University of the Basque Country UPV/EHU

Master in Economics: Empirical Applications and Policies

2019/2020

Master Thesis

Measuring School Income Segregation in Spain

Author:

LANDER SAN VICENTE LARRONDO

Supervisor:

CASILDA LASSO DE LA VEGA

Index

Abstract

In this work, our goal is to answer if Spanish students are differently located among schools depending on their standard of living. In other words, we analyse the segregation by income in schools in Spain.

To do so, we use data from Programme for International Student Assessment (PISA) databases. We consider all autonomous communities and types of school of Spain for year 2018. Although PISA does not collect data about the income level of students, it creates an index which measures the socioeconomic status of pupils. We use that index as a proxy of the income level for the analysis.

This index is standardized by PISA, therefore we use the Variance Separation Index (VSI) to measure the segregation, that is an invariant and absolute index. The results show that there is low income segregation level among the schools of Spain for 2018. We find the largest segregation level by income in schools in Madrid and the lowest in Cantabria. We also analyse the contribution of the type of school -public school, private-government school and private-government school- and we find that the school type may significate a big part of the school income segregation for some regions. Finally, we examine the segregation due to the language type of school in the Basque Country and we find that it represents around 28% of the total segregation.

Keywords: Segregation; PISA database; ESCS; Variance Separation Index.

Acknowledgement: We want to thank Agurtzane Lekuona for her help and support. Also, we consider convenient to thank the professor Petr Mariel for the suggestion about measuring the income segregation by the language model of schools.

1. Introduction

The Cambridge Dictionary defines segregation as “the act of keeping one person or thing separate from another person or thing”. The goal of this work is to analyse the segregation by income in Spanish schools. This topic is a recurring in papers (Rubia, 2013) and press (Hernández, 2019 for La Vanguardia) in Spain and the reason is obvious; it causes concern in the population for the negative effects it may cause, as the achievement gaps it may cause between advantaged and disadvantaged students, Owens (2018). However, this disruption has existed in the whole world for decades, see Douglas and Denton (1988) for instance.

Before delving into the analysis, there are two notions we have to understand about segregation. The first one is that there are different types of segregation, according to the criterion used to classify the groups, Reskin (1993). For instance, ethnic segregation means that demographic groups are classified according to ethnicity. Likewise, when groups are classified by gender, it refers to gender segregation. These examples have something in common, and it is that there is no natural order of groups. In other cases, individuals could be classified according to an ordered criterion, such as the educational level of the parents, depending if they have completed the primary, secondary or higher education. In these cases, it would not be correct to treat groups symmetrically.

In the same vein, we could also distinguish segregation by income. Even if this is a very important topic, there is wide disagreement about how to measure the income segregation, because the variable used to classify people is not discrete as the ones aforementioned above. There had been proposed many indices, some of them are based on ethnic segregation indices, like Jahn, Schmid, & Schrag (1947) with the Dissimilarity Index. But some other indices treated income as a cardinal variable, like the Neighbourhood Sorting Index of Jargowsky (1996). Another way to measure income segregation is to stablish a threshold or a

poverty line to create two income status groups, such as in Fong & Shibuya (2000), or to divide the groups by income percentiles, Watson (2009), but due to income is a dynamic variable, it is hard to control the distribution of the groups. And more importantly, the resulting segregation orders do not satisfy continuity, which is a basic property for any income segregation order.

In this work, we use the properties proposed by Lasso de la Vega & Volij (2019), leveraging that they “adapt the properties of standard ethnic segregation measures to the new context and investigate their implications”, instead of adapting an existing ethnic segregation to the context of income segregation. Through that method, we can see how the proposed properties characterize an absolute index of income segregation, which we introduce afterwards.

The second notion is the perception about the income segregation and its relation to income inequality. As we work about schools, lets explain this approach using schools as example:

Suppose two districts X and Y. Each of the districts have two schools A and B as follows.

District X	10€	100€
A	100	0
B	0	100

District Y	10€	1.000.000€
A	100	0
B	0	100

In both districts poor people attend school A and rich pupils attend school B, so it could be thought that in both situations the segregation by income is equal and maximum, according to the Scale Interpretability proposed by Reardon (2011). Nevertheless, the idea about income segregation we are working at is different. We not only measure the scale interpretability but also the income inequality between students. Therefore, for us in district Y the income segregation is higher as the income difference in schools is much higher. We will consolidate this idea with the axioms we will explain later. Let’s see another example about the difference of the pure segregation and the income segregation we propose along this work:

District X	200€	300€
A	100	0
B	0	100

District Y	100€	200€	300€	400€
A	50	50	0	0
B	0	0	50	50

Once again, some people could consider that there is pure and equal segregation in the schools of the two districts, as the rich and poor people are separated, Reardon (2011). However, we consider that segregation in district Y is larger, since the income is more distributed, in the same way as Denton & Massey (1988). These two examples show the main difference between the segregation by income and the ethnic segregation, and detail how our index works.

For our analysis, we use data from PISA. PISA is a worldwide study by the Organisation for Economic Co-operation and Development (OECD), which includes the evaluation of educational systems by measuring 15-year-old school pupils' performance on science, mathematics and reading of both member and non-member nations, Schleicher (2019). It was first performed in 2000 and then repeated every three years. Although there is no data about the income level of pupils' families, this programme creates the indicator of Economic, Social and Cultural Status (ESCS) index[1], that provides a comprehensive measure of student socioeconomic background according to the OECD, Rutkowski & Rutkowski (2013). In that way, they allocate a value to each student, that we use as a proxy of the income level. This value is standardized, in order to get a mean of 0 and standard deviation of 1 for the OECD countries, for that reason we use an absolute and invariant index[2].

