The version of record of this article has been published and is available in *British Journal of Educational Studies* (date of publication 23 July 2020) https://doi.org/10.1080/00071005.2020.1795078

School segregation in public and semiprivate primary schools in Andalusia

Claudia Prieto Latorre[1]; Oscar D. Marcenaro-Gutierrez[2]; Anna Vignoles[3]

Abstract

School segregation by socio-economic background is a feature of most education systems and impacts negatively on educational outcomes for poor children. Evidence on this issue is lacking for the Spanish primary education system and in particular the extent to which a) poor and rich students sort into different types of school and b) the extent of segregation within different school types. We measure the level of segregation of students from different socioeconomic backgrounds into public and semiprivate schools and within these two types of school using the Hutchens Index. The analysis is based on data for students attending 5th grade in the largest region in Spain (Andalusia). Our results indicate significant differences in the level of segregation across and within both school types. Students from higher socioeconomic backgrounds are more likely to attend semiprivate schools. This provides evidence of the inequity of providing public funding for semiprivate schools that are then disproportionately attended by wealthier students.

Key words: school segregation; socioeconomic level; public schools; semiprivate schools; Hutchens Index.

Funding

This work was supported by the Ministerio de Economía, Industria y Competitividad [Research Project ECO2017-88883-R]; Ministerio de Educación, Cultura y Deporte [Scholarship FPU2017-00432]; Centro de Estudios Andaluces [Research Contract PRY85/19]; Universidad de Málaga [PhD Program in Economics and Business].

Acknowledgements

The data used in this research have been provided by *the Consejería de Educación de la Junta de Andalucía*. This work has been supported by the *Ministerio de Economía, Industria y Competitividad* under Research Project ECO2017-88883-R and *Fundación Pública Centro de Estudios Andaluces under Research Contract PRY85/19*. Claudia Prieto Latorre acknowledges the scholarship FPU2017-00432 of the Ministry of Education, Culture and Sports [*Ministerio de Educación, Cultura y Deporte*] and the training received from the University of Malaga PhD Program in Economics and Business [*Programa de Doctorado en Economía y Empresa de la Universidad de Malaga*].

[1] Programa de Doctorado en Economía y Empresa, Universidad De Málaga, Málaga, Spain. E-mail: claudiaprieto@uma.es

[2] Department of Applied Economics (Statistics and Econometrics), Universidad de Málaga, Málaga, Spain. E-mail: odmarcenaro@uma.es

[3] Faculty of Education, University of Cambridge, Cambridge, UK. E-mail: av404@cam.ac.uk

1. Introduction

An education system may play a role in fostering social cohesion and social mobility. However, in a highly segregated school system, in which poor and rich students are educated in separate schools, it is less likely that these goals can be achieved. We analyse school segregation in primary schools in a region of Spain, Andalusia, exploring the extent of socio-economic segregation across different school types (public and semiprivate schools) and within these school types. Our analysis is focused on socio economic segregation since this factor has been found to have a strong association with students' academic performance (see Tourón et al. 2018 for evidence for Spain), though other sources of segregation are also of importance.

There is a growing policy focus on increasing socio-economic diversity within schools. The function of an education system is not simply to foster learning but also to develop cultural and social capital (Krüger, 2019). The latter arguably requires socio-economic integration of pupils. According to García-Huidobro (2010), three dimensions characterize an inclusive education system: access to education for different groups of students; academic achievement, and specifically equity in opportunities to learn; and lastly, integration, namely the extent to which different types of students are educated together. It is this last dimension on which our research is focused.

Another motivation for reducing socio-economic segregation in schools is its potential impact on pupil achievement via peer effects. The literature on the impact of the socio-economic background of peers on academic achievement is mixed (Epple and Romano, 2011) but there is some empirical evidence that school segregation and peers matter for student outcomes (e.g. Gibbons and Telhaj, 2016; Hanushek et al., 2003; Palardy, 2013). Evidence of long term negative peer effects include reduced likelihood of high school graduation and probability of enrolling in a 4 year-college (Harding, 2003; Wodtke et al. 2011; Palardy, 2013). Concentrating students from low income households in the same schools therefore potentially introduces additional disadvantage. Conversely, schools with disproportionate numbers of high-income students may provide an additional educational advantage to those students.

There exist also evidence of the direct effect of socioeconomic segregation on students' academic performance (Willms, 1986; Rumberger and Palardy, 2005; Danhier, 2018). Differences in school policy and practices between schools serving high and low income students explains the different academic performance (Rumberger and Palardy, 2005); in particular, socioeconomically advantaged schools provide students with more opportunities to learn (more homework and take more advanced classes) as well as teachers have higher expectations. Additionally, the impact of school composition is independent to the initial level of ability of the students (Willms, 1986), what means that the socioeconomic composition of the class has the same effect for high ability and low ability students. For this reason, measuring the extent of socio economic segregation in a school system is important from a policy perspective.

We use unique census data from the largest region of Spain (Andalusia) to investigate socio-economic segregation in the primary school system. This region is one of the worst in terms of education performance in Spain (which, in itself, performs slightly below the OECD average in PISA; OECD, 2014). Andalusia has a high dropout rate (around a third of boys and a quarter of girls dropped out before finishing secondary education in 2012; IECA, 2017). The concentration of students in the bottom tail of the distribution of test scores' is significantly higher in Andalusia than in the other Spanish

Regions (PISA 2012): around 27% of Andalusian students have low performance in mathematics (competency below level 2) while only 6% of students have high performance (level 5 or 6)[4]. This is particularly relevant given that Andalusia is the most populated region in Spain. Given the relatively poor levels of academic achievement in this region, measuring the level of socio economic segregation may provide insights into one factor that might potentially be linked to poor academic performance.

The empirical evidence presented here is novel in a number of ways. Our main contribution is to examine socioeconomic segregation in primary schools in a region of Spain: previous research on Spain has focused on secondary education. Additionally, to the best of our knowledge, this is the first work to use the rich census data on 5th grade Andalusian students who attend public and semiprivate schools. This has allowed us not just to provide evidence on the level of segregation overall but the extent to which it is related to a distinct feature of the Andalusian education systems, namely a highly subsidised semi-private school sector. For the analysis we use the Hutchens Index. The strength of this index is its decomposable property, which allows us to quantify the contribution of each type of school (public and semiprivate) to the overall level of segregation. This is relevant for the Spanish case because, despite the fact that the legislation regulating admission to both types of schools is based on the same criteria, previous evidence on secondary education has suggested differences in the level of socio economic segregation in public and semiprivate schools (Mancebón and Pérez, 2010; García-Montalvo, 2013; Murillo and Martínez-Garrido, 2018).

