

Measuring local socioeconomic segregation in schools with an international large-scale survey: the case of PISA in the Belgian French community

Abstract:

The Belgian education system is deeply marked by academic but also socioeconomic segregation. Assuming that free school choice is one of the parameters contributing to segregation, the Government implemented decrees to regulate this free school choice and altering school enrolment policy. At the moment of writing, there is still an obvious lack of evaluation of the effects of these measures on desegregation. Actually, such evaluation can be quite a challenge particularly in a context of budgetary constraints. The challenge is then to proceed with available data and methods. In this framework, international large-scale surveys can offer an alternative and complementary data source to the local administrative data collections. The aim of this article is to evaluate whether PISA can allow assessing segregation and its evolution in the Belgian French-speaking community, and overcome the limits of the local data. Two statistical approaches (dissimilarity index and multilevel modelling) are consequently used to measure the evolution of segregation from 2006 to 2012 in two databases (Administrative students count and PISA).

Keywords: education, segregation index, socioeconomic background, PISA, variance partition coefficient, Belgian French-speaking community

Introduction

Since the early 2000s, the availability of the successive PISA databases has allowed comparisons of countries in terms of school performances. For the Belgian French-speaking community, it was a real let-down. The community was not only a dunce but the champion of inequalities. Actually, its average performances were systematically below the OECD mean and it presents one of the widest gaps between low and high performers (Baye et al., 2009; Danhier, Jacobs, Devleeshouwer, Martin, & Alarcon, 2014; Jacobs, Rea, Teney, Callier, & Lothaire, 2009). More subtle analyses highlighted that the community's education suffered from a structural illness: the system was deeply marked by academic but also socioeconomic segregation (Baye & Demeuse, 2008; Demeuse & Friant, 2010; Dupriez & Vandenberghe, 2004; Hindriks & Lamy, 2013; Jacobs et al., 2009; Verschelde, Hindriks, Rayp, & Schoors, 2010).

The high level of segregation comes, at least for a certain part, from the organization of the Belgian education system as a quasi-market. Such type of organization refers to a free school choice and a per capita public funding of the schools. Studies have shown that the quasi-market functioning of the education systems in Belgium fosters several types of segregation between schools, resulting in the creation of both 'ghetto' and 'sanctuary' schools ((see Demeuse & Friant, 2010 for a synthesis). There are some evidences that, while proximity plays a role in the choice of the school, other factors such as reputation, socioeconomic composition and educational offer influence the parents when they make their choice.

Moreover, the structuration of the system as a 'waterfall' reinforce it: pupils from the academic curricula with 'non-satisfactory' school results tend to switch to vocational curricula, and these pupils are more often those with lower socioeconomic status (Friant, Derobertmeasure, & Demeuse, 2008). Thus a certain hierarchy of curricula is established according to a social status linked to school success for some and school failure for others. As each school does not offer all curricula, such a mechanism

creates segregation between schools. Vocational schools draw pupils with a lower socioeconomic status that have in this way been pushed out of general education (Friant et al., 2008).

In addition, schools are in competition with each other to attract pupils, as the pupils have not only a financial value (their number determines the subsidies awarded to each institution) but also a pedagogical value based on their more or less desirable personal traits (Bernard Delvaux & Joseph, 2006). This competition results in increased socioeconomic and academic segregation between schools.

School segregation is considered, by the scientific community as well as by the government, as a major problem that has to be tackled. From a philosophical point of view, it is not desirable to have an education system that separates children according to social group, be it socioeconomically or ethnically defined. From a pedagogical point of view, there is a scientific debate on the negative effects of segregation. Although no consensus has been reached, numerous studies have shown that composition plays an important role in a pupil's learning. In their literature review, van Ewijk and Slegers (2010) sum up three categories of explanations. Next to statistical misspecifications, compositional effect can result from direct peer interactions (discussions, motivation, disruptions or, for ethnic composition, tensions between races or language difficulties), teachers' practices (adjustments in teaching style or expectations) and school quality (problems in human resources management or funding).

Regulating the school choice

In its 'contract for school' of 2005 (Cfwb, 2005), the French-speaking community government identified segregation as one of the four major problems of its education and set the slogan 'No to ghetto schools' as one of its ten top priorities. Assuming that free school choice is one of the parameters contributing to segregation, the Government thought it should be possible to reduce school segregation by regulating

this free school choice and thus altering school enrolment policy. The regulation of enrolment procedures consequently began for the first year of the secondary education in 2008-09. A 3-year period of turmoil followed, with the promulgation of a new decree each year. This was the consequence of important conflicts between actors of civil society and political parties but also of a bad reception by a part of public opinion, (amongst others, some parents quite present in the media) of any measure they interpreted as an unacceptable reduction of their freedom of choice (see Ryelandt, 2013).

The first idea was to apply a 'first-come, first-served' principle. The schools had to keep a register of available spots and requests for enrolment, keeping track of each request in order of arrival, starting from a date known to everyone. This 'enrolment law', produced spectacular effects in the media, analysed in depth by Delvaux & Maroy (2009).

A 'social mixing decree' was applied the following year, introducing, in the few highly popular schools in which demand exceeded offer, some priority rules according to a socioeconomic index (SEI) and distance travelled, plus a random draw for the extra demand. This idea of random drawing fuelled a feeling of injustice for some parents. Moreover, strategies of multiple registrations caused major problems and dramatically reduced the chances of satisfying parents' preferences.

The enrolment procedure that is in application at the time of writing was set up in 2010. In order to correct the problems of the previous version, random drawing was abandoned. Parents' preferences are maximised by asking them to rank several schools, and enrolment management is centralized by a commission to avoid multiple registrations. These procedures are still contested by some parents, but its application since the start of the school year 2010-2011 has been carried out without major problems.

A posteriori, one can see two distinct objectives in this educational policy. The

first was to really enact free school choice, as it had become obvious that without clear enrolment rules, not everybody had the same opportunity to choose a particular school for their child. The second was to reduce socioeconomic segregation between schools. There seem to have been some kind of confusion between these two different objectives in the debates surrounding the regulation of enrolment procedures. While evaluating the attainment of the first goal does not per se require empirical data analysis (the existence of fair rules of enrolment is sufficient), the attainment of the second goal is far more complex and difficult to evaluate. Although it can be argued that it is too soon to evaluate these effects and that a spectacular desegregation effect cannot be expected in short terms, there is still an obvious lack of evaluation of the effects of these measures on desegregation.