The goal of our work is to analyse school income segregation in Spain and its regions as measured by the VSI and using the ESCS index as a proxy of income. We further examine to what extent the type of school, Public, Private-government and Private-independent schools is a source of segregation. In addition, we analyse if in the Basque Country the attendance of students to schools classified according to the language model contributes to the school segregation.

[1] Check Section 3 for more details.
[2] Check Section 2 for more details.

The work is organized as follows. The next Section introduces the index we use for the analysis and the list of axioms fulfilled by the index. After that, in Section 3 we describe the data and its characteristics. Section 4 shows the results and finally we highlight some conclusions about the analysis.

2. Variance Separation Index and Income Segregation Axioms

As we have explained in the previous section, the ESCS value is standardized by PISA, for that reason we use an absolute and invariant segregation index. This index is the Variance Separation Index (VSI) characterized by Lasso de la Vega & Volij, (2019). This index measures income segregation through the variance between schools.

A segregation index *S* defines a segregation order $\succcurlyeq$ as follows. For any two districts X,Y, $X \succcurlyeq Y$ if and only if $S(X) \geq S(Y)$.

For any district X= {c_1, …, c_k} where {c_1, …, c_k} is the list of schools in the district, the Variance Segregation Index is defined as follows

$$V(X) = \frac{1}{n_X}\sum_{c\in X} n_c(\mu_c - \mu_X)^2 \ ,$$

where n_c is the total enrolment of school c_c and μ_c is the mean income of the school. In the same way, n_X represents the total attendance of the district X and μ_X its mean income.

Now we present some properties that are desirable for an income segregation index. We begin with three fundamental axioms that transmit the idea of what a district means to be segregated. Particularly, these axioms express the idea that there cannot exist segregation unless there are at least two schools with different income distributions.

For any district X, R(X) is the district obtained from reallocating students so that the schools keep their initial enrolment while sharing the same relative income distribution. In other words, R(X) district does not have any segregation as all the schools have the same income distribution.

The first axiom *Equal Allocation Property (EAP)* demands that if students are reallocated with the purpose of all schools to have the same income distributions, the segregation does not increase. That is, EAP requires that for any district $X, X \succcurlyeq R(X)$. It does not talk about different allocations of students across income groups nor about different number of schools.

The second axiom recognizes a class of districts all of whose members display the same level of segregation. The *Equivalence of Single-School Districts (SSD)* stands that if X and Y are single-school districts, then X ∼ Y.

The next axiom deals with simple districts and egalitarian districts. Simple districts are those with no income variation within schools. That is, all pupils that attend the same school belong to the same income group, while egalitarian districts are known to have an egalitarian income distribution if all pupils have the same income. Egalitarian districts are equally segregated and less segregated than any other simple district, unless it is an egalitarian simple district.

The *Equivalence of Uniform Distribution Districts (UDD)* assumes two simple districts X and Y with the same income and number of pupils. Assume also that X has an egalitarian income distribution. Then Y ∼ X if and only if Y also has an egalitarian income distribution.

In consequence of these two axioms, the egalitarian districts and the single-school districts are equally segregated.

The next two axioms require invariance to certain changes in units of measurement. The first one expresses that changes in population that leave the relative attendances of the schools unchanged do not affect segregation. *Population Homogeneity (PH)* states that for any district X and scalar λ>0, X∼ λX.

School	*100€*	*200€*	*300€*	*400€*
A	120λ	50λ	30λ	20λ
B	30λ	30λ	30λ	30λ
C	10λ	20λ	50λ	100λ

The λX district is equally segregated as X district.

The next axiom affirms that changes in household incomes that do not differ in students' absolute income do not affect segregation. In other words, if household's income increases by the same constant amount, then the income segregation should not change. *Invariance to Uniform Income Additions (IUIA)* says that for any district X, and for any positive scalar λ, $X \sim X + \lambda$.

School	*100€+λ€*	*200€+λ€*	*300€+λ€*	*400€+λ€*
A	120	50	30	20
B	30	30	30	30
C	10	20	50	100

VSI satisfies the IUIA because it is an absolute index, as well as it satisfies the previous axioms (Lasso de la Vega & Volij, 2019). The next two axioms talk about the independence and separability. These axioms require segregation comparisons to be independent of irrelevant sub-districts.

For the first one, let's suppose a school district divided into two sub-districts. Let's also assume that a reorganization within each sub-district reduces segregation in both of them. For that, we should expect that the reorganization does not result in a higher districtwide segregation. If this would occur, we would perceive a rather perverse outcome of an otherwise well-intended policy. The next axiom requires that those outcomes never happen.

Independence (IND) requires the same population and total income for any two districts X and Y, and for any arbitrary district Z, $X \succcurlyeq Y \Leftrightarrow X \uplus Z \succcurlyeq Y \uplus Z$. This guarantees that any policy to reduce income segregation in a sub-district does not result in a higher districtwide segregation. The VSI satisfies IND. Check Lasso de la Vega & Volij (2019) for more details.

The next axiom although is similar, is different from independence. Consider now again a district composed of two sub-districts. Assume that a policy is applied to sub-district Y, transforming it into Z. Moreover, assume that the policy left attendance unchanged. The axiom affirms that even or not this policy increases the districtwide segregation, it does not depend on the segregation within sub-district X.

