0.00	100.00	55.39	19.93	0.00	100.00	62.37	18.68	0.00	100.00
parents education	3.44	1.58	1.00	6.00	3.42	1.65	1.00	6.00	3.63	1.65	1.00	6.00
class mean edu	3.44	0.89	1.40	5.93	3.42	0.93	1.00	5.88	3.63	0.81	1.00	5.88
class st.d edu	1.31	0.24	0.26	2.37	1.36	0.27	0.00	2.66	1.44	0.22	0.00	3.54
class cv edu	0.40	0.11	0.04	0.79	0.43	0.13	0.00	0.99	0.42	0.11	0.00	1.01
class poverty rate	0.16	0.15	0.00	0.88	0.11	0.12	0.00	1.00	0.14	0.12	0.00	1.00
neigh mean edu	3.44	0.46	1.87	5.02	3.42	0.52	1.06	5.14	3.66	0.42	1.72	5.05
neigh st.d edu	1.51	0.13	0.86	1.84	1.56	0.14	0.18	1.94	1.62	0.08	1.04	1.94
neigh poverty rate	0.16	0.08	0	0.6	0.11	0.07	0	0.88	0.14	0.07	0.02	0.69
segregation index	0.36	0.26	0.00	1.56	0.35	0.26	0.00	2.34	0.24	0.14	0.00	1.51
assets	27.47	13.51	0.00	43.10	25.26	13.91	0.00	43.10	26.40	13.51	0.00	43.10
class mean assets	27.47	7.09	5.23	42.40	25.26	7.14	1.50	41.63	26.40	6.55	0.47	42.21
class st.d assets	11.56	2.08	0.73	17.07	12.04	2.02	1.27	20.44	11.86	1.61	0.67	25.46
class cv assets	0.46	0.17	0.02	1.44	0.52	0.18	0.03	1.99	0.48	0.15	0.02	2.49
class poverty rate	0.16	0.16	0.00	1.00	0.16	0.16	0.00	1.00	0.17	0.14	0.00	1.00
neigh mean assets	27.47	4.49	9.86	35.47	25.25	4.51	6.41	38.94	26.55	4.15	7.46	35.36
neigh st.d assets	12.74	1.09	6.66	15.5	13.17	0.95	7.57	17.04	12.97	0.71	9.66	14.81
neigh poverty rate	0.16	0.1	0.03	0.63	0.16	0.1	0	0.72	0.17	0.09	0.05	0.81
segregation index	0.24	0.20	0.00	2.04	0.23	0.18	0.00	1.28	0.19	0.12	0.00	3.65
prop. male	0.50	0.50	0.00	1.00	0.48	0.50	0.00	1.00	0.49	0.50	0.00	1.00
repetitors	0.21	0.41	0.00	1.00	0.24	0.43	0.00	1.00	0.21	0.41	0.00	1.00
class size	26.67	5.86	4	42	26.17	6.11	2	47	26.09	5.85	3	58

Source: Operativo Nacional Educativo, Ministerio de Educación Y Cultura, Argentina

Table 13: Percentiles of main variables

Variables	p10	p50	p90	mean	sd
Panel of schools					
test score in Spanish	33.33	60.00	86.67	60.1	19.35
parental education	1.50	3.50	6.00	3.55	1.64
class mean edu	2.41	3.46	4.83	3.55	0.91
class sd edu	1.00	1.40	1.73	1.38	0.29
class poverty rate	0.00	0.10	0.33	0.14	0.15
neigh. mean edu	2.97	3.60	4.00	3.57	0.46
neigh. sd edu	1.43	1.61	1.69	1.58	0.12
neigh. poverty rate	0.07	0.12	0.22	0.14	0.08
Full sample					
test score in Spanish	36.67	66.67	90	63.96	19.24
parental education	1.5	4	6	3.67	1.65
class mean edu	2.56	3.58	4.97	3.67	0.92
class sd edu	0.97	1.42	1.76	1.39	0.33
class poverty rate	0	0.09	0.33	0.13	0.14
neigh. mean edu	3.1	3.67	4.22	3.67	0.47
neigh. sd edu	1.45	1.6	1.68	1.59	0.09
neigh. poverty rate	0.06	0.11	0.21	0.13	0.07

Source: Operativo Nacional Educativo, Ministerio de Educación y Cultura, Argentina

References

Ainsworth, J. W. (2002). Why does it take a village?: The mediation of neighborhood effects on educational achievement, *Social Forces* **81**: 117–152.