2. Literature review

On average, children from high socio economic status (SES) backgrounds tend to have higher levels of academic achievement (Lucas, 2001; Sirin, 2005; Martins and Veiga, 2010; Hanushek and Woessmann, 2011). However, the strength of the relationship between SES and academic outcomes varies by context (Breen and Jonsson, 2005) and this may be related to segregation in the school system.

In the case of Spain, there is a positive correlation between SES and academic achievement, a relationship which has been observed in repeated rounds of PISA (MECD, 2016). Cordero et al. (2013) also stress the importance of parental education in predicting a child's educational achievement in Spain, a finding confirmed for Andalusia by Gutiérrez (2014), (see also Tourón et al. (2012) using TIMSS 2011 data, Corral Blanco et al. (2012) using PIRLS 2011 and Tourón et al. (2018) using PISA 2015). Although the influence of SES on students' performance is lower in Spain than in many other countries, the correlation has increased from 2003 to 2012, especially in public schools (no trend over time has been found in private schools; see García-Montalvo, 2013). Martins and Veiga (2010) note that the strong correlation between SES and student achievement in Spain reflects the relatively high level of socioeconomic inequality.

One potential source of inequality in education achievement is socio-economic segregation of pupils. Previous literature from Spain points to sorting of students such that students' SES determines which school they attend (see also Mancebón and Pérez

[4] For Spain the share of low achievers in mathematics in PISA 2012 is 23%. 8% obtain a competency level of 5 or 6. In countries such as Japan and Finland (at the top of the PISA 2012 distribution), the proportion of top performers is up to four times higher than in Andalusia (23.7% in Japan and 15.3% in Finland) while the share of students performing at Level 2 or below is half the Andalusian rate (11.1% in Japan and 12.3% in Finland).

(2010) and García-Montalvo (2013)). Numerous studies have found that students benefit from being around high-achieving peers - see Sacerdote (2001), Hanushek et al. (2003), Ammermueller and Pischke (2009) and Vardardottir (2013) - , though many studies have found modest impacts (Epple and Romano, 2011). In Spain, there is some evidence that the socioeconomic composition of classmates exerts a significant influence on students' performance (Cordero et al., 2011), but socioeconomic class composition has lower effect in Spain than in other countries such as France (Danhier, 2018).

In terms of the international context, Spain is not the most segregated country and sits just above the OECD average (Gutiérrez et al. 2017). When compared with other European countries, segregation of the bottom decile of students in Spain is also around the average (the Gorard Index in Spain for this decile is 0.41 and 0.43 for France, for instance) but the level of segregation in Spain is significantly greater for high SES students (for the top decile, the Gorard Index in Spain is 0.46, while the OECD average is 0.41 according to Murillo and Martínez-Garrido, 2018). Within Spain, there are also regional differences in the level of school segregation (Mancebón and Pérez, 2010; Murillo et al. 2018), although this evidence is for secondary education. Specifically, Murillo et al. (2018), using data from PISA 2015, found that the level of segregation of low SES students (bottom quartile of the SES) in Andalusia (Hutchens Index of 0.12) was below the average for Spain (0.17).

This paper contributes to this literature with new empirical evidence for the primary school system in Andalusia.

3. Institutional features of the Spanish education system

The Spanish education system classifies schools according to ownership and funding source. There are three distinct school types: public schools, semiprivate schools and private schools. Specifically, in 2007-08, 76% of students attended public schools, 21% attended semiprivate schools and 3% private schools, with a higher proportion of pupils in semi-private schools in urban areas distribution of non-public schools between rural and urban areas. We focus on schools which receive public funding, i.e. public and semiprivate schools, though the latter can also charge fees. We remove private schools from our analysis because there is a high proportion of missing data on the student socioeconomic status variable which could introduce a source of bias[5]. Spain also operates a school choice system: parents can theoretically choose which school to send their child to. For oversubscribed schools, access is regulated at national level in accordance with a prescribed student-assignment system. Students get priority when they have siblings attending that school, when their parents work there and when their home or parents' workplace is close to the school. Nevertheless, it is important to note that there are formal and informal mechanisms to circumvent the theoretical equality established in the law.

The over-subscription criteria based on proximity to the school will drive socio-economic segregation. If houses nearer over-subscribed schools are more expensive as parents seek to access such schools, this will increase socio-economic segregation in schools.

[5] In the case of public and semiprivate schools, there is not statistical significant differences between the proportion of missing data in each type of school, i.e. the missing data are randomly distributed.

Another source of potential socio-economic segregation is the fact that semiprivate schools can charge fees[6]. Specifically, public funding financed on average 73% of the resources spent on students attending semiprivate schools (academic year 2009/10)[7], while 23% came from tuition fees paid by families and the rest (3.82%) other private income (donations, payments of parents associations, etc.). This means that if we compare the family expenditure per student in other educational goods and services (extracurricular activities at school, such as learning activities and leisure, complementary services, such as canteen, transport, accommodation, etc., and goods and services, such as school textbooks, uniforms, etc.) during a full academic year, we can see that family expenditure in semiprivate schools is almost twice that in public schools (1193 euros in semiprivate and 627 euros in public schools). Public funds are used in semi private schools to ensure that their educational offer adheres to the national curriculum and to subsidize student fees. The schools in turn must use admissions criteria that are the same as those used by public schools (Murillo et al., 2018).

4. Data

By law the Andalusian Education Act (*Ley de Educación de Andalucía, Ley* 17/2007) requires that the competencies of students be assessed before finishing primary education. The Andalusian Agency of Educational Assessment (AGAEVE) conducts an annual test called Diagnostic Assessment (DA). This assessment is not high stakes, as students' progress is not determined by performance on this test. The census dataset used in this research is focused on the 2008-09 wave, specifically 5th grade Andalusian students, namely 78,413 students, in 2,398 schools, of which 1,676 are public and 402 semiprivate.

The data include questionnaire data provided by students, families, teachers and head teachers. From this data, students' SES is measured using the index of Economic, Social and Cultural Status (ESCS). The ESCS is a synthetic index based on the highest level of parental education, highest parental occupation and home possessions. The home possessions index is used as a proxy for family income and it contains information related to family wealth, cultural possessions, home educational resources and the number of books at home. The three indicators composing the ESCS index have been found to have a causal impact on students' performance independently. Although the effect of household income on achievement is the most studied in the literature (Sirin, 2005), the inclusion of parental education and parental occupation indicators better captures students' socioeconomic background. Indeed, Ilie et al. (2017) found that both these latter indicators are stronger predictors of students' performance than self-reported household income.

Missing data on ESCS reduces the sample size to 63,155 students. Before analysing the level of segregation, some descriptive statistics provide insights which we explore further in our analysis. Table 1 shows the percentage of students attending public and semiprivate schools by ESCS. For example, the proportion of students attending

[6] There are other differences between semiprivate and public schools. Recruitment of teachers in the former does not have to be based on the official state exams for teachers, whereas that is a requirement in the case of public schools. Public schools are secular. Semiprivate schools can be religious or secular. Around 70% of semiprivate schools are religious.