Separability (SEP) states that for any three districts X, Y, Z such that $n_Y = n_Z$, $X \uplus Y \succcurlyeq X \uplus Z \Leftrightarrow R(X) \uplus Y \succcurlyeq R(X) \uplus Z$.

To finish, we have an axiom that requires similar districts to have similar levels of segregation. The *Continuity (CONT)* property says: let $X= \{c_1,\ldots, c_K\}$ be a district and let $X^n=\{c^n_1,\ldots, c^n_K\}$, for n=1, 2,… be a sequence of districts such that $c^n_k \rightarrow c_k$ for k=1,… ,K. For any district Y, if $X^n \geq Y$ for all n, then $X \geq Y$, and if $Y \geq X^n$ for all n, then $Y \geq X$.

VSI satisfies these last two axioms and all previous ones, what we consider valid for our analysis. In addition, *VSI* is the only index that satisfies all the proposed axioms. Check Lasso de la Vega & Volij (2019).

3. Database

For our analysis, we use data from PISA. As mentioned in the introduction, PISA creates the ESCS index to measure the socioeconomic background of pupils, because they do not collect data about the income level. “In many countries, the quality of the education a student acquires can still best be predicted by the student’s or his or her school’s socioeconomic background”, Schleicher (2019).

PISA examines the home possessions index for the ESCS. This index is a composite score based on three indicators: highest parental occupation (HISEI), parental education (PAREDINT), and home possessions (HOMEPOS) including books in the home, PISA (2018, p. 39). According to PISA’s technical report “Scaling procedures and construct validation of context questionnaire data” (chapter 16), these are the compositions of the three indicators:

- HISEI: This indicator gathers both fathers and mother´s occupational data. This data was collected through two ISCO coded options (OCOD1, OCOD2).

- PAREDINT: Pupils answered to the questions about their parent’s education level.

- HOMEPOS: Students answered to 16 household item questions about their home, including three country-specific household items that are seen as local measures of family wealth within the country’s context. Moreover, pupils specified the amount of possessions and books at home. From those answers, there were derived five indices:
  - Family wealth possessions (WEALTH)
  - Cultural possessions (CULTPOSS)
  - Home educational resources (HEDRES)

- ICT resources (ICTRES)
- Home possessions (HOMEPOS).

HOMEPOS is a summary index of all household and possession items.

The reason to use these three indicators is that the socio-economic status is known to be based on education, occupational status and income, Schleicher (2019). Given that PISA does not provide any income measure, the existence of household items has been used as a proxy for family wealth. Assuming that, we use this index as a proxy of the income level for our analysis.

For some students who had missing data on one component, PISA imputed the missing variable. The other two variables were used to predict the missing one, using a regression by adding a random component to the predicted value. If there were more than one missing value, ESCS was not computed and it was assigned a missing value for the index. After the imputation, the three components were standardised for OECD countries with an OECD mean zero and a standard deviation of one, OECD (s.f.)

We work with the ESCS that was constructed as the arithmetic mean of the three indicators after their standardization, Avvisati et al. (2019). For that reason, we use an absolute index like the VSI in order to measure income segregation.

Spain has 17 autonomy provinces and two autonomous cities, Ceuta and Melilla. For our analysis, we use the PISA database of 2018 about Spanish schools and students. In the database we can find different variables and the answers to questionnaires, all collected in the codebook and the questionnaire files published by PISA (all links and documents listed at the bibliography section).

There is a questionnaire directed towards students and another aimed at the schools, which is answered by teachers or personal of the high school. PISA sends one of the questionnaires to some schools and other questionnaires to other ones, and after that they collect the answers and weigh them in order to cover the population of that school, region or whatever corresponds, PISA (2018).

The student's database includes data about students, but the only data we need is the ESCS index measurement of each student. The schools' database

includes more interesting variables. Here there is the list of variables we consider for the analysis:

School Variables	
Region	72400 is the value assigned for Spain and then we find the autonomous communities one by one. 72401 (Andalusia):72419 (Melilla)
STRATUM	Provides information about the type and language model of schools.
SCHLTYPE	School Ownership: Value of 1 for Private independent schools, value of 2 for Private-Government schools and 3 for Public schools
SC011Q01TA	Schooling available to students in their location: 1= two or more other schools in the area that compete for students; 2=Only one other school that compete; 3=No other school that compete for students in the area
W_SCHGRNRABWT	Adjusted school base weight
Students Variables	
ESCS	Index of economic, social and cultural status

This database works with sample weights. All our analysis is carried out using these sample weights.

4. Results

In this section we illustrate some descriptive statistics and the VSI results for Spain in 2018.

We start the analysis with the PISA indicator of Economic, Social and Cultural Status index. In Table 1 below we can find the mean and the standard deviation of the ESCS for every region.

TABLE 1: ESCS by Regions in 2018.