Akerlof, A. (1997). Social distance and social decisions, *Econometrica* **65**(5): 1005–27.

Allen, R. & Vignoles, A. (2006). What should an index of school segregation measure? Centre for the Economics of Education Discussion Paper 60.

Berlinski, S., Galiani, S. & Gertler, P. (2006). The effect of pre-primary education on primary school performance. IFS Working Paper W06/04, Institute for Fiscal Studies.

Borjas, G. (1995). Ethnicity, neighborhoods, and human capital externalities, *American Economic Review* **85**(3): 365–90.

Brock, W. A. & Durlauf, S. N. (2001). Discrete choice with social interactions, *Review of Economic Studies* **68**(2): 235260.

Brock, W. A. & Durlauf, S. N. (2002). A multinomial-choice model of neighborhood effects, *American Economic Review* **92**(2): 298–303.

Case, A. C. & Katz, L. F. (1991). The company you keep: the effects of family and neighborhood on disadvantaged youths. Working Paper No. 3705, National Bureau of Economic Research.

Cutler, D. & Glaeser, E. (1997). Are ghettos good or bad?, *Quarterly Journal of Economics* **112**(3): 827–72.

DINIECE (2003). Informe resultados one 2003. 3er, 6to y 9no ao egb - fin del nivel medio o polimodal, *Technical report*, Sistema Nacional de Evaluación de la Calidad de la Educación, Ministerio de Cultura y Educación.

Duflo, E., Kremer, M. & Robinson, J. (2004). Understanding technology adoption: fertilizer in western kenya preliminary results from field experiments. Unpublished Manuscript, Massachusetts Institute of Technology.

Duncan, G. (1994). Families and neighbors as sources of disadvantages in the schooling decision of white and black adolescents, *American Journal of Education* **1003**(3): 20–53.

Ermisch, J. & Francesconi, M. (2001). Family matters: Impacts of family background on educational attainments, *Economica* **68**(270): 137–156.

Falk, A. & Ichino, A. (2006). Clean evidence on peer effects, *Journal of Labor Economics* **24**(1): 3957.

Fertig, M. (2003). Who's to blame? The determinants of German students' achievement in the PISA 2000 study. IZA Discussion Paper 739, Institute for the Study of Labor (IZA).

Fiszbein, A. (1999). Institutions, service delivery and social exclusion: A case study of the education service in Buenos Aires. LCSHD Paper Series No. 47, Human Development Department, World Bank.

Ginther, D., Haveman, R. & Wolfe, B. (2000). Neighbiurhood attributes as determinants of children's outcome: How robust are the relationships?, *Journal of Human Resources* **35**(4): 603–42.

Gorard, S. & Taylor, C. (2002). What is segregation? A comparison of measures in terms of 'strong' and 'weak' compositional invariance, *Sociology* **36**(4): 875–95.

Hanushek, E. A. (1971). Teacher caracteristics and gains in student achievement: Estimation usign micro-data, *American Economic Review* **61**(2): 280–288.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools, *Journal of Economic Literature* **24**(3): 1141–77.

Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003). Does peer ability affect student achievement?, *Journal of Applied Econometrics* **18**(5): 527–44.

Hanushek, E. A. & Rivkin, S. G. (2006). School quality and the black-white achievement gap. NBER Working Paper 12651.

Herrera-Gomez, M. (2005). Multilevel methods: Application to the determinants of students performance. Universidad Nacional de Tucumán.

INDEC (2001). *Censo Nacional de Población y Vivienda 2001*, Instituto Nacional de Estadísticas y Censos, Buenos Aires, Argentina.