[7] Data referred to 2009-10 non university studies (INE, 2010, Survey of Financing and Expenditures of Private Teaching, 2010). This information is not disaggregated by educational level.

public and semiprivate schools from the lower quartile of the ESCS should be 25%. However, we see that in fact 29% of students attending public schools are from the bottom quartile of the socio economic distribution while in the case of semiprivate schools only 11% of students come from the bottom quartile. We see that just 19% of students attending public schools are from the top quartile of the socio economic distribution, while 44% of students attending semiprivate schools are from the top quartile. These results, perhaps unsurprisingly, suggest a higher concentration of students from low ESCS households in public schools and a higher concentration of students from high ESCS backgrounds in semiprivate schools.

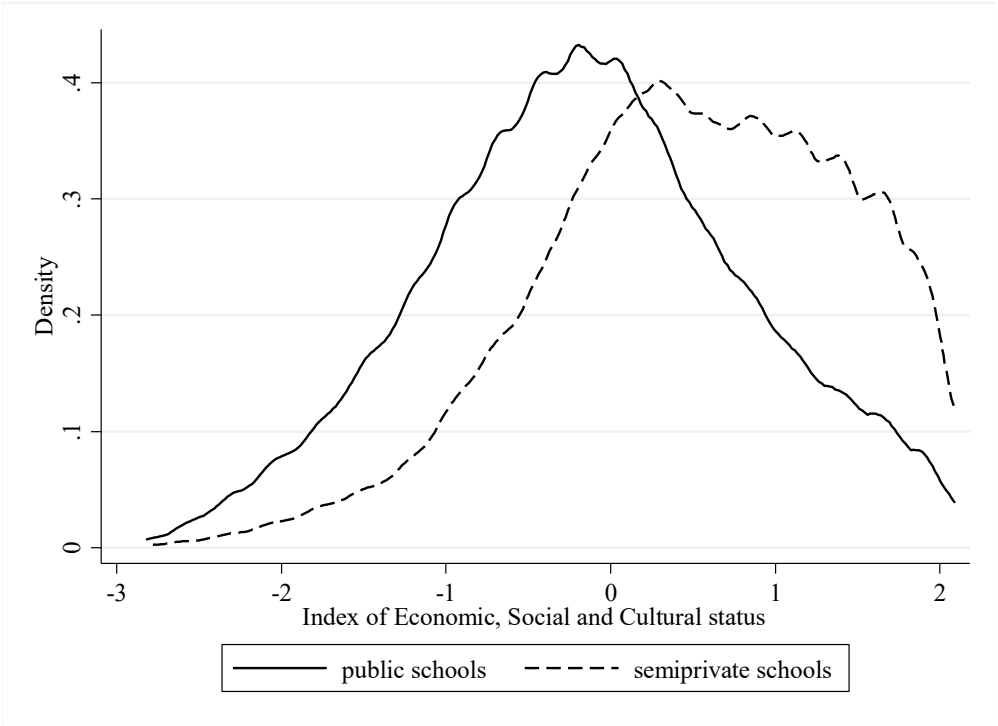
Table 1. Percentage of students who attend public and semiprivate schools by ESCS.

	Public schools	Semiprivate schools
Percentage of students from the **lower tercile of the ESCS**	38.50%	15.89%
Percentage of students from the **lower quartile of the ESCS**	29.17%	10.93%
Percentage of students from the **lower decile of the ESCS**	11.90%	3.62%
Percentage of students from the **top tercile of the ESCS**	27.16%	54.08%
Percentage of students from the **top quartile of the ESCS**	19.35%	43.93%
Percentage of students from the **top decile of the ESCS**	7.01%	19.59%

Note: Cut-points are shown in the first column of the table. These cut-points have been defined by the Andalusian distribution of the ESCS of the students attending public or semiprivate schools in 5th grade in 2008-2009.
Source: Authors own calculation.

Figure 1 shows the distribution of the ESCS index for Andalusian students attending public and semiprivate schools in 5th grade and confirms that students attending semiprivate schools are concentrated in the higher levels of the ESCS distribution.

Figure 1. Distribution of the ESCS (socioeconomic and cultural level) of students attending public and semiprivate schools.



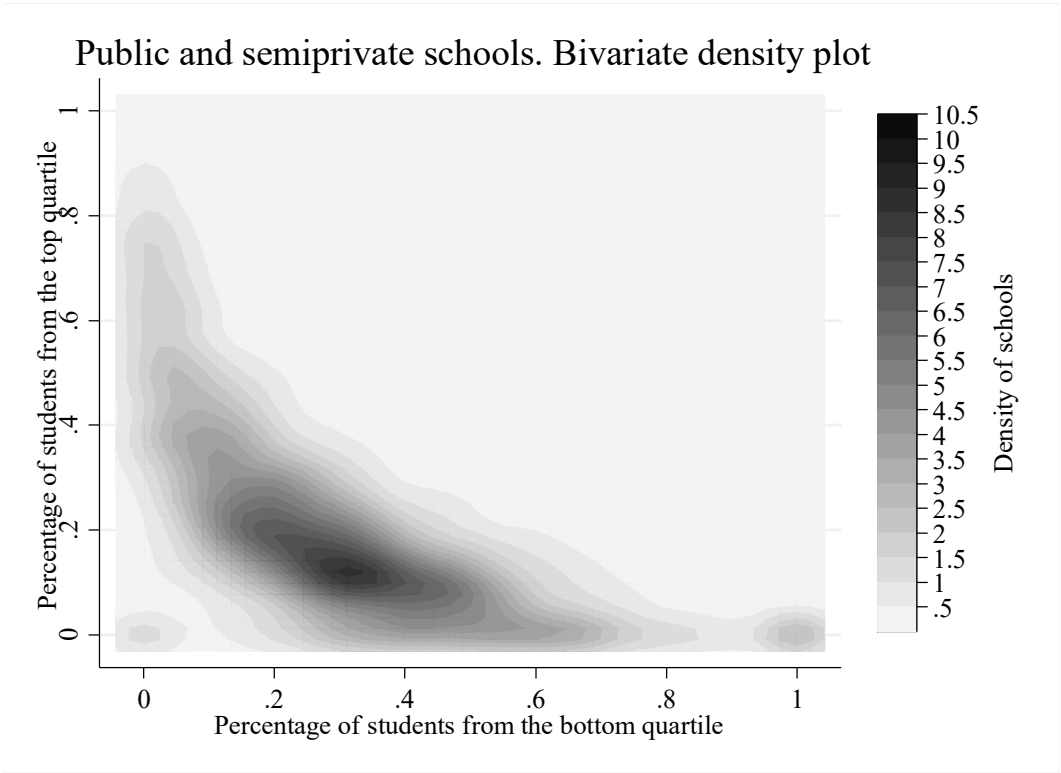
Source: Authors' own calculations.
Data source: Andalusian students attending 5th grade of primary education in 2008-09. Data provided by AGAEVE (Andalusian Agency of Educational Assessment).