Region	ESCS Mean	ESCS Std. Dev.
Andalusia	-0.32	1.04
Aragon	0.14	0.98
Asturias	-0.05	1.02
Balearic Islands	-0.08	0.99
Canary Islands	-0.38	1.01
Cantabria	0.02	0.91
Castile and Leon	-0.002	0.99
Castile La Mancha	-0.24	1.06
Catalonia	0.10	0.97
Extremadura	-0.34	1.04
Galicia	-0.06	1.00
La Rioja	-0.12	1.00
Madrid	0.18	1.02
Murcia	-0.34	1.07
Navarre	0.04	1.00
Basque Country	0.14	0.88
Valencian Community	-0.18	1.03
Ceuta	-0.55	1.11
Melilla	-0.55	1.25
Spain	**-0.09**	**1.03**

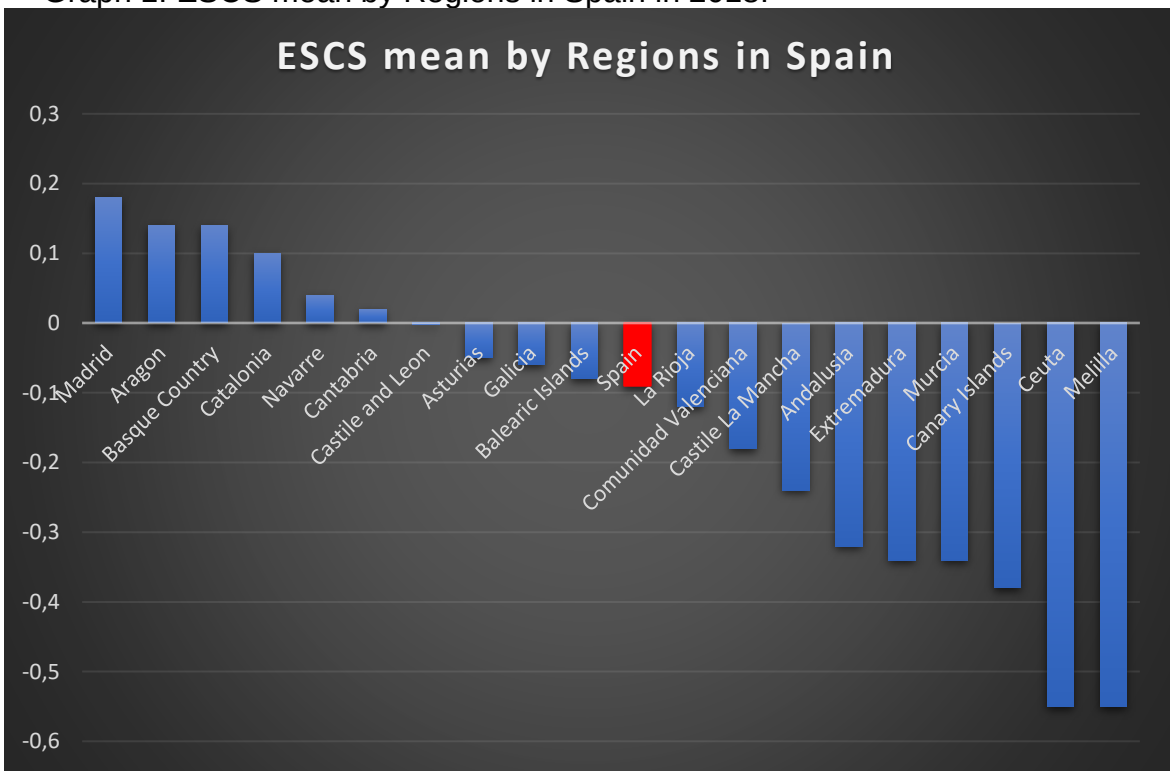
Source: Own elaboration with data from PISA (2018).

Remember that PISA standardizes the ESCS indicator so that the mean for the OECD countries takes a value equal to 0 and a standard deviation equal to 1. So, from that point on, negative values mean that the region or country is socioeconomically under the OECD mean level. The ESCS values are between -1 and 1. Hence, we can see that Spain is socioeconomically under the OECD mean level. 13 out of 19 regions have negative values (68.42%), but Ceuta and Melilla for example have a low weight, that is why Spain is still close to the zero mean. Spain also has a large standard deviation. That means that the values are not concentrated near the mean.

If we go over the regions in a general view, we can observe that all have large standard deviations. That means that all the regions have large differences between the most socioeconomic advantaged student and the less ones.

The most iconic regions, Madrid and Catalonia, have 0.18 and 0.10 positive values respectively. However, there are more regions with positive and large values, as Graph 1 shows.

Graph 1: ESCS mean by Regions in Spain in 2018.



Source: Own elaboration with data from PISA (2018).

So many regions have the ESCS mean over Spain mean, but only 6 of them have positive mean, all in the north half of the country. This could be for many reasons, such as the differences in GDP, income and employment in absolute and per capita terms, Caballero (2019). Likewise, the regions in the worst situation are in the South, or even farther like Canary Islands or Ceuta and Melilla.

Now we check whether those ESCS differences have an effect on school segregation. We use the Variance Separation Index to measure the school segregation. Table 2 below shows the results.

TABLE 2: Variance Separation Index in Spain and its regions in 2018.

Region	Variance Separation Index	ESCS indicator; Variance ESCS (total inequality)
Andalusia	0.2316 (21%)	1.0947
Aragon	0.1757 (18%)	0.9757
Asturias	0.2541 (25%)	1.0369
Balearic Islands	0.1888 (20%)	0.9673
Canary Islands	0.2242 (22%)	1.0282
Cantabria	0.1383 (17%)	0.8285
Castile and Leon	0.1629 (17%)	0.9820
Castile La Mancha	0.2207 (20%)	1.1289
Catalonia	0.2383 (25%)	0.9393
Extremadura	0.1885 (18%)	1.0637
Galicia	0.1749 (18%)	0.9962
La Rioja	0.1607 (16%)	0.9990
Madrid	0.3267 (32%)	1.0268
Murcia	0.2375 (20%)	1.1596
Navarre	0.2024 (20%)	1.0186
Basque Country	0.1627 (21%)	0.7823
Valencian Community	0.1984 (19%)	1.0394
Ceuta	0.2748 (22%)	1.2328
Melilla	0.4660 (30%)	1.5629
Spain	**0.2706 (25%)**	**1.0627**

Source: Own elaboration with data from PISA (2018).