Ioannides, Y. M. & Zabel, J. E. (2003). Neighborhood effects and housing demand, *Journal of Applied Econometrics* **18**(1): 563584.

Jargowsky, P. A. (1997). *Poverty and place: Ghettos, barrios, and the American city*, Russell Sage Foundation, New York.

Kingdon, G. (1996). The quality an efficiency of public and private schools: A case study of urban India, *Oxford Bulletin of Economics and Statistics* **58**(1): 55–80.

Kingdon, G. (2006). Public-private-partnership in education in India.

Kremer, M. & Levy, D. M. (2003). Peer effects and alcohol use among college students. Working Paper, Harvard University.

Lazear, E. (1999). Culture and language, *Journal of Political Economy* **107**(6): S95–S126.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem, *The Review of Economic Studies* **60**(3): 531–42.

Manski, C. (2000). Economic analysis of social interactions, *The Journal of Economic Perspectives* **14**(3): 115–36.

MCE (1996). Nivel socioeconómico familiar y rendimiento escolar 1994/95. Sistema Nacional de Evaluación de la Calidad de la Educación, Ministerio de Cultura y Educación.

MCE (1997). Composición social de la escuela y rendimiento de los alumnos. Working Paper 28, Sistema Nacional de Evaluación de la Calidad de la Educación, Ministerio de Cultura y Educación.

Miguel, E. & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities, *Econometrica* **72**(1): 159217.

Naradowski, M. & Andrada, M. (2001). The privatization of education inargentina, *Journal of Education Policy* **16**(6): 585595.

Sacerdote, B. (2001). Peer effects with random assignments: Results for Dartmouth roommates, *Quaterly Journal of Economics* **116**(2): 681–704.

Soetevent, A. R. (2006). Empirics of the identification of social interactions: An evaluation of the approaches and their results, *Journal of Economic Surveys* **20**(2).

Stewart, E., Stewart, E. & Simons, R. (2007). The effect of neighbourhood context on the college aspirations of African American adolescents, *American Educational Research Journal* **44**(4): 896–919.

Taylor, C., Gorard, S. & Fitz, J. (2000). A re-examination of segregation indices in terms of compositional invariance, *Social Research Update* **29**: 1–6.

Wilson, W. (1996). The economic structure of higher education: Subsidies, customer-inputs, and hiercharchy. Williams Project on the Economics of Higher Education Discussion Paper 40.

Wilson, W. (1997). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, University Chicago Press, Chicago.

Winston, G. & Zimmerman, D. (2003). Peer effects in higher education. NBER Working Paper 9501.

Woessmann, L. (2004). How equal are educational opportunities? Family background and student achievement in Europe and the United States. IZA Discussion Paper 1284, Institute for the Study of Labor (IZA).

Zhang, X. & Kanbur, R. (2001). What difference do polarisation measures make? an application to china.

Zimmerman, D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment, *Review of Economics and Statistics* **85**(1): 98–23.