Bivariate kernel density estimation contour plot

The following figures analyse the distribution of socioeconomically advantaged and disadvantaged students across schools. The axes show the share of students within each school type. The horizontal axis measures the share of students from the bottom ESCS quartile. The vertical axis measures the percentage of students from the top ESCS quartile. The contour plot then measures the density of schools for each x,y combination as measured by the right hand axis. White indicates low density, black indicates high density.

Figure 2 shows the plot for the whole sample of schools (public and semiprivate schools). An even distribution of students would exist if the share of students in each school is 25% from the top and bottom quartile of the ESCS distribution. In the data there is a relatively high number of schools (dark grey colour) in which 25% of students come from the bottom quartile as expected but just 15% of students from the top quartile. Overall this suggests that there are some schools with a concentration of high ESCS students and many more schools where high ESCS students are underrepresented or absent altogether.

Figure 2. Contour plot of the bivariate distribution of the share of students from the 1st and 4th quartile (of the ESCS distribution) across schools.

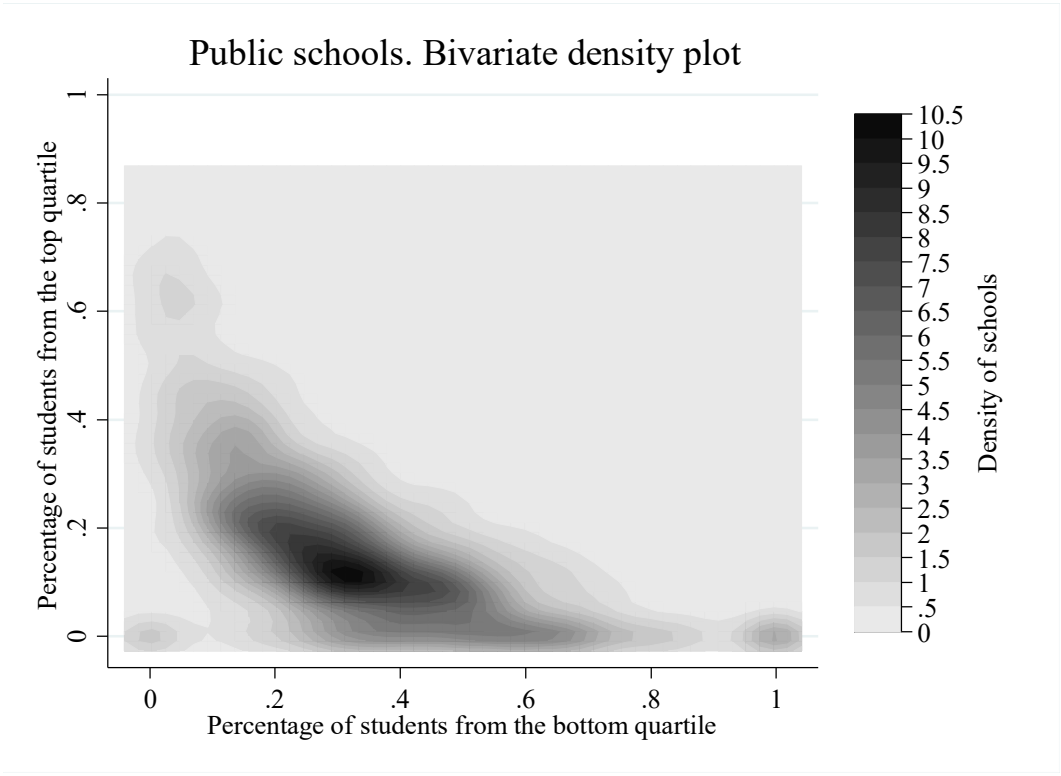


Source: Authors' own calculations.
Data source: Andalusian students attending 5th grade of primary education in 2008-09. Data provided by AGAEVE (Andalusian Agency of Educational Assessment).

Figures 3 and 4 show a similar plot by school type. For public schools (Figure 3), there is a cluster of schools with around 35% of students from the lowest ESCS quartile and only around 15% of students from the top quartile. The distribution of students across semiprivate schools is very different (Figure 4), there is a higher density of semiprivate schools with around 40% of students from the top ESCS quartile and only around 5% of disadvantaged students from the bottom ESCS quartile.

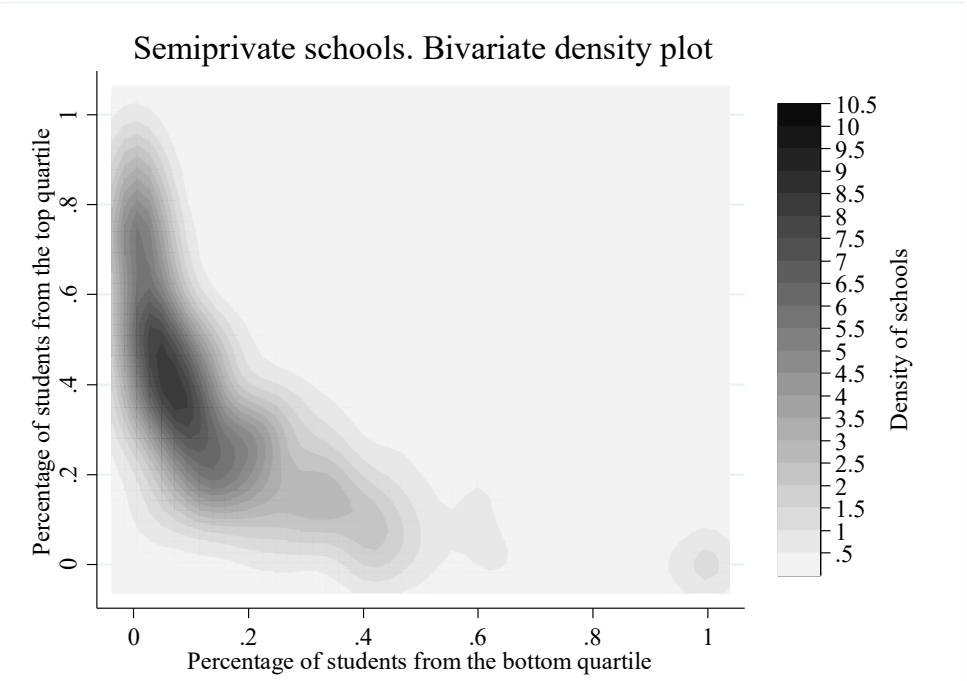
In both the public and semiprivate school samples there are schools in which there are only students from a disadvantaged background, but the density of such schools is higher in the public school sample. There is also a cluster of semiprivate schools in which 80% of students are from the top quartile. Overall, social segregation across both public and private schools is evident in these data.

