Institut d'Economia de Barcelona

IEB Working Paper 2018/14

# CLASS COMPOSITION EFFECTS AND SCHOOL WELFARE: EVIDENCE FROM PORTUGAL USING PANEL DATA

João Firmino

**Human Capital**

IEB Working Paper 2018/14

# CLASS COMPOSITION EFFECTS AND SCHOOL WELFARE: EVIDENCE FROM PORTUGAL USING PANEL DATA

João Firmino

The **IEB** research group in **Human Capital** aims at promoting research in the Economics of Education. The main objective of this group is to foster research related to the education and training of individuals and to promote the analysis of education systems and policies from an economic perspective. Some topics are particularly relevant: Evaluation of education and training policies; monetary and non-monetary effects of education; analysis of the international assessments of the skills of the young (PISA, PIRLS, TIMMS) and adult populations (PIAAC, IALS); education and equality, considering the inclusion of the disabled in the education system; and lifelong learning. This group puts special emphasis on applied research and on work that sheds light on policy-design issues. Moreover, research focused in Spain is given special consideration. Disseminating research findings to a broader audience is also an aim of the group. This research group enjoys the support from the **IEB-Foundation.**

The **Barcelona Institute of Economics (IEB)** is a research centre at the University of Barcelona (UB) which specializes in the field of applied economics. The **IEB** is a foundation funded by the following institutions: Applus, Abertis, Ajuntament de Barcelona, Diputació de Barcelona, Gas Natural, La Caixa and Universitat de Barcelona.

Postal Address:
Institut d'Economia de Barcelona
Facultat d'Economia i Empresa
Universitat de Barcelona
C/ John M. Keynes, 1-11
(08034) Barcelona, Spain
Tel.: + 34 93 403 46 46
ieb@ub.edu
http://www.ieb.ub.edu

**IEB Working Paper 2018/14**

# CLASS COMPOSITION EFFECTS AND SCHOOL WELFARE: EVIDENCE FROM PORTUGAL USING PANEL DATA

João Firmino

ABSTRACT: Using student-level longitudinal data of 6th graders I estimate class composition effects impacting on individual academic achievement. The richness of the dataset allows to tackle endogeneity stemming from between and within-school non-random sorting of students and of teachers and other confounding factors through the inclusion of many control covariates that characterize the students' cumulative process of learning and several fixed effects, namely school, teacher and cohort fixed effects. I find that increasing the percentage of high achievers, in a 6th grade class, has a negative effect on student performance. Larger shares of low-income classmates improve performance for non-low-income students. The shares of male and foreign students yield non or faintly significant results. Using the setup of a particular school, representative of the sample, I also compute improving classrooms' allocation of students by rearranging the existing students through the existing classes using the estimates of the education production function and different social welfare functions. This way I assess how the actual distribution of students across classes of a given school-grade deviate from what can be considered an improving distribution of classmates. Pareto improving allocations were not found, nevertheless utilitarian welfare functions yield marginally improving allocations.

JEL Codes: I210, I240, I28
Keywords: Class composition, peer effects, student achievement, classroom, welfare

João Firmino
Nova School of Business and Economics
Universidade Nova de Lisboa
Campus de Campolide
1099-032, Lisbon, Portugal
Email: joao.mgs.firmino@gmail.com

# 1 Introduction

Classes of students enrolled in primary to secondary schools are constantly formed every academic year across the world. From the enormous amount of possible combinations of students across classes that may occur within any given school one must be chosen by the respective school authorities so that schooling services can take place. The school agents charged with setting up classes – hence of making a final allocation choice – are likely to be constrained by specific school and education system legal restrictions, to be pressed more or less vigorously by parents' preferences regarding the classmates they want their own children to be exposed to, and the agents themselves may carry different beliefs and different social concerns of what should constitute an optimal allocation of students across classes. It is then important to understand whether different class compositions (which reflect the allocation of students across classes) cause variations in students' general cognitive achievement (arguably one of the most sought outcomes in any contemporaneous education system) in order to improve the quality of the allocation choice.