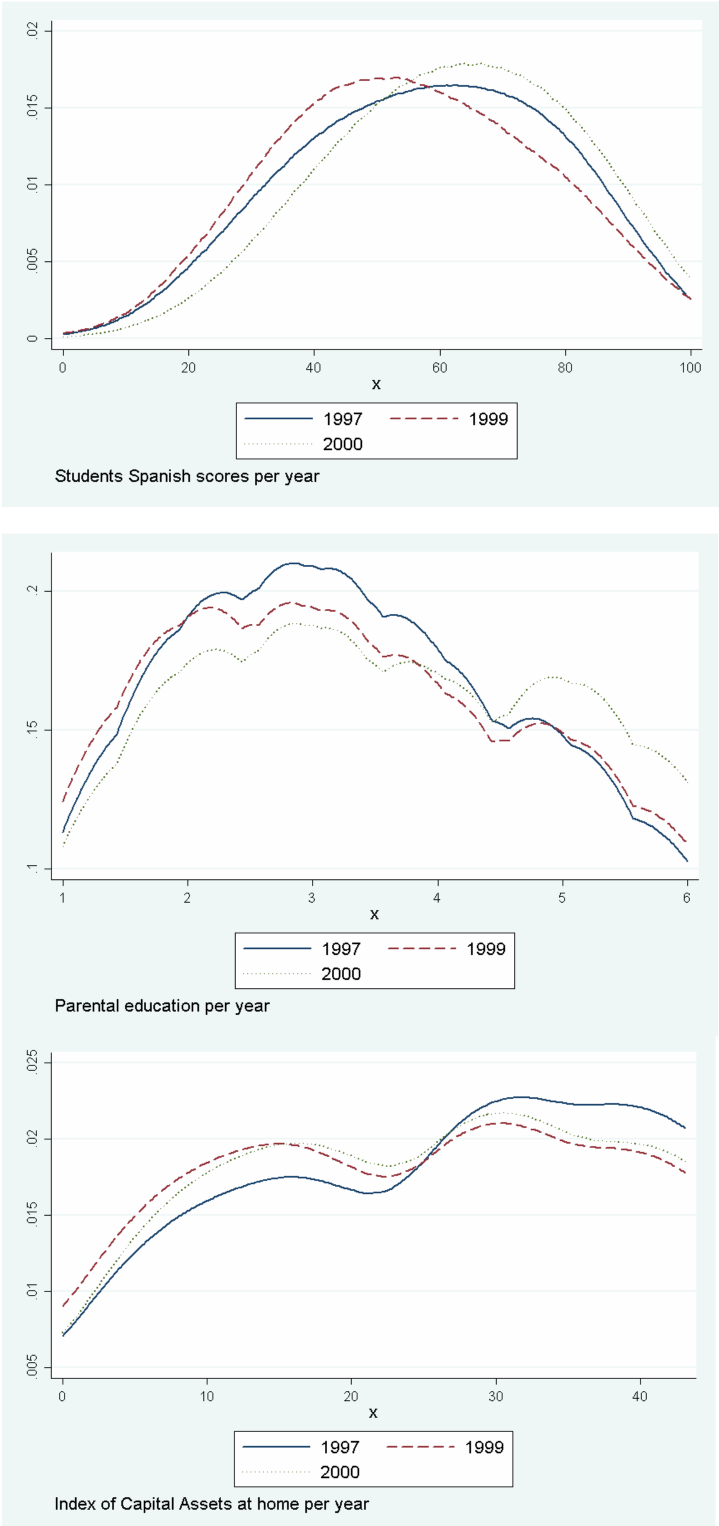


Figure 5: Distributions of main variables