Figure 3. Contour plot of the bivariate distribution of the share of students from the 1st and 4th quartile (of the ESCS distribution) across public schools.



Source: Authors' own calculations. Data source: Andalusian students attending 5th grade of primary education in 2008-09. Data provided by AGAEVE (Andalusian Agency of Educational Assessment).

Figure 4. Contour plot of the bivariate distribution of the share of students from the 1st and 4th quartile (of the ESCS distribution) across semiprivate schools.



Source: Authors' own calculations.
Data source: Andalusian students attending 5th grade of primary education in 2008-09. Data provided by AGAEVE
(Andalusian Agency of Educational Assessment).

5. Methodology

Although not widely used in the literature, we use the Hutchens Index (Hutchens 2001, 2004) as it has several desirable properties appropriate for this context. The index analyses the uneven distribution of a binary classification of the population across social units (schools) according to a criterion. In our case, we divide the population into two groups (high and low socioeconomic status) based on the ESCS distribution. We use the Hutchens Index to assess the segregation of low and high ESCS students both across all schools and within different school types. The Hutchens Index is defined as follows:

$$H = 1 - \sum_{i=1}^{2078} \sqrt{\frac{a_i}{A}\,\frac{b_i}{B}} \qquad (1)$$

where i is the number of schools, a_i is the proportion of students within school i from a low ESCS household, b_i is the proportion from a high ESCS household, and A and B are the proportion of students in the population who belong to the low and high ESCS groups respectively.

The value of the Hutchens index ranges from 0 to 1, where 0 means complete evenness and 1 complete unevenness. The index compares the proportion of students from each ESCS group within the school with the proportion expected across the whole sample. In the case that the representation of ESCS groups within schools is exactly the same as that in the population, quotients would be one and the index value 0. Conversely, if high and low ESCS students are entirely segregated into separate schools, the index would take the value of 1. Besides being easily interpretable, fixed boundaries allow comparison across different distributions.

According to Hutchens (2001), this index satisfies all the properties which are desirable for a measure of dissimilarity. The index is also consistent with the segregation curve approach, which is an advantage (Hutchens, 1991). We now turn to consider the specific properties of this numerical measure of segregation.

The Hutchens Index (2004) is scale invariant. A proportional increase in the number of students in any group (rich or poor), in such a way that the proportion within each school remains the same, does not affect unevenness. Neither does a proportional division of students who attend a school into subgroups (insensitivity to proportional division). Students of a particular school can be exchanged with those who attend another school and, although the particular contribution to the segregation level of those specific schools would be exchanged too, the index value does not vary as it measures segregation across all schools (symmetry in groups). Labels of the type of student can be traded without affecting the value of the index, that is, the rich and poor categories can be exchanged without changing the index value (symmetry in types). Lastly, if we add a school that does not include either type of student, the index is not altered (the zero member property). The only property under which there is a change in the value of the index is the so called "disequalizing movement", which is defined by the movement of a student (rich or poor) from one school to another. In this case, the shares of poor and rich students will change and consequently so will the value of the index.

The most commonly used index for measuring school segregation is the Dissimilarity Index (Duncan and Duncan, 1955). This index is consistent with segregation curves and meets most of the defined properties of the Hutchens index but it does not strictly accomplish the axiom of movement between groups (Allen and Vignoles, 2007).

Another axiom that both the Dissimilarity Index and the Hutchens Index does not meet is "strong composition invariance", although there is a disagreement on the suitability of this feature (see Gorard and Taylor, 2002; Allen and Vignoles, 2007). An index is strongly composition invariant if the value remains constant when, for example, the number of low ESCS pupils increases or decreases proportionally within each school. In this research we use a relative measure of ESCS which therefore cannot change so this issue is less relevant.

The Hutchens Index property of greatest importance for our purposes is its additivity in groups. The Hutchens enables us to split the total level of segregation in the system into different subsystems (semiprivate and public schools). An alternative formulation of the Hutchens Index (Jenkins et al. 2008) is:

$$H = \sum_{i=1}^{2078} Contribution_i = \sum_{i=1}^{2078} \left(\frac{a_i}{A} - \sqrt{\frac{a_i}{A} \frac{b_i}{B}} \right) \quad (2)$$

The level of segregation can be decomposed into the extent of unevenness within and across schools. Based on this characteristic of the index, the property of additivity in groups decomposes the segregation across schools into the segregation within groups plus the segregation between groups, where within group segregation is weighted by the size of each subsystem:

$$H = H_{within} + H_{between} \quad (3)$$

$$H_{within} = w_{public} * H_{public} + w_{semiprivate} * H_{semiprivate} \quad (4)$$

where $w_g = \sqrt{\frac{a_g}{A} \frac{b_g}{B}}$, g is the type of schooling (public or semiprivate); a_g is the proportion of low ESCS students in each g type of schooling; b_g is the proportion of high ESCS students in each g type of schooling; *A* and *B* are the proportion of students in the population who belong to a low and high ESCS groups respectively.

The property of additivity in groups is especially useful, enabling the role of semiprivate schools in school segregation to be studied. Previous literature notes differences in the level of segregation between public and semiprivate secondary schools (Mancebón and Pérez, 2010; Llera and Pérez, 2012; Elosua Oliden, 2013; Murillo et al., 2018).

6. Results

The focus of our analysis is not only on students from a low ESCS background, and the extent to which they are segregated into different schools, but also on the distribution of students from high ESCS households. Clearly the potential problems associated with segregation of students from a low income background are of great policy concern, as the low performance of peers may impact negatively on students' achievement and hence constrain social mobility. However, one can argue that the

segregation of higher income students is also of concern, given that it is likely to impact on social cohesion. Although there is a larger literature that has addressed the segregation of disadvantaged students, there is less evidence on the distribution of high income students in public school systems. To consider segregation across the ESCS distribution, we focus on three alternative ESCS thresholds (to check robustness) to identify measure the poorest and the richest students. For the poorest students, we focus upon the bottom tercile, bottom quartile and bottom decile of the ESCS distribution; and for the richest students we focus on the top tercile, top quartile and top decile.

We begin with a graphical analysis of the segregation curve of the whole population of 5th grade students. Figure A1 in the Appendix shows the distribution across schools of students from the first, second, third and fourth quartile of the ESCS distribution. Graphically, the horizontal axis represents the cumulative proportion of students from that particular quartile and the y axis shows the cumulative share of students that each school has from the remaining three quartiles (b_i/B). For example, the segregation curve of the distribution of students from the first quartile would be represented on the horizontal axis as the cumulative percentage of students from the first quartile of the ESCS distribution in schools and on the vertical axis the cumulative share of students from the second to the fourth quartile.