Concretely, in this paper, I address a first question of which class compositions impact 6$^{th}$ grade student general achievement and whether these affect heterogeneously different types of students. The first part of the question has been analyzed before (e.g. Hoxby, 2000a; Hanushek, Kain, Markman & Rivkin, 2003; and Yao, Ohinata, & Ours, 2016), however it has been done focusing in just a few specific dimensions of class composition at a time whereas in this paper I analyze six different dimensions simultaneously. Moreover, studies of heterogeneous effects are scarcer (Burke & Sass, 2013 is a recent reference for heterogeneous effects with respect to the specific previous achievement class compositional dimension) and this paper documents heterogeneous effects for all considered compositional dimensions.

It is important to note that the above question refers to general cognitive achievement – the outcome of interest is the average of mathematics and reading scores – and not to one particular educational outcome. Clearly, from a theoretical point of view and from the perspective of education systems that might compose different classes to different subjects, interest should not be discarded regarding the identification of class compositional effects that may be particular to a specific type of cognitive skills, say skills associated with the exact sciences from which mathematics may be a good representative, or skills related with humanities possibly representable by reading. However, given that across the world (and particularly in Portugal), education systems seem to employ as common practice the policy of composing classes to be kept constant across disciplines and given that it is reasonable to assume that such education systems (and their local school authorities in particular) favor the acquisition of overall cognitive skills by the students instead of skills of just one particular type, then the policy implications stemming from the identification of class composition effects impacting general cognitive achievement are the most interesting to all relevant actors within such education systems.

Equipped with credible causal estimates of the class composition effects I address a second question in this paper that, to the best of my knowledge, has not yet been tackled before, of whether it is possible to find welfare improving allocations of students across classes, at a school representative of the system, in comparison to its current allocation, as implied by the former estimates. Estimation of class composition effects, even if correctly identified in terms of their causality, may not be sufficient to guide class formation policies as the former only possess a *ceteris paribus* interpretation. The moment a reallocation of students is idealized based on a given *ceteris paribus* effect of a class composition dimension, it is uncertain what will be the overall welfare effect on all students of the school because any reallocation exercise has the potential to affect not only those that are reallocated, but, as well, those that were *and* will be exposed to the reallocated ones.

The answers to both questions complement each other: simulating a reallocation of students across classes hinges on the correct identification of the class composition effects, and, taking class composition effects by themselves are likely to not allow a clear perception of the overall welfare effect on the entire school-by-grade population that will occur once a reallocation of students is idealized. Clarification of this complementarity is important from a policy perspective as it enables improving the task of grouping students across classes.

This work is structured as follows. The next section covers a review of the literature. Section 3 details the Portuguese institutional context and the micro dataset. Section 4 specifies the econometric methodology to tackle potential endogeneity issues resulting from non-random allocation of students and teachers between schools and across classes within each school; it also formalizes the school planner problem and how the representative school is chosen. Section 5 presents the estimates of the class compositions effects as well as the school planner results. Finally, Section 6 concludes.

## 2 Literature Review

Class compositional effects' identification is discussed via the education production function framework. Todd & Wolpin (2003), Lazear (2001), Pritchett & Filmer (1999), Hanushek (2008), Hanushek (1979), Hanushek (1970), and even, to some extent, the earlier Coleman Report (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, et al., 1966) are examples of both empirical and theoretical studies based on the education production function framework which conveys the simple idea that the education outcome of a given student must be the result of a set of educational inputs. Simple as it may seem at first sight, the education production function framework also permits one to realize that identification of the effects of each input is likely to be hampered by correlations between them. In other words, correct identification of inputs' effects – and of class composition effects in particular – relies on not omitting any correlating educational input. In what regards the class composition effects' identification it is crucial to take in account confounding inputs, such as class size (Wößmann & West, 2006 and Bosworth, 2014) and teachers – whose within-school allocation may correlate with classes' characteristics – through teacher observables and teacher fixed effects (Hanushek, Kain, Markman & Rivkin 2003; Sund, 2009; and Burke & Sass, 2013). In fact, whether Lazear (2001) shows that for given quality level of the classmates (the fraction of time each student is able to follow the teacher without disturbing the class) increasing class size exponentially decreases class learning time, it can be argued that, similarly, fixing class size and letting the proportion of classmates with a lower level of attention to increase the same result should be expected. Adding to this that there is evidence that school authorities act in a compensatory way when setting up classes – West & Wößmann (2006) – namely by allocating students with the *a priori* weakest education inputs (low past test scores, accumulated retentions and low-income background) to the smallest classes (the authors stress that such within-school compensatory schemes are likely to exist in countries with external exams which is the case of Portugal during the period studied), then one has to recognize the potential danger from omitting class size when the interest lies on identifying class composition effects (however this relies on the assumption that class size indeed impacts on education performance which the empirical evidence does not completely support – Hoxby, 2000b – unless when very large changes in class size are introduced – Duflo, Dupas,