In this evaluation, we think that two dimensions should be taken into account. The first one is the relevance (Bouchard & Plante, 2002; Demeuse, Demierbe, & Friant, 2011), that's to say the conformity link between its objectives and the needs of the system. The second one is the effectiveness (Bouchard & Plante, 2002; Demeuse et al., 2011), that's to say the conformity link between the objectives and the results. In this paper, we propose to give some tools in order to evaluate the effectiveness of this policy. Did this policy really reduce socioeconomic segregation between schools? As Gorard, Taylor and Fitz (2003) have shown, this is not a simple question to answer. According to the method used, there may be several biases that will lead to opposite conclusions. It is therefore extremely important to have the best possible measurement of school segregation.

Moreover, such measure can in practice be quite a challenge particularly in a context of budgetary constraints. Indeed, there is little chance that government could deploy new tools to measure the decrees' effects. The challenge is then to proceed with available data and methods. In this framework, international large-scale surveys can offer an alternative and complementary data source to the local administrative data

collections. The aim of this article is to evaluate whether PISA can allow assessing segregation and its evolution in the Belgian French-speaking community, and overcome the limits of the local data.

Methodology

Like Delvaux (2005, p. 276), we define school segregation as the *spatial separation* of students endowed with characteristics differently valued by the society. Let us briefly develop this definition. Separation can take different forms. Massey & Denton (1988) proposed five dimensions to study residential segregation: evenness, exposure, concentration, centralization and clustering. Although these dimensions remain pertinent to study school segregation and can help to understand the available indexes, it is the issue of evenness which we address in this article. This is, indeed, this dimension that has been mainly approached in the French-speaking community.

Then, the *characteristics differently valued* can be of different kinds as can be the consecutive segregations. From an equity point of view, it seems to be pertinent to introduce characteristics which the individual cannot escape (Baye et al., 2005), as ethnicity or socioeconomic background. If these phenomena are closely linked, at the very least in Belgium, the French-speaking government chose to tackle the problem from the socioeconomic angle. Such a choice is consistent with the French-speaking tradition. Indeed, it has been shown that whereas the Flemish community largely uses references to ethnic or language characteristics, the French-speaking community focuses its actions on the basis of socioeconomic background (Jacobs & Rea, 2005). We will consequently restrict our study to socioeconomic segregation.

Segregation indexes

School evenness can be graphically represented by a Lorenz curve, namely by plotting the cumulative school proportion of students with a specific characteristic (called here, the socioeconomically disadvantaged group) against the cumulative school proportion

of students without this characteristic (the socioeconomically advantaged group). Such a graphical representation has some advantages. As long as the distributions do not intersect, it allows a simple ranking without any a priori judgement and without any loss of information. However, when the distributions cross each other, segregation curves do not provide an unique ranking anymore and decisions are needed to decide which situation are the most segregated (Allen & Vignoles, 2007; Hutchens, 2004). Moreover, because this approach is based on graphical comparisons, they become difficult when the number of curves increases.

In order to overcome such issues, several numerical indexes have been developed. Actually, selecting an index requires defining what is segregation (Massey & Denton, 1988) and assumes a measurement theory (see Hutchens, 2004; James & Taeuber, 1985). Each available index has a specific behaviour and some 'desirable' features. Consequently, different theoretical bases will produces different rankings. On the contrary, some indexes with different theoretical background will produce very similar results and rankings (Massey & Denton, 1988; White, 1986). In other words, some choices of indices will conduct to different conclusions in terms of segregation evolution while others will not. Knowledge is then necessary to skilfully select an index and the choice can be crucial (or not). Moreover, the use of only one index could not be sufficient to cover the complexity of segregation (Duncan & Duncan, 1955; Massey & Denton, 1988). Without covering the huge literature of segregation indexes, we will briefly present and take a critical look at 2 indexes used in the Belgian debate about socioeconomic segregation between schools.

The first measurement is the dissimilarity index (D), largely used to measure evenness. For example, Dupriez & Vandenberghe (2004) used it for the Belgian French-speaking community. It can be computed as follow:

$$D = \frac{1}{2} \sum_i \left| \frac{c_j}{C} - \frac{!c_j}{!C} \right| = \frac{\sum_j t_j |p_j - P|}{2TP(1 - P)} \quad (1)$$

Where p_j and j_j are respectively the proportion and the population of disadvantaged student in the j^{st} school, whilst t_j is the total enrolment of this school and P and C are the overall aforesaid proportion and population, whilst T is the total number of students. The ' $!$ ' has to be read as 'not' and refers to the complementary group of advantaged students. In the classic Duncan & Duncan's review (1955, p. 211), we read that graphically, it is the 'maximum vertical distance between the diagonal and the curve' and that it can be interpreted as the proportion of disadvantaged students who should change of school to reach an even repartition of these students amongst schools. Strictly speaking, it is the proportion of student to move without replacement (Cortese, Falk, & Cohen, 1976). A look inside the formula tells us that the weighted sum of the school deviations from the overall composition is divided by its maximum and consequently that the index varies between 0 and 1 (for maximum segregation). It furthermore tells us that the deviation from the overall proportion is linear, that is to say, there is no additional payoff for bigger departures from the overall proportion (Zoloth, 1976).

Two criticisms of to this index are worth noting. Firstly, D does not fully comply with the principle of transfer (James & Taeuber, 1985). While an exchange of student between schools with composition on either side of the overall proportion of disadvantaged students affect the index, an exchange between schools on the same side of the overall proportion does not affect it. Secondly, Gorard & Taylor (2002) credited the dissimilarity index with a 'weak' composition invariance. Actually, when the number of disadvantaged students doubles in each school, D remains constant if the number of advantaged students does not change. On the other hand, if both the numbers of advantaged and disadvantaged students change, D also varies even though the repartition of disadvantaged student remains the same. According to the authors, such a feature is problematic when a part of the advantaged students becomes disadvantaged.