Figure A1 shows that there is segregation of students across the full ESCS distribution. However, the segregation curves of the first and fourth quartiles are furthest from the 45 degree line (the 45 degree line represents no segregation at all). Thus, very high or low ESCS students are more segregated, whilst students from middle income households are less segregated.

Tables 2 and 3 show the extent of socioeconomic segregation in Andalusian primary schools for students from the bottom (Table 2) and the top (Table 3) of the ESCS distribution. The overall level of segregation of low ESCS students across all primary schools is given in the first data column of Table 2. The next two columns show the extent to which that segregation is attributable to school type attended (public or semiprivate). 80% of the total segregation of low ESCS students in the Andalusian primary education system is attributable to segregation of pupils *within* each type of school. The remaining 20% is explained by segregation between public and semiprivate schools. Table 3 suggests a similar pattern for high ESCS students. Much of the overall level of school segregation observed is therefore within school type. It is not simply that poorer students attend public schools and richer students attend semiprivate schools. In fact, there is a high level of segregation across schools within the group of students attending semiprivate schools and likewise within the public school students.

Tables 2 and 3 also show the level of segregation for different thresholds of the ESCS distribution. Broadly the most deprived students are more segregated across schools. The Hutchens Index has a value of 0.203 for the bottom decile of students (measured by the ESCS) against 0.137 for the bottom third of students. This difference derives from a higher level of segregation of the most deprived students within each school type. A similar pattern is observed for higher ESCS pupils: the richest students are more segregated.

Returning to Table 2, the next set of columns show the extent of segregation in public and semiprivate schools respectively. Note that the Hutchens index for public schools is based on the distribution of students from low ESCS backgrounds across different public schools, and the segregation in semiprivate schools is likewise based on

the distribution of low ESCS students across semiprivate schools. For example, in the case of semiprivate schools, only around 3.6% of students come from the lower decile of the ESCS distribution (because of segregation across school type). However, if these 3.6% of students were evenly distributed across each semiprivate school, the Hutchens index for semiprivate schools would have a value of 0 and complete evenness would be attained within that type of school. From Table 3 we can also see that the level of segregation of high ESCS students is actually greater in public schools than in semiprivate schools but only for the top decile of the ESCS distribution. For the other thresholds there is no significant difference in segregation levels for high ESCS pupils in public and semiprivate schools.

Overall the decomposition of the Hutchens Index reveals that the distribution of students within public schools and semiprivate schools (the "within component") has a relatively greater weight than the differences between school types (the "between component") when explaining the level of segregation. Specifically, for the three thresholds of both the low and high ESCS groups, the "within component" ranges from 80% to 85% while the "between component" ranges between 15% and 20%.

Table 2. Socioeconomic segregation of low ESCS students using the Hutchens Index.

				Public schools			Semiprivate schools		
	H	**H_{within}**	**$H_{between}$**	**H_{public}**	**w_{public}**	**Contribution**	**$H_{semipirvate}$**	**$w_{semiprivate}$**	**Contribution**
Lower tercile	0.137	0.111 (80%)	0.026 (20%)	0.093	0.790	0.169	0.209	0.177	-0.032
Lower quartile	0.145	0.120 (82.46%)	0.025 (17.54%)	0.099	0.810	0.170	0.243	0.165	-0.025
Lower decile	0.203	0.178 (87.52%)	0.025 (12.48%)	0.148	0.832	0.208	0.384	0.142	-0.005

Note: Cut-points are established in the first column of the table. These cut-points have been defined by the Andalusian distribution of the ESCS of the students attending public funded schools in 5th grade in 2008-2009.
Source: Authors own calculation.

Table 3. Socioeconomic segregation of high ESCS students using the Hutchens Index.

				Public schools			Semiprivate schools		
	H	**H_{within}**	**$H_{between}$**	**H_{public}**	**w_{public}**	**Contribution**	**$H_{semipirvate}$**	**$w_{semiprivate}$**	**Contribution**
Top tercile	0.151	0.121 (79.88%)	0.030 (20.12%)	0.118	0.728	-0.013	0.144	0.242	0.164
Top quartile	0.171	0.137 (80.25%)	0.034 (19.75%)	0.143	0.704	-0.005	0.141	0.262	0.176
Top decile	0.242	0.206 (85.04%)	0.036 (14.96%)	0.239	0.660	0.045	0.157	0.304	0.197

Note: Cut-points are established in the first column of the table. These cut-points have been defined by the Andalusian distribution of the ESCS of the students attending public fund schools in 5th grade in 2008-2009.

Source: Authors own calculation.

The tables indicate that low ESCS students are more segregated in semiprivate schools: high ESCS students are more segregated in public schools. In other words, low ESCS students tend to cluster in some, presumably lower achieving, semiprivate schools, whereas high ESCS students cluster in relatively few public schools. The fact that low ESCS students are less likely to attend semiprivate schools and high ESCS students are less likely to attend public schools, reinforces the extent of the segregation we see in the system. Although the differences are not large, in the segregation of high ESCS students across schools is greater than the segregation of low ESCS students. For example, comparing the Hutchens for students from the top and bottom ESCS deciles, the Index has a value of 0.242 and 0.203, respectively. This is also a feature of the Spanish upper secondary/high school system. More socio-economically advantaged pupils are relatively more segregated and hence more likely to be taught in separate schools with more similar peers.

As a robustness check we also use an alternative measure, namely the Dissimilarity Index[8]. The Dissimilarity Index cannot be decomposed but we can use it to calculate the level of segregation for the whole sample and for the sub-samples of public and semiprivate schools. The component of the Hutchens Index that can be compared with the Dissimilarity Index is the "within" component, although it is of course calculated differently and hence has a subtly different meaning. The value of the Dissimilarity Index for our data is higher than the Hutchens Index (around 0.25), but this is in line with previous literature (Jenkins et al., 2008; Gutierrez et al., 2017). The level of segregation of low ESCS students is greater in semiprivate schools than in public schools using the Dissimilarity Index, consistent with the Hutchens. The Dissimilarity Index also confirms that the top decile of students are particularly highly segregated in public schools.

6. Conclusions

This paper adds to the previous empirical evidence on segregation in the Spanish education system by providing the first high quality estimates (based on census data) of socio-economic segregation for primary education in Andalusia. We measure the extent of socio economic segregation of Andalusian students across primary schools, focusing particularly on the level of segregation in public and semiprivate schools. The Hutchens Index has been used given its attractive property of being able to decompose the segregation level by school type. Around 20% of the total segregation we see in the Andalusian primary education system is attributable to this sorting across school types; 80% of the segregation observed occurs due to social sorting of pupils within the two different types of school. Previous evidence, also based on the results of the Hutchens Index, for the secondary school system in Spain (Murillo et al., 2018), found that disadvantaged students are more segregated in semiprivate schools than in the public ones, in line with the results we find for primary education. We also find that the richest students (top decile of the ESCS distribution) are more segregated in public schools, suggesting there are a relatively limited number of public schools in which we see clusters of very socio-economically advantaged students.