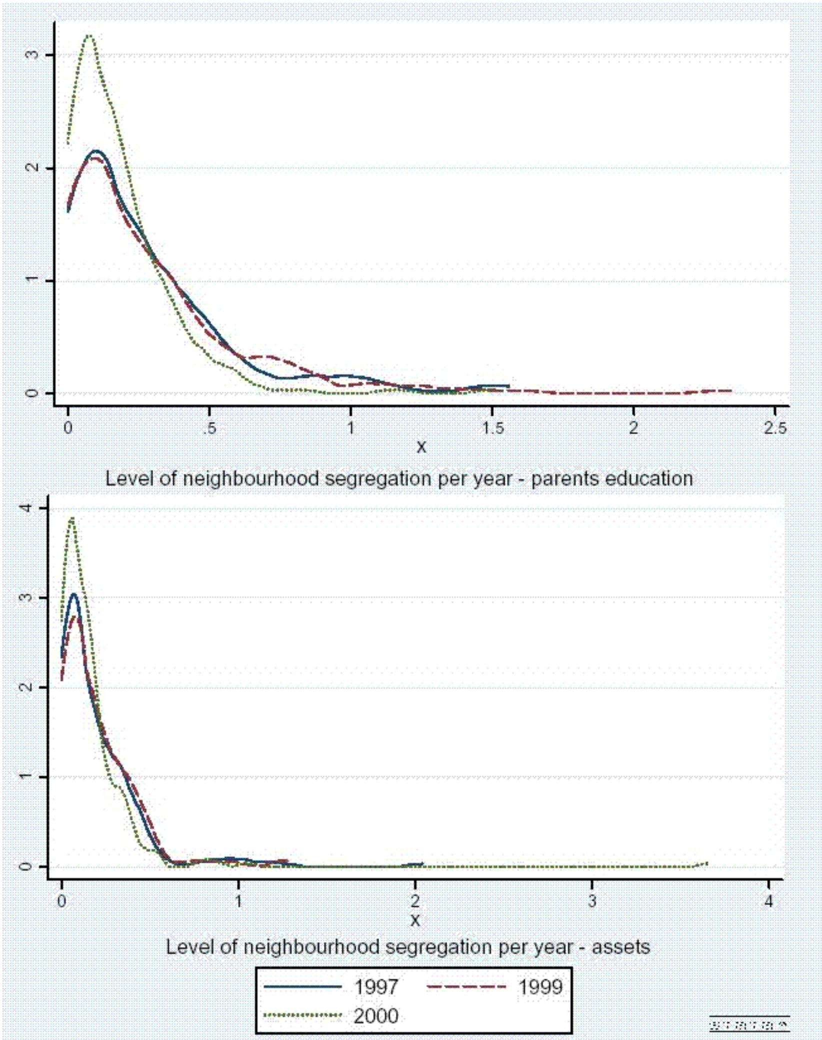
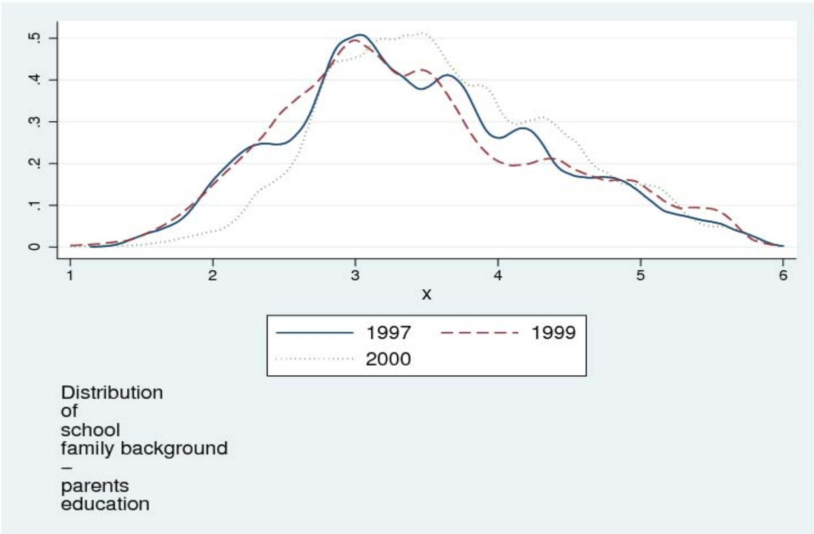