Gorard & Taylor (2002) proposed to use another old client which has the advantage to be strongly compositionally invariant: the segregation index (GS). It is equivalent to the Delta index (Duncan, 1961) when we replace the size of geographic areas by the population size of schools. According to Massey & Denton (1988), it measures the concentration dimension of segregation. It has been regularly used to measure segregation in the French-speaking community (Baye et al., 2005; Demeuse & Friant, 2010). This index is computed by:

$$S = \frac{1}{2} \sum_j \left| \frac{c_j}{C} - \frac{t_j}{T} \right| = (1 - P) \cdot D \quad (2)$$

As it showed, this index can be derived from the dissimilarity index. Since the $1-P$ term is absent of the latter formula, the index does not vary if the repartition of disadvantaged students remains constant and is consequently said to be strongly compositionally invariant. Moreover, it can be interpreted as the proportion of disadvantaged students who should be exchanged to reach even repartition of them in schools. Actually, this exchange proportion was one of the derived indicators used by Cortese & al. (1976) to help interpreting the dissimilarity index. Nevertheless, the index is not bounded anymore but varies from 0 to $1-P$ (Allen & Vignoles, 2007). Finally, it is an asymmetric index: its value differs for disadvantaged students and for advantaged ones. To summarize, the derivation of the segregation index from the dissimilarity index can help interpret our results and make the compositionally invariant issue less crucial.

Variance partition component

Another approach to measure segregation is provided by variance partitioning and multilevel techniques (Goldstein & Noden, 2003; White, 1986; Willms & Paterson, 1995). In multilevel analysis, it is usual to begin with the intercept-only (or unconditional) model to observe the way the variance is distributed at the specified levels. Such a model can be expressed by the following equation:

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij} \quad (3)$$

Where Y_{ij} is the characteristic of interest, γ_{00} is the grand mean of students' reading performances, u_{0j} the school deviation from the grand mean and r_{ij} the individual deviation from school mean. It allows us to compute the variance partition component (VPC), equally called intraclass correlation (ICC):

$$VPC = \tau_{00}/(\tau_{00} + \sigma^2), \quad (4)$$

Where τ_{00} is the school level variance and σ^2 the student level one. The VPC can be easily interpreted as the part of the total variance which is attributable to the differential recruitment of schools according to Y_{ij} . Let us note that the latter can be continuous. The VPC varies between 0 and 1 (for maximum segregation). Unlike D, it is affected by all the exchanges between schools. Moreover, it has some interesting features as the possibility to model sampling design and to add weights at each level.

Goldstein & Noden (2003) and then, Leckie et alii (Leckie, Pillinger, Jones, & Goldstein, 2012) proposed using a three level model with binary dependant variable. The variance of the school residual is considered as segregation index. No variable can be entered at the student or schools levels: such a modelling would result in measuring the ability to predict dependent variable at the school level and not segregation (Goldstein & Noden, 2004; Gorard, 2004). At the third level, variables can be entered to explain the possible differences between school areas. This feature is essential for areas comparison and avoids attributing disparities between areas to school segregation. Nevertheless, it requires sufficient number of areas, on pain of significant biases (Maas & Hox, 2005). In PISA, there is no variable available to identify Belgian school areas. Consequently, we will be unable to separate the area discrepancies and the school segregation. We model student as the first level and the whole school (administrative unit) as the second one.

<i>Database</i>	<i>2006</i>	<i>2007</i>	<i>2008</i>	<i>2009</i>	<i>2010</i>	<i>2011</i>	<i>2012</i>
Students count							
<i>SES missing rate</i>	4.16 %	3.07 %	3.38 %	3.74 %	3.14 %	3.81 %	3.35 %
<i>Student</i>	51084	52213	50751	47992	47123	46653	48020
<i>School</i>	531	529	532	502	504	503	497
PISA							
<i>SES missing rate</i>	1.37 %			1.74 %			1.94 %
<i>Student</i>	2816			2879			2778
<i>School</i>	94			96			95

Table 1: Sampling outcome

Data and socioeconomic indicators

We have at our disposal two sources of data to measure socioeconomic segregation in secondary schools of the French-speaking community of Belgium: an administrative file and an international survey. In order to delimit a subpopulation equally identifiable in both sources, we restrict our analysis to the 15-year students registered in the regular fulltime secondary schools (mainly in the 3rd and 4th grade). Given that the successive decrees regulating school choices are only implemented in the 1st grade of secondary education, such a population could be inappropriate to see any changes.

Nevertheless, the main aim of this article is to compare results from different sources and tools, not to resolve the discussion about the effect of the decrees regulating school choices. Although we are only able to measure short term effects, we can expect some change from 2012, when the first after-decree cohort reaches 15 year. Let us moreover note that the decree only resolve the allocation of places for supernumerary registrations in complete schools (Cantillon, 2013). Only limited changes are then expected.

Administrative local data

Since the 2004-05 academic year, the administration saves a set of exhaustive student data on each 15th January. This database, called 'students count', is used for distributing funding between the Belgian communities, the management of the French-

speaking education (as affirmative action) and statistics. Its access is restricted because of privacy protection but researchers can be allowed to use it for specific project and limited time. It is important to note that students count was intended for management, not statistics or analyses, and permit roughly to know where each student is in the educational system. However, while its uses have expanded, the number of variables has remained limited. For this study, we use successive students counts from 2005-06 to 2011-12.

Strictly speaking, the students count does not include any individual socioeconomic variable. Of course, there is a socioeconomic index (SEI) but this index is an aggregated measurement. Actually, within the framework of affirmative action, an SEI was computed for each Belgian statistical sector, namely the smallest Belgian administrative unit. A comprehensive socioeconomic index was initially developed on the basis of 12, then 11 variables, taking into account both the requirements imposed by the decree dated 30 June 1998 and the scientific literature that finds them reliable as indicators for academic and/or social success (see Demeuse, 2002). This synthetic factor was developed to 'cover the complexity of socioeconomic reality of sectors' (Demeuse, 2002, p. 229) and actually covers the following dimensions: income, qualifications, living conditions, occupation and employment. Once the sector indexes are computed, each student receives the value from his sector of residence. This index is a normal distribution metric variable that varies between -3.5 and +3.5. It is recalculated every three years on the basis of the latest statistical data available.