The second column in Table 2 shows income segregation disaggregated by regions jointly with the percentage of total inequality captured by the segregation

in parenthesis. As can be seen in the table, the mean of Spain is 0.27, which does not show too many school income segregation. Madrid and Melilla have a value larger than the mean, but Melilla has a low weight to be considered in an analysis. Catalonia has a 0.24 income segregation value, that is quite low to be Catalonia. We find the lowest income segregation in Cantabria.

According to the percentages, the 25% of the total school inequality is due to income segregation. In general, all the regions are between 0.15 and 0.30 points more or less, what means the income segregation to be low. We find Madrid with the largest percentage (32%), followed by Catalonia and Asturias, with the 25% each.

The difference between the income segregation and the total inequality corresponds to the inequality within schools by region. With these results, we can conclude whether there is income inequality within schools. In this case, the variance is much higher than between schools. This means that the main source of the variance in the ESCS indicator is the within-schools. We can then conclude that students are not separated by schools by the socioeconomic status, but they are mixed among them.

The variance of Spain is almost three times larger in terms of within schools than between them. Only Madrid, Andalusia and Murcia have larger values, but two of those have a high weight in the country's population.

In the last column, we find the ESCS indicator variance by region. Through these results, we conclude the inequality by income in the regions. We use this to conclude the segregation by socioeconomic status in schools in general terms for the region.

While the ESCS variance is 1.0627, there are some regions whose variance is over that number. In first places we find the two autonomous cities Ceuta and Melilla, whose ESCS level is -0.55 in both cases. Those cities besides having the worst ESCS situations they also have the largest variation in the index. This means that there are some students whose socioeconomical situation is even worse than the mean of the region.

We can extrapolate this idea to other regions like the Basque Country, which is the region with the less ESCS variation. In this case, the total variation is 0.782, that is composed by 0.16 variation between schools and 0.62 within schools. In the case of this region, the ESCS mean is 0.14, with the lowest standard deviation. In general, in the region the socioeconomic level is better than in other regions and there is low segregation level by income between schools.

Once we have analysed the school segregation by regions, we go deeper in the analysis and check the segregation by type of schools. In Spain we distinguish three type of schools: Public schools, Private-government schools and Private-independent schools.

Our goal is to observe the distribution to conclude whether this distribution affects the segregation by income, that is, whether the type of school separates richer pupils from poorest ones. It would be understandable to think that the less income the student has, it will study in a public school. Likewise, richer pupils are expected to go to private independent schools. The following table 3 shows the student distribution in Spanish schools.

TABLE 3: Student distribution in Spanish schools in 2018:

Region	Public school	Private Government-dependence school	Private independent school	TOTAL
Andalusia	2,282,533 (82.13%)	496,697 (17.87%)	-	2,779,230 (100%)
Aragon	250,946 (66.42%)	96,352 (25.50%)	30,493 (8.07%)	377,791 (100%)
Asturias	182,116 (72%)	69,329 (27.41%)	1,506 (0.60%)	252,951 (100%)
Balearic Islands	226361 (69.86%)	72,071 (22.24%)	25,603 (7.90%)	324,035 (100%)
Canary Islands	502,752 (77.47%)	72,949 (11.24%)	73,272 (11.29%)	648,973 (100%)
Cantabria	117,143 (74.72%)	36,187 (23.08%)	3,441 (2.19%)	156,771 (100%)

Castile and Leon	434,023 (67.95%)	179,595 (28.12%)	25,096 (3.93%)	638,714 (100%)
Castile La Mancha	536,000 (83.32%)	81,256 (12.63%)	26,046 (4.05%)	643,302 (100%)
Catalonia	1,535,817 (65.24%)	680,533 (28.91%)	137,683 (5.85%)	2,354,033 (100%)
Extremadura	267,917 (81.53%)	60,691 (18.47%)	-	328,608 (100%)
Galicia	499,684 (76.56%)	125,002 (19.15%)	27,995 (4.29%)	652,681 (100%)
La Rioja	61,426 (61.67%)	38,181 (38.33%)	-	99,607 (100%)
Madrid	1,077,402 (57.08%)	462,561 (24.51%)	347,456 (18.41%)	1,887,419 (100%)
Murcia	371,228 (72.74%)	139,087 (27.26%)	-	510,315 (100%)
Navarre	127,681 (60.67%)	82,770 (39.33%)	-	210,451 (100%)
Basque Country	278,718 (50.54%)	272,777 (49.46%)	-	551,495 (100%)
Valencian Community	1,025,774 (68.72%)	358,249 (24%)	108,720 (7.28%)	1,492,743 (100%)
Ceuta	26,647 (77.02%)	7,950 (22.98%)	-	34,597 (100%)
Melilla	31,637 (88.92%)	3,941 (11.08%)	-	35,578 (100%)
Spain	**9,835,805 (70.36%)**	**3,336,178 (23.87%)**	**807,311 (5.78%)**	**13,979,294 (100%)**

Source: Own elaboration with data from PISA (2018).

This table may insinuate that in some regions there are no private-independent school, but by a long shot, that is because this database is made by sample weights. In those regions, PISA has not coincided in private-independent schools, that is why we have no data.