[8] Results available on request.

Additionally, our estimates suggest a somewhat higher level of segregation in primary education in Andalusia than the one found by Murillo et al. (2018) for secondary education[9]. This may be driven by the combination of residential sorting by neighbourhood and smaller schools in the primary phase leading to higher levels of SES segregation than when pupils from multiple primary schools (and neighbourhoods) enrol together in the much larger secondary schools.

Our evidence would suggest that the system of semiprivate schools, and indeed socio-economic segregation in the public school system, is likely to be exacerbating inequalities in education achievement as a result of segregating pupils into different schools. The concentration of disadvantaged students in different schools may prevent them from developing the social references and norms of wealthier students and hinder their educational aspirations (Krüger, 2019). Moreover, grouping high income students together could prevent them from being exposed to experiences that help them understand the relevance of redistributional policies (Krüger, 2019). High ESCS families are also more likely to be active in their demands for high quality education, and a segregated education system may end up with poor quality schools in which there are clusters of more disadvantaged students who do not have the voice to demand improvements (Krüger, 2018).

What are the policy implications of this work? A particular concern is that both public and semiprivate schools receive public funding, but the latter can demand additional financial contributions from families. Although sorting of students into public and semiprivate schools only accounts for 20% of the segregation we see across the primary system as a whole, it is nonetheless an important source of segregation that is likely to be linked to families' ability to pay the extra costs associated with attending a semiprivate school. One policy response would be to eliminate this inequity. However, admission criteria unrelated to willingness to pay are one of the main sources of segregation in both public and semiprivate schools. Proximity to a school is a key criterion which determines admission. Given geographical clustering of families from a similar socio-economic background (not least due to housing costs), this automatically results in socio-economic segregation across schools. Policy might usefully consider alternative admission criteria or geographic based admission criteria that seek to ensure a more socio-economically balanced intake.

[9] Although their results are not strictly comparable with our findings, because they use a sample of Andalusian students (instead of census data).

References

Allen, R., and Vignoles, A. (2007). What should an index of school segregation measure? *Oxford Review of Education*, *33*(5), 643-668.

Ammermueller, A., and Pischke, J. S. (2009). Peer effects in European primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics*, *27*(3), 315-348.

Breen, R., and Jonsson, J. O. (2005). Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility, *Annual Review of Sociology*, 31, 223-243.

Cordero, J. M., Crespo, E., Pedraja, F. and Santín, D. (2011). Exploring educational efficiency divergences across Spanish regions in PISA 2006. *Revista de Economía Aplicada*, *19*, 57.

Cordero, J. M., Crespo, E., and Pedraja, F. (2013). Rendimiento educativo y determinantes según PISA: Una revisión de la literatura en España. *Revista de Educación* 362, 273-297. doi: 10.4438/1988-592X-RE2011-362-161.

Corral Blanco, N., Zurbano Fernández, E., Blanco Fernández, Á., García Honrado, I., and Ramos Guajardo, A. B (2012Estructura del entorno educativo familiar: su influencia sobre el rendimiento y el rendimiento diferencial. In *PIRLS-TIMS 2011. Estudio Internacional de progreso en comprensión lectora, matemáticas y ciencias, (Vol II*), 9-31. Ministerio de Educación, Cultura y Deporte, Madrid

Danhier, J. (2018). How big is the handicap for disadvantaged pupils in segregated schooling systems? *British Journal of Educational Studies*, *66*(3), 341-364.

Duncan, O. D., and Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, *20*(2), 210-217.

Epple, D. and Romano, R.E., 2011. Peer effects in education: A survey of the theory and evidence. In *Handbook of Social Economics* (Vol. 1, pp. 1053-1163). North-Holland.

Elosua Oliden, P. (2013). Diferencias individuales y autonómicas en el estatus socioeconómico y cultural como predictores en PISA 2009. *Revista de Educación, 361*, 646-664.

García-Huidobro, J. E. (2010). Educación inclusiva y formación democrática. *Educación secundaria, UNICEF, Buenos Aires*.

García-Montalvo, J. (2013). Crisis, igualdad de oportunidades y resultados educativos en España: una visión retrospectiva desde PISA 2012. In *PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español,* (Vol. II), 32–66. Ministerio de Educación, Cultura y Deporte, Instituto Nacional de Evaluación Educativa, Madrid.

Gibbons, S. and Telhaj, S. (2016). Peer effects: Evidence from secondary school transition in England. *Oxford Bulletin of Economics and Statistics*, *78*(4), 548-575.

Gorard, S., and Taylor, C. (2002). What is segregation? A comparison of measures in terms of 'strong'and 'weak'compositional invariance. *Sociology*, *36*(4), 875-895.

Gutiérrez, Ó. D. M. (2014). *El rendimiento educativo del alumnado andaluz a examen* (Vol. 9). Centro de Estudios Andaluces, Sevilla

Gutiérrez, G., Jerrim, J., and Torres, R. (2017). School segregation across the world: has any progress been made in reducing the separation of the rich from the poor? UCL Institute of Education Working Paper

Harding, D. J. (2003). Counterfactual models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology*, *109*(3), 676-719.

Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, *18*(5), 527-544.

Hanushek, E. and Woessmann, L. (2011). The economics of international differences in educational achievement. In Hanushek, E., Machin, S. and Woessmann, L. (Eds.), *Handbook of the Economics of Education*, Vol. 3 (pp. 89–200). Amsterdam: Elsevier B.V.

Hutchens, R. M. (1991). Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations. *Mathematical Social Sciences*, *21*(1), 31-51.

Hutchens, R. M. (2001). Numerical measures of segregation: desirable properties and their implications. *Mathematical Social Sciences*, *42*(1), 13-29.

Hutchens, R. M. (2004). One measure of segregation. *International Economic Review*, *45*(2), 555-578.

IECA (2017). *Premature dropout rate by sex in Andalusia and Spain*. http://www.juntadeandalucia.es/institutodeestadisticaycartografia/indsoc/indicadores/1038.htm. Last time accessed: August 2018.

Ilie, S., Sutherland, A., and Vignoles, A. (2017). Revisiting free school meal eligibility as a proxy for pupil socio-economic deprivation. *British Educational Research Journal*, *43*(2), 253-274.

INE (2010). Survey of Financing and Expenditures of Private Teaching. Available at: https://www.ine.es/dynt3/inebase/es/index.htm?padre=4185

Jenkins, S. P., Micklewright, J., and Schnepf, S. V. (2008). Social segregation in secondary schools: how does England compare with other countries? *Oxford Review of Education*, *34*(1), 21-37.