Demeuse (2002) identifies the main arguments justifying such a method for calculating the socio-economic index from the district where the pupils live and for determining the schools that will benefit on the basis of their population rather than the zones where they are located. The fact that the socio-economic index is not created from data collected directly from the pupils in the schools is because this approach was rejected by the legislator for at least two reasons. The first is related to respecting the

private life of the pupils and their parents: both the law of 8 December 1992 restricting individual collection of information about the characteristics of the family environment, and educational staff, are particularly reticent about putting on record information about pupils' socioeconomic background. The second is related to how such data are encoded: this is expensive and relatively unreliable. This solution was selected on the basis of the results of former scientific studies (Demeuse, 2002; Ross, 1983), which show that an indirect indicator of the socio-economic status "predicts' pupils' educational difficulties as well as the variables collected directly from families."

Such a procedure has some limits. Firstly, there is a problem of data availability for the sector index computation. For the last students count, some variables are quite out of date (the oldest one date back to 2001), some are only available at a widest administrative unit as the municipality and some are not available for sectors with a low population density to ensure privacy protection. Secondly, due to legal requirement, some variables that could be weakly correlated with the factor have to be kept in the model. Such a choice can confronts us with a validity issue. Thirdly, the use of data at the sector level introduces a bias. Strictly speaking, in the case of perfect socioeconomic homogeneity within the sectors, no bias would be introduced. Nevertheless, because sectors include a more or less heterogeneous population, the variance of this socioeconomic variable is artificially reduced (Bernard Delvaux, 2003). Moreover, students from homogeneous sector will be better represented by the index than students from heterogeneous ones. Finally, the index presents a noticeable part of missing values due to missing index for some sectors and errors in the process of automatic recognition of addresses. As **Erreur ! Nous n'avons pas trouvé la source du renvoi.** shows, the proportions of missingness for the ISE variable vary from 3.14 to 4.17 per cent. After listwise deletion, our final subpopulation encompasses about 50000 15-year-old students clustered in 500 schools.

Programme for International Student Assessment (PISA)

PISA is a research project led by the OECD which aims to assess the student ability “to use their knowledge and skills to meet real-life challenges.” (OECD, 2012, p. 22) This large-scale survey has been conducted every three years since 2000. The most recent round for which data is available was executed in 2012 and concerns specifically the mathematics skills. There were sixty five participating countries, representing approximately 510000 assessed 15-year-old students (grade 7 or higher).

It is possible to identify the two Belgian communities in the public database. For our purpose, we selected only the subsample for the French-speaking community in the successive PISA 2006, 2009 and 2012 files. The sampling design consists of a two-stage stratified one: schools¹ are sampled according to their size but are first separated between explicit strata (form of education², public/private dichotomy) and ordered by implicit strata (Index of Over-aged Students in 2006, National/International School, Retention Rate and Vocational-Special Education in 2009³); students are randomly sampled in selected schools to obtain 35 respondents by schools (or less if there are not enough valid 15-year-old students) (OECD, 2009, 2012).

During PISA process, data can be excluded or missed at different steps. Firstly, OECD provides exclusion rules to take off some schools and students. Secondly, it is possible that some schools and students do not participate because they refuse or are absent at the time of testing. Unweighted school participation rates (before and after replacement) are consequently computed. Let us note that only schools with at least 25 per cent respondents are included in the PISA data but cannot be distinguished through lack of information about school response rate. Finally, some students can fail to respond to some items in the questionnaire. After listwise deletion, our final samples

¹ The schools sampled in PISA are “whole schools”, namely the administrative units. Let us note however that in 2009, a part-time vocational school are not included in the linked classical school anymore (OECD, 2012, p. 77).

² Regular, special education, part-time education.

³ At the time of writing, the last technical report for the 2012 PISA round was not yet available. Some information are the provided for 2006 and 2009 only.

cover, respectively for 2006, 2009 and 2012, 2816 and 2879 and 2778 respondents in 94, 96 and 95 schools.

PISA provides an individual socioeconomic variable. This statistical construct is called the *index of economic, social and cultural status (ESCS)*. It synthesizes information from three sources: the highest level of parental occupation, the highest level of parental education and the number and kind of home possessions. OECD (2009, 2012) reports 0.68 and 0.67 reliability scores (standardised Cronbach's alpha computed with weighted samples) for Belgium in 2006 and 2009. Some slight modifications are introduced in the indexes computation in order to make them comparable across cycles.

Weighting and confidence intervals

The PISA database is provided with a set of sampling weights w_i in order to deal with the informative design. Properly speaking, informativeness is a property of a specific model in a specific design. This means that for a model including a set of variables, some variables (not included in the model) stay correlated to the outcome variable. In PISA, weights are provided in such a way, firstly, to deal with the over- and under-sampling of some strata of the population, secondly, to take the potential lack of accuracy in sampling frame into account and thirdly, to adjust for school and student non-response (OECD, 2012). Moreover, replicate weights are present in the database. They allow computing confidence intervals to summarize the uncertainty linked to the indexes we use.