Looking the results obtained for the whole country, we can observe that more than the 70% of the students attend public schools, while the 23.87% go to private government-dependence schools and 5,78% to private-independent schools. We also had “no-response” data because this PISA questionnaire was answered by 15 years old students. That is why we must understand that some of them might not know exactly what type of schools they belong to. Nevertheless, we have dropped all missing data for the analysis.

The Basque Country is the region where less students attend public schools, only the 50%. In the Basque Country also, we find the larger private-government dependence school attendance, with the 49.46% of the students of the region. Whereas in Madrid, we find the highest private-independent school attending rate, 18.41%, far above the rest of regions.

Now that we know the distribution of pupils, let’s examine the ESCS by types of school and regions in Spain. It will be interesting to conclude whether the socioeconomic status is related with the school type attendance.

Table 4: ESCS by type of school in Spain in 2018.

Region	ESCS mean in Pu. Sch.	Std. Dev.	ESCS in Priv.-Gov. School	Std. Dev.	ESCS mean in Pr. Ind. Sch.	Std. Dev.
Andalusia	-0.42	1.07	0.01	0.87	-	-
Aragon	-0.15	0.99	0.26	0.90	0.51	0.86
Asturias	-0.22	1.00	0.39	0.92	1.00	0.53
Balearic Islands	-0.22	0.99	0.20	0.91	0.31	0.85
Canary Islands	-0.57	0.97	0.13	0.89	0.34	0.88
Cantabria	-0.09	0.89	0.36	0.90	-0.05	0.70
Castile and Leon	-0.11	1.00	0.19	0.94	0.45	0.82
Castile La Mancha	-0.39	1.04	0.43	0.90	0.41	0.94
Catalonia	-0.09	0.98	0.37	0.81	0.98	0.68
Extremadura	-0.48	1.02	0.02	0.99	-	-

Galicia	-0.18	1.01	0.11	0.84	0.81	0.88
La Rioja	-0.30	0.99	0.13	0.96	-	-
Madrid	-0.13	1.02	0.25	0.91	0.92	0.65
Murcia	-0.49	1.08	0.06	0.96	-	-
Navarre	-0.22	1.04	0.44	0.82	-	-
Basque Country	-0.04	0.90	0.29	0.84	-	-
Valencian Community	-0.42	1.00	0.19	0.86	0.56	0.89
Ceuta	-0.78	1.06	0.19	0.92	-	-
Melilla	-0.70	1.22	0.65	0.73	-	-
Spain	**-0.29**	**1.03**	**0.22**	**0.86**	**0.76**	**0.79**

Source: Own elaboration with data from PISA (2018).

Public schools have negative values for all regions. That means that the students of public schools are in mean socioeconomically disadvantaged in comparison with the OECD mean. For private-government schools, the ESCS results are all positive, some of them far from the zero mean. The mean for Spain is 0.22, which means that students of private-government schools tend to be socioeconomically advantaged with respect to the OECD mean.

At last, we observe that pupils of private-independent schools are much more socioeconomically advantaged, as the ESCS mean of Spain is 0.76, being 1 the maximum. Therefore, we could conclude that in mean there are socioeconomical differences between students from different type of schools.

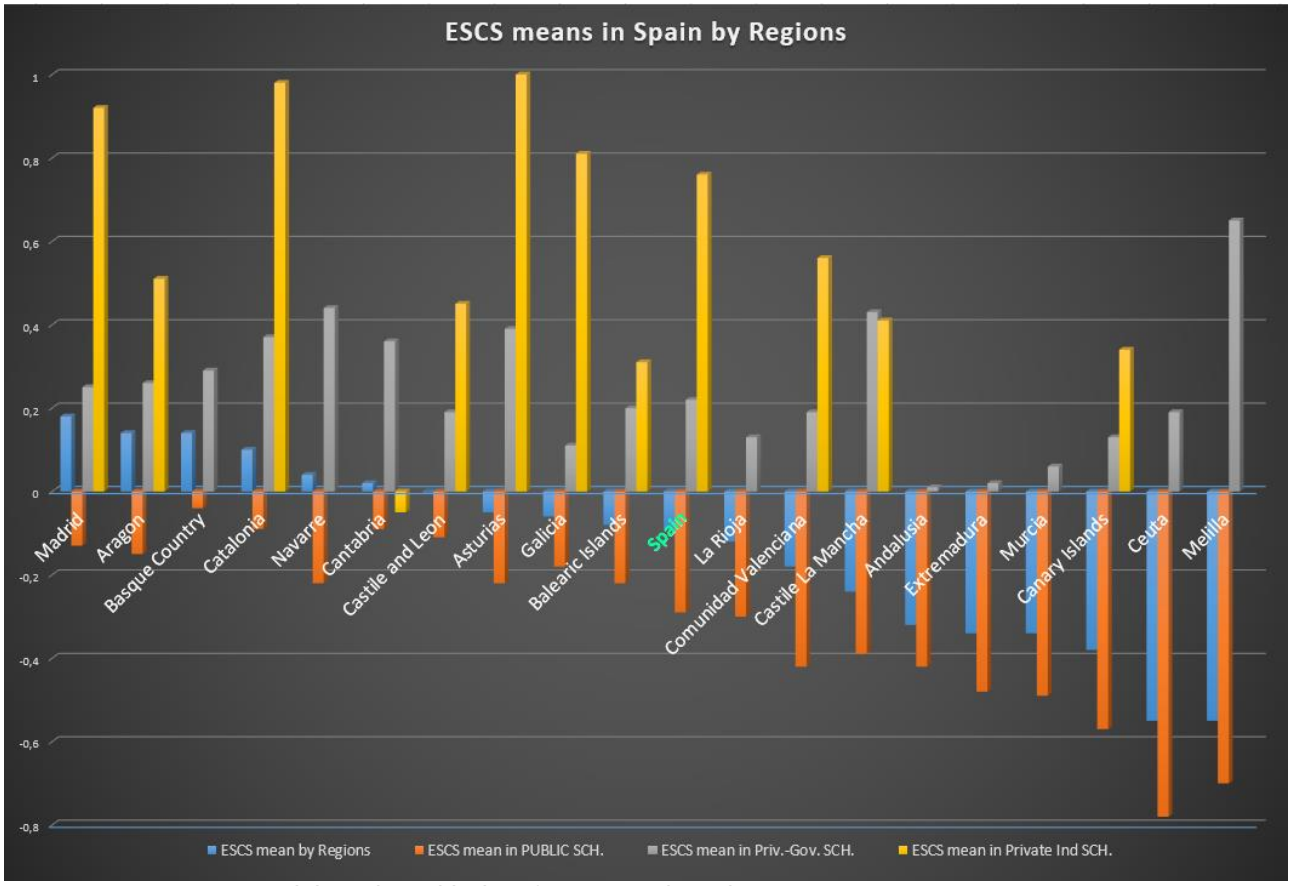
On the other hand, we observe that the standard deviation of the variance index is larger in public schools and lower in private-independent schools. This means that there is more within-school segregation in public schools than in private-independent schools.