Figure 6: Distributions of segregation index



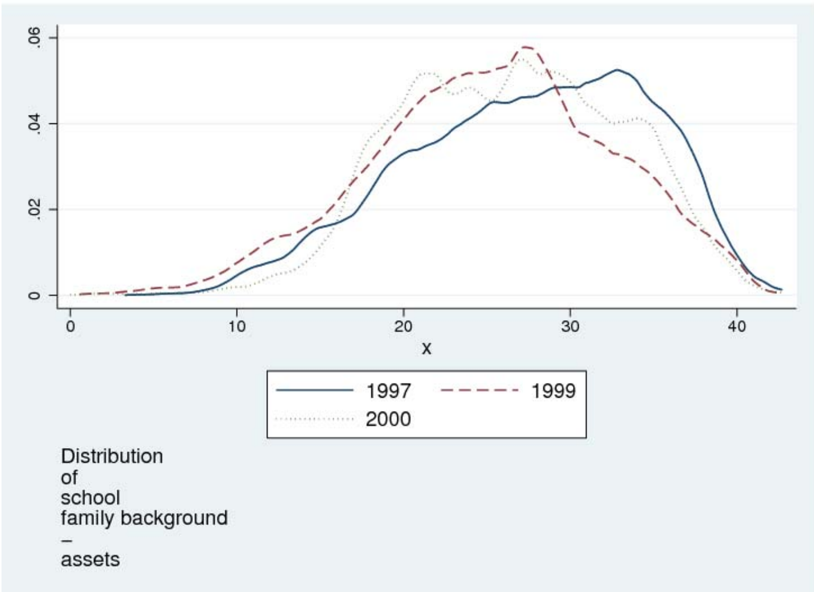


Figure 7: Distribution of school average parental education

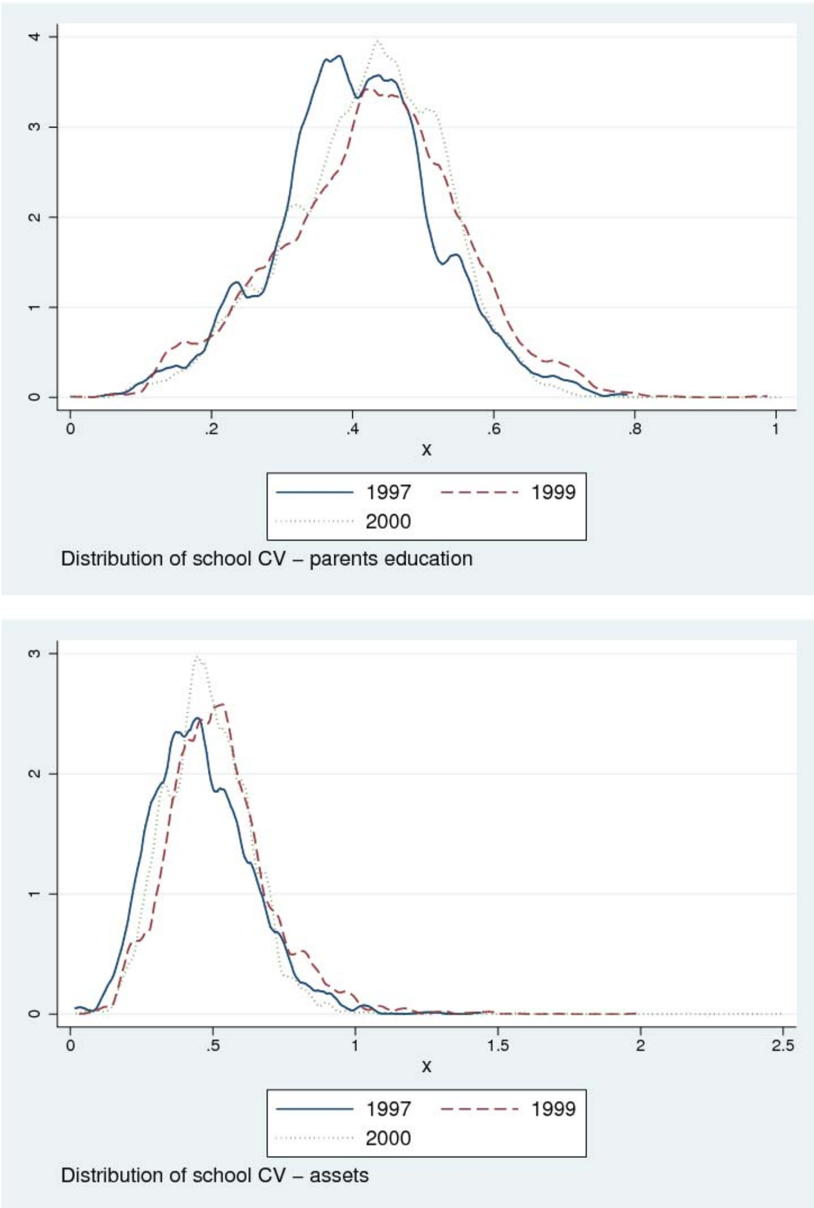


Figure 8: Distribution of school dispersion parental education