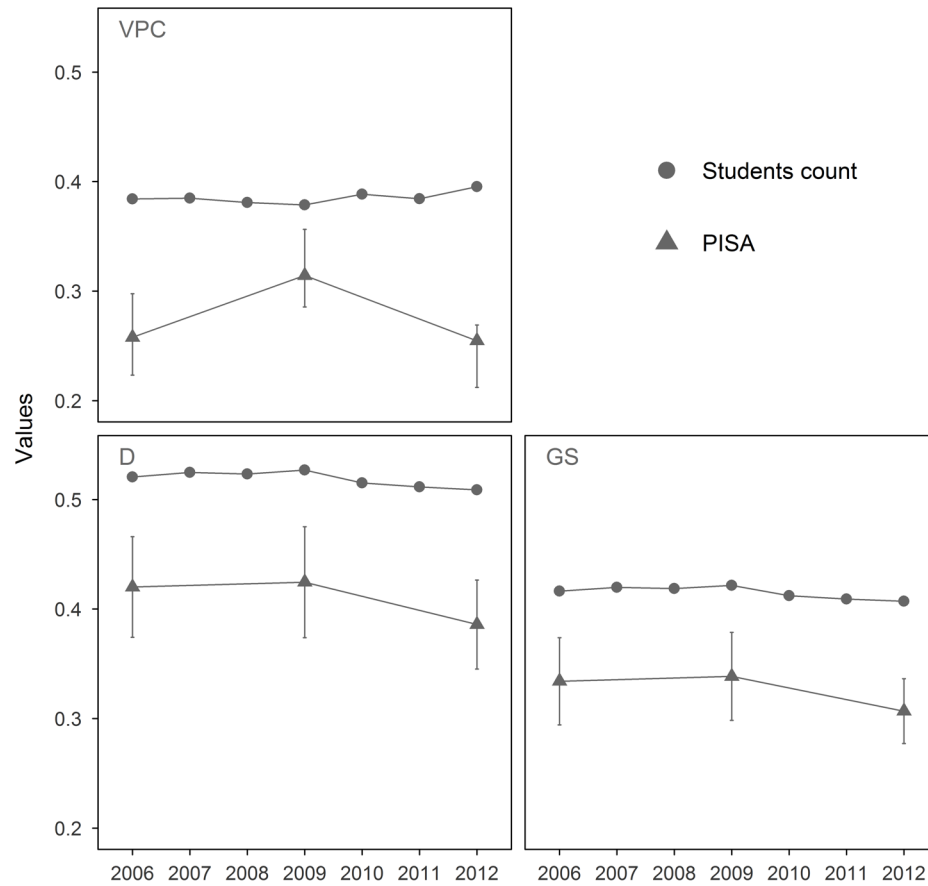
The dissimilarity index has been slightly modified to encompass weighting. In the equation 1, $t_j = \sum_i w_i$ and $p_j = \sum_i w_i |_{SESD=1} / t_j$ are computed in each school whilst $T = \sum_j t_j$ and $P = \sum_i w_i |_{SESD=1} / T$ are the parameter for the overall sample. To obtain confidence intervals, it is advised to compute sampling variance with replicate weights. Fay's method is the variant of Balanced Repeated Replication used by the OECD. Setting that ϑ^* is the estimator computed with sampling weights and ϑ_t^* the same

estimator computed with one of the eighty replicate weights from the database, the variance of ϑ^* is $1/20 \sum_{t=1}^{80} (\vartheta_t^* - \vartheta^*)^2$ (Adams & Wu, 2002).

For the VPC, conditional student weights are used at the first level and standardized ones (for the French-speaking community) at the second level. MLwiN uses method 2 for rescaling conditional student weights (Centre for Multilevel Modelling, 2011; Pfeffermann, Skinner, Holmes, Goldstein, & Rasbash, 1998). For this index, the use of replicate weights is problematic. The method requires weights at both levels but replicates weights are provided at the student level in the database. Some authors used bootstrapping to obtain confidence intervals for the VPC (Willms & Paterson, 1995)⁴. Bootstrapping requires mimicking the sampling method that produced the data to draw many resamples. When it works, the distribution of the estimators computed separately on each resample is asymptotically equivalent to the real estimator distribution. In this article, we set up a basic procedure. We assumed that school weights are the inverse of sampling probabilities and, on this basis, we recreated the whole population of schools (about 500 according to the students count) with their respective sizes. In this population, we drew 9999 independent resamples (proportionally to the school sizes) in which we computed segregation indices. Considering the estimators' distribution as the true one, we selected the 25th and the 975th percentiles as the bounds of or 95% confidence interval.

⁴ Simulation-based techniques like Markov Chain Monte Carlo (MCMC) procedures are also available to compute easily and efficiently a confidence interval for the VPC (Brown, 2012) but unfortunately, at the time of writing, weighting is not available for such procedures in MLwiN (Centre for Multilevel Modelling, 2011).

Figure 1: Segregation indexes



Results

All the analyses were executed in the R environment. Multilevel modelling were run with the R2MLwiN package (Zhang, Charlton, Parker, Leckie, & Brown, 2012). Because dichotomized variables are needed to compute segregation index, disadvantage students are arbitrarily defined as the 20 per cent of students with the lowest socioeconomic background. All the results are presented in figure 1.

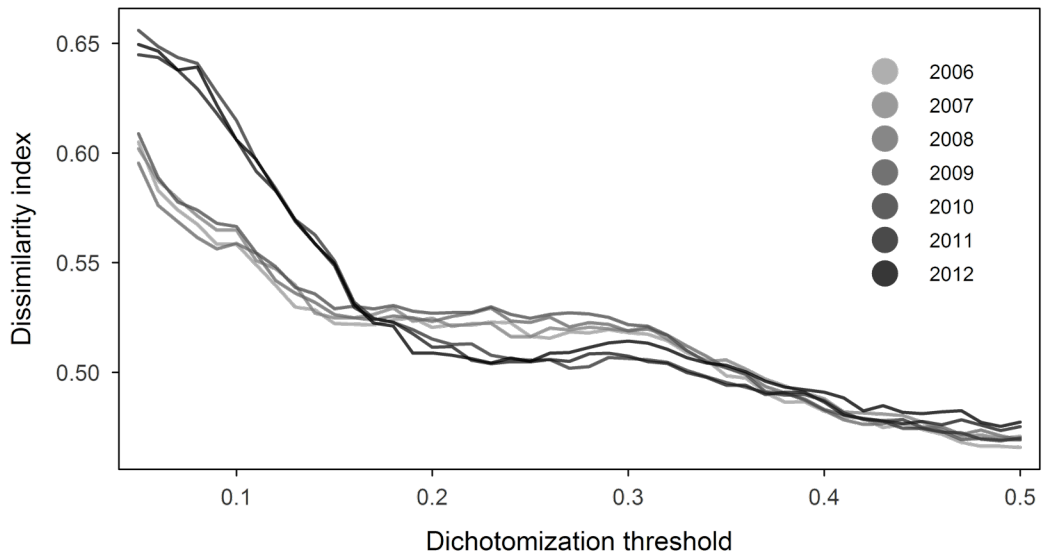
The first obvious thing is that indexes computed on one database is significantly different from those computed on the other. The values for segregation indexes based on the students count do not even fall into the confidence interval of those based on PISA. Computing segregation with the aggregated measurement of SEI at the sector

level seems to overestimate systematically the true level of segregation. Consequently, conclusion about the level of segregation observable in the schools of the French-speaking community will drastically change depending on the used database. According to the GS on the students count, 41 % of students should move to a different school to reach an even socioeconomic distribution of students between schools. According to PISA, this would concern only 31 % (95 %-CI: 28 %, 34 %).