At last, and before analysing these results through a graphic, we have to point out that Cantabria has a -0.05 ESCS mean value for private-independent schools.

There can be several reasons for this, but as we have no additional data about this issue, we cannot state anything.

In the following Graph 2, we do not only see the ESCS mean evolution but also the relation between ESCS region mean and ESCS region mean by school type.

Graph 2: ESCS mean in Spain by regions and type of school in 2018.



Source: Own elaboration with data from PISA (2018).

Blue columns are the ESCS mean by regions. That shows what we have seen in the previous Graph 1. Next to those, we find the orange columns, which represent the ESCS mean of public schools. We can observe the negative tendency they have for the less ESCS region mean. There are some exceptions like the Basque Country, Catalonia, Cantabria, Castile and Leon and Galicia. Those regions' public school ESCS mean is larger than the expected value considering the negative tendency. Nevertheless, they all still are negative values.

The next grey columns show the private-government schools' ESCS mean, which has no continuity though the graph. That is, it is not related to the region's ESCS mean. Some regions have larger values and others lowers.

At last, we find the yellow columns, that symbolizes the ESCS mean in private-independent schools. It is more complicated to conclude a tendency due to not all regions have private-independent school data. These values are expected to be the highest ones, but we can clearly see that not always is like that. In Cantabria and Castile la Mancha, the private-government school ESCS mean value is larger.

In the following Table 5 we find results of the income segregation according to the type of schools and its implication in the total Income Segregation. In other words, in the following table we find whether students are segregated by income among type of schools and how much of this segregation explains the total school segregation.

Table 5: Income Segregation between type of schools in Spain in 2018.

Region	Income Segregation according to type of schools	% of the total Income Segregation
Andalusia	0.0344	14.85
Aragon	0.0474	26.99
Asturias	0.1292	50.83
Balearic Islands	0.0421	22.30
Canary Islands	0.1179	52.58
Cantabria	0.0366	26.47
Castile and Leon	0.0284	17.47
Castile La Mancha	0.1375	62.32
Catalonia	0.0889	37.29
Extremadura	0.0463	24.56
Galicia	0.0622	35.58
La Rioja	0.0358	22.29
Madrid	0.0753	23.06
Murcia	0.0582	24.51
Navarre	0.0813	40.15
Basque Country	0.0206	12.67
Valencian Community	0.1044	52.63
Ceuta	0.1789	65.09
Melilla	0.3450	74.03
Spain	**0.0931**	**34.40**

Source: Own elaboration with data from PISA (2018).

The second column exhibits the segregation between type of schools, which are the public school, the private government-dependence school and the private independence school.

The results show that in some cases the type of school is not relevant for the segregation between type of schools, while it is for other cases. Let's see that in regions as the Basque Country, Andalusia or Castile and Leon the segregation by type of school does not even suppose the 20% of the income school segregation. However, in Valencian Community, Castile La Mancha or Asturias it signifies more than the 50% of the total income school segregation.

To finish with the analysis, we test whether the language and the type of school are relevant in the income segregation for the Basque Country. We do this analysis for the Basque Country because the language may be significant for students and their parents when deciding where they study. In the Basque Country the culture is very important and there are two types of schools in respect of the language: AB model that mixes Spanish and Basque languages and the D model, that teaches only in Basque except the Spanish language subject. In the next Table 6 we find results about the type of schools mentioned above in the work and the type of language of the teaching.

Table 6: Type of schools and language in the Basque Country in 2018.

	Number of students	ESCS mean	Income Segregation according to language and type of schools	% of the total Income Segregation
Public, AB	24,922	-0.5846	0.0460063	0.2828
Public, D	246,422	-0.0034		
Private-government, AB	152,488	0.2429		
Private-government, D	127,663	0.3531		

Source: Own elaboration with data from PISA (2018).