Krüger, N. (2018). An Evaluation of the Intensity and Impacts of Socio-economic School Segregation in Argentina. *Understanding School Segregation: Patterns, Causes and Consequences of Spatial Inequalities in Education*, 103-122. London, Great Britain: Bloomsbury Academic.

Krüger, N. (2019). Socioeconomic school segregation as a dimension of educational exclusion: Fifteen years of evolution in Latin America. *Education Policy Analysis Archives*, *27*, 8.

Llera, R. F., and Pérez, M. M. (2012). Colegios concertados y selección de escuela en España: un círculo vicioso. *Presupuesto y gasto público*, *67*, 97-118.

Lucas, S. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects. *American Journal of Sociology, 106*, 1642-1690.

Mancebón, M., and Pérez, D. (2010). Una valoración del grado de segregación socioeconómica existente en el sistema educativo español. Un análisis por comunidades autónomas a partir de PISA 2006. *Regional and Sectorial Economic Studies*, *10*(3), 129-148.

Martins, L., and Veiga, P. (2010). Do inequalities in parents' education play an important role in PISA students' mathematics achievement test score disparities? *Economics of Education Review, 29*(6), 1016–1033. doi: 10.1016/j.econedurev.2010.05.001

MECD (2016). *PISA 2015. Informe Español. Programa para la Evaluación Internacional de los Alumnos.* Ministerio de Educación, Cultura y Ciencia. Retrieved from http://www.educacionyfp.gob.es/inee/dam/jcr:e4224d22-f7ac-41ff-a0cf-876ee5d9114f/pisa2015preliminarok.pdf

Murillo, F. J., Belavi, G., and Rodríguez, L. M. P. (2018). Segregación escolar público-privada en España. *Papers. Revista de Sociologia*, *1*(1), 1-31.

Murillo, F. J., and Martínez-Garrido, C. (2018). Magnitud de la Segregación escolar por nivel socioeconómico en España y sus Comunidades Autónomas y comparación con los países de la Unión Europea. *Revista de la Asociación de Sociología de la Educación (RASE)*, *11*(1), 37-58.

OECD (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014)*. PISA, OECD Publishing. doi: 10.1787/9789264208780-en

Palardy, G. J. (2013). High school socioeconomic segregation and student attainment. *American Educational Research Journal*, *50*(4), 714-754.

Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers college record*, *107*(9), 1999.

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

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453. doi: 10.3102/00346543075003417

Tourón, J., Lizasoaín, L., Castro, M., and Navarro, E. (2012). Alumnos de alto, medio y bajo rendimiento en Matemáticas en TIMSS. Estudio del impacto de algunos factores de contexto. *PIRLS-TIMSS 2011: Informe Español. Análisis secundario*, *2*, 187-215.

Tourón, J., López-González, E., Lizasoain Hernández, L., García San Pedro, M. J., and Navarro Asencio, E. (2018). Alumnado español de alto y bajo rendimiento en ciencias en PISA 2015: análisis del impacto de algunas variables de contexto. *Revista de Educación*, 380, 156-174.

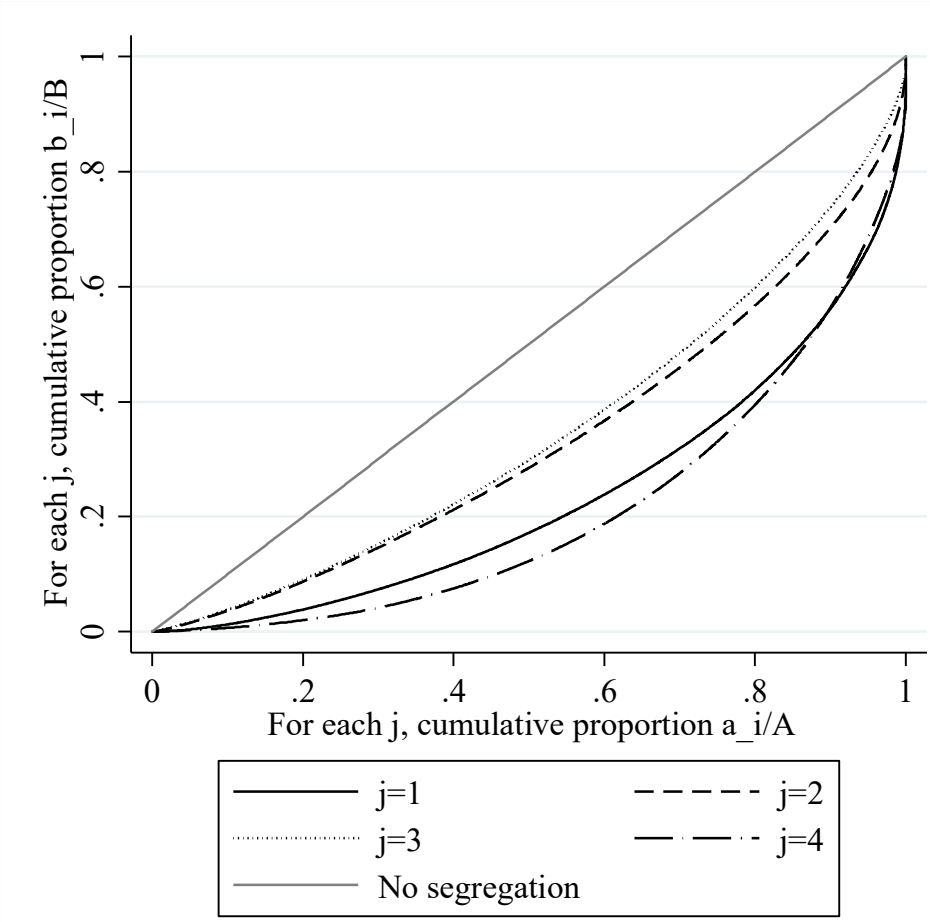
Vardardottir, A. (2013). Peer effects and academic achievement: a regression discontinuity approach. *Economics of Education Review*, 36, 108–121. doi: /10.1016/j.econedurev.2013.06.011

Willms, J. D. (1986). Social class segregation and its relationship to pupils' examination results in Scotland. *American sociological review*, *51*(2), 224-241.

Wodtke, G. T., Harding, D. J., and Elwert, F. (2011). Neighbourhood effects in temporal perspective: The impact of long-term exposure to concentrated disadvantage on high school graduation. *American Sociological Review*, *76*(5), 713-736.

Appendix.

Figure A1. Segregation curves of students across schools according to their socioeconomic level (j =first quartile, second quartile, third quartile and fourth quartile).



Source: Authors' own calculations.
Data source: Andalusian students attending 5th grade of primary education in 2008-09. Data provided by AGAEVE (Andalusian Agency of Educational Assessment).