Although databases disagree concerning the level of segregation, they could agree regarding the evolution. Here again, this is not so clear. The D index shows a slight decrease after 2009. Between 2009 and 2012, the segregation has encountered a 3.4 % relative decrease according to students count. The higher decrease of 9.1 % is PISA remains non-significant. It is more difficult to attribute such diminution to the decrees. Actually, we expected to see the first changes for the 15-year student only in the last year available (2012). Nevertheless, segregation decreases as early as 2010. Consequently, two hypotheses can be proposed. On the one hand, the media coverage given to social mix issues had effect on segregation even in grade not concerned by the decree; on the other hand, the decrease in segregation could be due to other changes.

The VPC index on the students count shows a low relative increase. Between 2009 and 2012, it gains 4.2 % highlighting an increase in segregation during this period. After a steep increase before 2009, it dramatically decreases after (18.8 % in relative terms). We are hence confronted with opposite evolution of the segregation index, depending on which database is being used. This could reflect some instability in the computation rather than real changes. The use of weight of a huge range at the second level could explain this instability. However, replication without weights shows the same pattern, although the changes from one year to another are more limited. The use of weights in multilevel modelling remains an open debate (Carle, 2009). Moreover, some outliers in 2009 could explain such increase but one "school" outliers has already

Figure 2: The choice of the threshold and its consequences on D (Students count database)

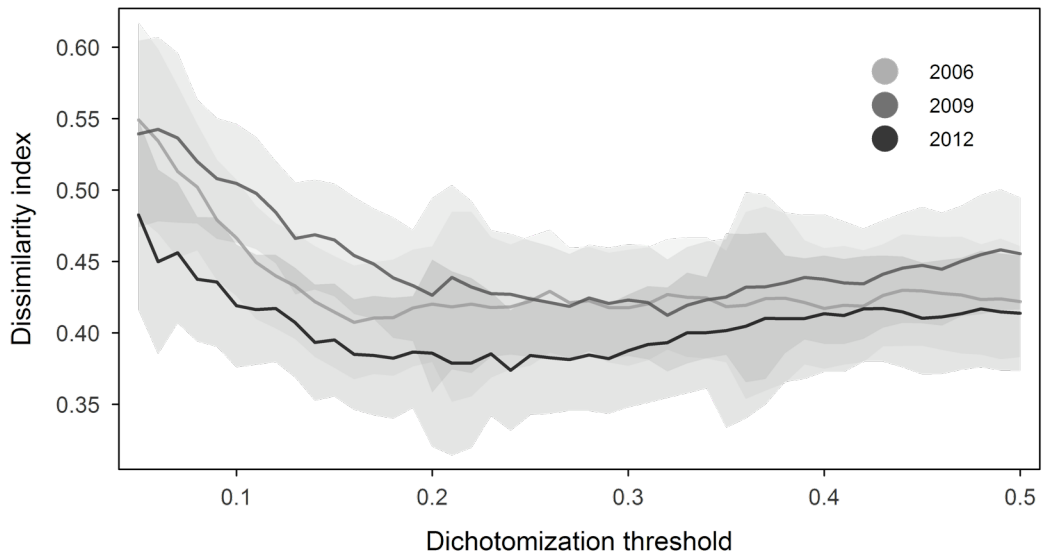


been modelled and screening of residuals plots do not show any critical problem.

The opposite direction of VPC and D indexes in students count and the behaviour of VPC on PISA is challenging. One explanation could be that both indexes are sensitive to different changes. Let us bear in mind that we have continuous variables available. The multilevel index allows us to directly use the continuous variables. On the contrary, for the dissimilarity index, we have to use a dichotomous indicator. Dichotomization of continuous variable is largely used in social sciences but raises some problems. Actually, this procedure 'alters the nature of individual differences.' (MacCallum, Zhang, Preacher, & Rucker, 2002) It has negative consequences in terms of effect size, power and reliability. From a theoretical point of view, we could argue that socioeconomic disadvantage is not continuous but categorical. This is not the exact score which counts but being above or below a specific threshold. Nevertheless, the choice of the threshold would be arbitrary.

Erreur ! Nous n'avons pas trouvé la source du renvoi. and 3 represent the variations of the dissimilarity index relative to the chosen threshold for the 7 years of the students count database and the three ones in PISA. It does not only show that

Figure 3: The choice of the threshold and its consequences on D (PISA database)



different thresholds produce different values for D but it also demonstrates that different choices could lead to different conclusions in terms of increasing or decreasing segregation. For example, we can compare what is the difference between 2009 and 2012 with the .1 and .2 thresholds. In PISA, the conclusion will be in the same direction (the segregation decrease) although not at the same extent since the difference is significant with the .1 threshold. In contrast, segregation increase in students count when we chose the .1 threshold but decrease with the .2 one. Consequently, when dichotomous indicator is used, different thresholds have to be explored. Let us note that the behaviour of VPC does not seem to be odd once we compare it to the pattern of the dissimilarity index with a .1 threshold in both databases.

Discussion and Conclusion

The evaluation question is a logical step in the cycle Conception – Implementation – Evaluation, which should give relevant, reliable and objective information about a public policy (Demeuse et al., 2011). Nevertheless, evaluating a public policy is never easy. We have seen that it is necessary to have the adequate tools in order to measure

an increase or a decrease in segregation. Unfortunately, there is still no optimal tool in French-speaking Belgium.