In the Column 2 we find the number of students that belong to each type of school. We distinguish four different school type here, depending if they are private-government or public schools and in the language they teach. We can see that there are more students for the D model schools, being the Public D model school the one with the most of students. The Public AD has a low participation in relative terms.

The Column 3 shows the ESCS mean for each type of school. We can easily see that public schools have in mean negative values and private-government schools positive and larger values. Moreover, D model schools have larger values than AB model schools.

The next Column 4 shows the income segregation level according to language and type of schools. We can see that this value represents the 28%, as the last Column 5 shows, of the whole region's income segregation. It could be considered that is significant the type and the school language to explain the income segregation in the Basque Country.

The next and the last Table 7 shows the same results but summarized to the language of the school, without considering the type of the school.

Table 7: Language school types in Basque Country in 2018.

	Number of students	ESCS mean	Income Segregation according to language of schools	% of the total Income Segregation
AB	177,410	0.1266	0.0000151	0.0001
D	374,085	0.1183		

Source: Own elaboration with data from PISA (2018).

The Column 2 shows that in D model schools there are much more students than in AB model schools. The Column 3 displays the ESCS mean for both school models. Here the results may be confusing considering the previous Table 6, but we have to take a look to the number of students. Then, we can understand that

the AB schools have more ESCS mean. To finish, the language of the school is not relevant to explain the Income Segregation of the Basque Country.

On the one hand, these results may clash with other paper's results. The reason could be that the ESCS is not a good income proxy. On the other hand, it would be accurate to make a disaggregated analysis of the provinces of the Basque Country, but our data is not significant in that sense, so we do it for the whole region.

5. Conclusions

In this work we have analysed the school income segregation by regions and type of schools for Spain in 2018. To obtain the results, we have used data of PISA, which creates the ESCS index that we have used as a proxy for the income level of families. As the ESCS value is standardized by PISA, we needed an invariant and absolute index.

There are several papers in the literature dealing with the income segregation, but we have not adapted ethnic segregation indices to measure the income segregation as others do. We have adapted the ethnic segregation axioms to the income segregation. The index we have used for the analysis is the Variance Separation Index, that is the only index that fulfils the axioms we have adapted from ethnic segregation axioms. This index measures the variance between schools to measure the income segregation.

The results of the analysis show that there exists low income segregation. On the one hand, for Spain, only the 25% of the total inequality is explained by the inequality between schools. On the other hand, the reason of that segregation level may be the segregation between type of schools, such as for Asturias, Canary Islands, Castile La Macha or Valencian Community, where more than the 50% of the between school segregation is explained by type of school segregation.

To finish, we affirm that in Spain for year 2018, the main reason of the income inequality was within schools, that is, rich and poor pupils did not attend different schools but they were mixed among them. Moreover, the language model of the school explains 28% of the Income Segregation of the Basque Country.

For further research it could be interesting to try to understand better the ESCS values. In the PISA database there is data about the number of students that belong to socioeconomically disadvantaged homes and the competitions of schools in areas, for example.

References

Avvisati, F., Le Donné, N., & Paccagnella. (2019). *A meeting report: cross-cultural comparability of questionnaire measures in large-scale international surveys.* Measurement Instruments for the Social Sciences.

Caballero, D. (27 de 10 de 2019). De norte a sur: la España autonómica de ricos y pobres. *ABC*.

Denton, N., & Massey, D. (1988). The Dimensions of Residential Segregation. *Oxford University Press, 67*(02), 281-315.

Fong, E., & Shibuya, K. (2000). The spatial separation of the poor in Canadian cities. *Demography, 37*(04), 449-459.

Hernández, M. (28 de 03 de 2019). Segregación escolar: de la dificultad a la oportunidad. *La Vanguardia*.

Jahn, J., Schmid, C., & Schrag, C. (1947). The measurement of ecological segregation. *American Sociological Review, 12*(03), 293-303.

Jargowsky, P. (1996). *Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas.* American Sociological Association.

Lasso de la Vega, C., & Volij, O. (2019). *The measurement of income segregation.* International Economic Review, online first, https://doi.org/10.1111/iere.12466.

Musterd, S. (2005). Social and ethnic segregation in Europe: Levels, causes, and effects. *Journal of urban affairs, 27*(03), 331-348.

OECD. (s.f.). *OECD iLibrary*. Obtenido de https://www.oecd-ilibrary.org/sites/0a428b07-en/index.html?itemId=/content/component/0a428b07-en

Owens, A. (2018). Income segregation between school districts and inequality in students' achievement. *Sociology of Education, 91*(01), 1-27.

PISA, 2. (2018). *Chapter 16: Scaling procedures and construct validation of context questionnaire data.*

Reardon, S. (2011). *Measures of income segregation.* Stanford Center for Education Policy Analysis.

Reskin, B. (1993). Sex segregation in the workplace. *Annual review of sociology, 19*(01), 241-270.

Rubia, F. (2013). La segregación escolar en nuestro sistema educativo. *Forum Aragón: Revista Digital de FEAE-Aragón sobre organización y gestión educativa*, 47-52.

Rutkowski, D., & Rutkowski, L. (2013). *Measuring Socioeconomic Background.* Research in Comparative and International Education.

Schleicher, A. (2019). *PISA 2018: Insights and interpretations.* Organisation for Economic Co-operation and Development (OECD).

Schleicher, A. (2019). *PISA 2018: Insights and Interpretations.*

Watson, T. (2009). Inequality and the measurement of residential segregation by income in American neighborhoods. *Review of Income and Wealth, 55*(03), 820-844.