Here, we have shown that the choice of index and database could lead to different conclusions. Caution is then required. Such choice is crucial and will depend on the question to reply. The students count is a rich database that allows observing the evolution of segregation on different grades and in different geographic area. Nevertheless, it contains no individual socioeconomic index. PISA could then provide extra information, especially an individual socioeconomic index. Nevertheless, only 15-years students are concerned and geographic decomposition is impossible. Concerning the indices, the VPC used all the continuous information of the socioeconomic index but lead to different conclusion depending on the database. However, its computation became complicated in the case of complex sampling design. Index previously used in the French-speaking community, as Dissimilarity index (D) or Gorard's segregation index (GS), show consistent pattern across databases although this consistent is apparent. When the threshold is modified, conclusions can change and show opposite direction. With replicate weights, its computation is easy and validated. The use of a variety of index is recommended with this dichotomous solution.

In this discussion, let us take some distance from these technical considerations. There are actually nine dimensions of such a policy that could be evaluated (Bouchard & Plante, 2002): relevance, appropriateness, effectiveness, efficiency, impact, coherence, synergy, durability and flexibility. Relevance addresses the question whether the objectives address an identified need. Appropriateness considers whether external constraints are taken into account. Effectiveness asks whether the policy attained its goals. Efficiency examines effectiveness while taking into account the costs. Impact refers to the question whether the policy has unexpected effects. Coherence considers whether the implemented means are susceptible to complete the objectives. Synergy addresses whether people are coordinating with each

other in order to obtain the objectives. Durability examines whether the effects last in time. Flexibility examines whether it is possible to adapt and ameliorate the policy.

The question at the basis of this paper is actually a question of effectiveness: does the enrolment regulation procedure attain the goal of reducing socioeconomic segregation between schools? However, as we have seen in the introduction, we could also ask the question of the definition of the objectives: are the objectives well defined? Is this policy really designed to reduce school segregation? And if it is, then the question of the coherence arises: do we really hope that such an adaptation at the margin could actually have a substantial effect on socioeconomic segregation (coherence)? In other words, there are other dimensions to look at in the question of evaluating the regulation of enrolment procedures in French-speaking Belgium than only effectiveness.

There is no wonder that the task of evaluating this policy is difficult as long as its objectives are unclear. Moreover, we can hypothesize that, as the implemented policy is modest compared to some formulated objectives, the fear of disappointing results puts the brakes on the will to evaluate (Broccolichi, 2011).

Acknowledgement

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement 28360, for the EQUOP-project 'Equal opportunities for migrant youth in educational systems with high levels of social and ethnic segregation - assessing the impact of school team resources' (principal investigator: Dirk Jacobs).

The authors wish to express their gratitude to Ariane Baye, Guillaume Chauvet, Eric Deny, Jean Hindriks, Dirk Jacobs, Bart Meuleman, Christian Monseur, Catherine Vermandele and Marijn Verschelde for helpful comments and assistance.

Reference

- Adams, R., & Wu, M. (2002). *PISA 2000 Technical Report*. Paris: OECD Publishing.
- Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? *Oxford Review of Education*, 33(5), 643-668. doi:10.1080/03054980701366306
- Baye, A., Benadusi, L., Bottani, G., Bove, G., Demeuse, M., Garcia de Cortazar, M., ... Vandenberghe, V. (2005). *L'équité des systèmes éducatifs européens. Un ensemble d'indicateurs*. Liège: Service de pédagogie théorique et expérimentale.
- Baye, A., & Demeuse, M. (2008). Indicateurs d'équité éducative. Une analyse de la ségrégation académique et sociale dans les pays européens. *Revue française de pédagogie*, 165(4), 91-103.
- Baye, A., Fagnant, A., Hindryckx, G., Lafontaine, D., Matoul, A., & Quittre, V. (2009). Les compétences des jeunes de 15 ans en Communauté française en sciences, en mathématiques et en lecture. Résultats de l'enquête PISA 2006. *Cahiers des Sciences de l'Education*, (29-30), 1-248.
- Bouchard, C., & Plante, J. (2002). La qualité : Mieux la définir pour mieux la mesurer. *Les Cahiers du Service de Pédagogie Expérimentale*, (11-12), 219-236.
- Broccolichi, S. (2011). Lacunes et mésusages dans l'évaluation des politiques d'éducation prioritaire. In M. Demeuse, D. Frandji, D. Greger, & J.-Y. Rochex (Éd.), *Les politiques d'éducation prioritaire en Europe. Tome II : Quel devenir pour l'égalité scolaire ?* (p. 317-336). Lyon: ENS Éditions.
- Cantillon, E. (2013). Mixité sociale: le rôle des procédures d'inscription scolaire. In P. Maystadt, E. Cantillon, L. Denayer, P. Pestieau, B. Van Der Linden, & M. Cattelain (Éd.), *Le modèle social belge: quel avenir?* (p. 847-864). Charleroi: Presses Universitaires de Charleroi.
- Carle, A. C. (2009). Fitting multilevel models in complex survey data with design

- weights: recommendations. *BMC Medical Research Methodology*, 9(1), 9-49.
- Centre for Multilevel Modelling. (2011). *Weighting in MLwiN*. Consulté à l'adresse <http://www.bristol.ac.uk/cmm/software/support/support-faqs/weighting.pdf>
- Cfwb. (2005). *Contrat pour l'école. 10 priorités pour nos enfants*. Bruxelles: Cfwb.
- Cortese, C. F., Falk, R. F., & Cohen, J. K. (1976). Further Considerations on the Methodological Analysis of Segregation Indices. *American Sociological Review*, 41(4), 630-637. doi:10.2307/2094840
- Danhier, J., Jacobs, D., Devleeshouwer, P., Martin, E., & Alarcon, A. (2014). *Vers des écoles de qualité pour tous ? Analyse des résultats à l'enquête PISA 2012 en Flandre et en Fédération Wallonie-Bruxelles*. Bruxelles: Fondation Roi Baudouin.
- Delvaux, B. (2003). La ségrégation scolaire, phénomène avéré mais mal connu : apports et limites des données disponibles en Belgique francophone. *Cahiers du Service de Pédagogie expérimentale (Ulg)*, 15-16, 113-124.
- Delvaux, B. (2005). Ségrégation scolaire dans un contexte de libre choix et de ségrégation résidentielle. In M. Demeuse, A. Baye, M.-H. Straeten, J. Nicaise, & A. Matoul (Éd.), *Vers une école juste et efficace* (p. 275-295). Bruxelles: De Boeck.
- Delvaux, B., & Joseph, M. (2006). Hiérarchie scolaire et compétition entre écoles : le cas d'un espace local belge. *Revue française de pédagogie*, (156), 19-27. doi:10.4000/rfp.269
- Delvaux, B., & Maroy, C. (2009). Débat sur la régulation des inscriptions scolaires en Belgique francophone: où se situent les désaccords? *Les Cahiers de Recherche en Éducation et Formation*, (68), 1-30.
- Demeuse, M. (2002). *Analyse critique des fondements de l'attribution des moyens destinés à la politique de discrimination positive en matière d'enseignement en*

Communauté française de Belgique (Thèse présentée en vue de l'obtention du titre de Docteur en Sciences psychologiques, sous la direction de Marcel Crahay). ULg, Liège.

Demeuse, M. (2011). Peut-on vraiment évaluer les politiques d'éducation prioritaire en Europe ? In *Les politiques d'éducation prioritaire en Europe. Tome II : Quel devenir pour l'égalité scolaire ?* (p. 281-292). Lyon: ENS Éditions.

Demeuse, M., Demierbe, C., & Friant, N. (2011). Quelle évaluation des politiques d'éducation prioritaire en Europe ? In D. Frandji, D. Greger, & J.-Y. Rochex (Éd.), *Les politiques d'éducation prioritaire en Europe. Tome II : Quel devenir pour l'égalité scolaire ?* (p. 293-316). Lyon: ENS Éditions.

Demeuse, M., & Friant, N. (2010). School segregation in the French Community of Belgium. In *International perspectives on countering school segregation* (p. 169-187). Antwerpen/Apeldoors: Garant.

Duncan, O. D. (1961). *Statistical geography: Problems in analyzing areal data*. Free Press.

Duncan, O. D., & Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, 20(2), 210-217. doi:10.2307/2088328

Dupriez, V., & Vandenberghe, V. (2004). L'école en Communauté française de Belgique : de quelle inégalité parlons-nous ? *Les cahiers de recherche du Girsef*, 27. Consulté à l'adresse <http://www.uclouvain.be/307046.html>

Friant, N., Derobertmeasure, A., & Demeuse, M. (2008). Les liens entre recherche prospective, description du système éducatif et pilotage: un exemple en Communauté française de Belgique. Présenté à 20e colloque de l'ADMEE-Europe - Entre la régulation des apprentissages et le pilotage des systèmes: évaluations en tension. Consulté à l'adresse <http://hdl.handle.net/2013/UMONS-DI:oai:di.umons.ac.be:2353>

- Goldstein, H., & Noden, P. (2003). Modelling Social Segregation. *Oxford Review of Education*, 29(2), 225-237.
- Goldstein, H., & Noden, P. (2004). A response to Gorard on social segregation. *Oxford Review of Education*, 30(3), 441-442. doi:10.1080/0305498042000260539
- Gorard, S. (2004). Comments on 'Modelling social segregation' by Goldstein and Noden. *Oxford Review of Education*, 30(3), 435-440. doi:10.1080/0305498042000260520
- Gorard, S., & Taylor, C. (2002). What is Segregation? A Comparison of Measures in Terms of 'Strong' and 'Weak' Compositional Invariance. *Sociology*, 36(4), 875-895. doi:10.1177/003803850203600405
- Gorard, S., Taylor, C., & Fitz, J. (2003). *Schools, Markets and Choice Policies*. London: Routledge.
- Hindriks, J., & Lamy, G. (2013, septembre 9). Retour à l'école, retour à la ségrégation. Itinera Insitute.
- Hutchens, R. (2004). One Measure of Segregation. *International Economic Review*, 45(2), 555-578. doi:10.1111/j.1468-2354.2004.00136.x
- Jacobs, D., & Rea, A. (2005). Construction et importation des classements ethniques. *Revue européenne des migrations internationales*, 21(2), 35-59. doi:10.4000/remi.2487
- Jacobs, D., Rea, A., Teney, C., Callier, L., & Lothaire, S. (2009). *L'ascenseur social reste en panne*. Bruxelles: Fondation Roi Baudouin.
- James, D. R., & Taeuber, K. E. (1985). Measures of Segregation. *Sociological Methodology*, 15, 1-32. doi:10.2307/270845
- Leckie, G., Pillinger, R., Jones, K., & Goldstein, H. (2012). Multilevel Modeling of Social Segregation. *Journal of Educational and Behavioral Statistics*, 37(1), 3-30. doi:10.3102/1076998610394367

- Maas, C. J. M., & Hox, J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1(3), 86-92.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19-40. doi:10.1037/1082-989X.7.1.19
- Massey, D. S., & Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, 67(2), 281-315. doi:10.1093/sf/67.2.281
- OECD. (2009). *PISA 2006 technical report*. Paris: OECD Publishing.
- OECD. (2012). *PISA 2009 technical report*. Paris: OECD Publishing.
- Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H., & Rasbash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(1), 23-40.
- Ross, K. N. (1983). *Social area indicators of educational need*. Hawthorn, Vic: Australian Council for Educational Research.
- Ryelandt, N. (2013). *Les décrets « inscriptions » et « mixité sociale » de la Communauté française*. Bruxelles: CRISP.
- Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134-150. doi:10.1016/j.edurev.2010.02.001
- Verschelde, M., Hindriks, J., Rayp, G., & Schoors, K. (2010). Explaining social segregation in Belgium: an index decomposition approach.
- White, M. J. (1986). Segregation and Diversity Measures in Population Distribution. *Population Index*, 52(2), 198-221. doi:10.2307/3644339
- Willms, J. D., & Paterson, L. (1995). A multilevel model for community segregation*. *The Journal of Mathematical Sociology*, 20(1), 23-40.

doi:10.1080/0022250X.1995.9990150

Zhang, Z., Charlton, C. M. J., Parker, R. M. A., Leckie, G. B., & Brown, W. J. (2012).

R2MLwiN (Version v0.1). University of Bristol: Centre for Multilevel Modelling.

Zoloth, B. S. (1976). Alternative Measures of School Segregation. *Land Economics*,

52(3), 278-298. doi:10.2307/3145527