& Kremer, 2015). But even if school authorities have no external accountability incentives to act in a compensatory way, within-school sorting may still be at work. The school agents responsible for setting up classes may reflect their own priors relative to what is an optimal allocation of students, hence students' characteristics may determine, via school agents' beliefs, the composition of their classes (see Collins & Gan, 2013).

Furthermore, one needs to also take in consideration that current education outcomes of a given student may not only be the result of student's present inputs, but also of those from the past. Todd & Wolpin (2003), in particular, systematizes this idea, i.e. that one should as well conceive as important inputs applied during past periods to the limiting initial point of the birth of the student. As they explain, from the moment a student is born parents' decisions may express strategic behavior – conditional on their particular preferences in terms of what they idealize their children's path should be, and on the resources available to them – which should take in account the student's initial conditions and, from then on, re-update the strategic behaviors based on the realization of the cumulative learning outcomes. This should be reflected on to which schools parents, and from a certain age the students themselves, opt to enroll, and then to which classes they may pressure school authorities to enlist the students. Overall, this demand side behaviors may, in part, determine allocation of students between schools, and then between classes within a given school.[1]

In a nutshell the literature warns against non-random allocation of students and teachers between-schools and, as well, between-classes within-schools. Whether some studies have made use of first differencing, instrumental variables or school fixed effects (or a combination of these) to properly identify class compositional (or class size) effects, e.g. Hoxby (2000a), Hoxby (2000b), Akerhielm (1995), Jürges & Schneider (2004), Wößmann & West (2006), and West & Wößmann (2006); this study makes use of an extensive set of students' characteristics, past education outcomes, and of combinations of fixed effects (at the school, teachers, and cohort levels) to overcome endogeneity. Thus, this study is methodologically closer to Hanushek, Kain, Markman, & Rivkin (2003), Sund (2009), and Burke & Sass (2013).

Most of the empirical results found in the literature relate to the estimates of the impact of the proportion of classmates with either a given level of ability or of predetermined achievement level. Hoxby (2000a), with results at the grade level, reports that the performance of a given student increases between 0.1 to 0.5 points if the cohort average exam score increases by 1 point. Hanushek et al. (2003) and Sund (2009) point to gains in achievement by the average student from having peers with higher levels of prior mean achievement. Burke & Sass (2013), in turn, point to gains for a given student in having better peers, but not too much better ones, thus suggesting a break of monotonicity between the relation of own achievement and that of the peers.

Related with the previous results are those on tracking policies, i.e. on grouping students with the same level of ability or previous years' test scores, which basically set the proportion of top achieving classmates close to 100% for the top achieving students and to 0% for the low achieving ones. Specifically, in Sacerdote (2011) half of the surveyed research indicates that the impact of tracking students is positive. Duflo, Dupas, & Kremer (2011) by means of a randomized experiment in Kenyan schools besides providing evidence that increasing the percentage of top achieving classmates had a positive influence on all types of students they also observed that sorting students

[1] Note that geographic segregation of parents in terms of education, income, and other socio-economic variables reinforces the idea that students are not randomly allocated across schools.

homogeneously, i.e. tracking them, across classes caused all types of students to perform better. That is, there has been relatively recent literature that seems to support the view that grouping students of similar academic levels has a stronger positive impact that surpasses that of depriving the low achievers, in particular, from the positive effect of interacting with the best peers. The hypothesis being that low achievers benefit, in absolute value, more from better tailored teaching than what lose from not being exposed to the latter.

With respect to other dimensions of class composition, Hoxby (2000a) reports that, at the grade level, (i) a larger share of females improves performance in general for both males and females, and (ii) peer effects are stronger and positive within racial groups. The latter point is challenged by Card & Giuliano (2016) who find that top achievers from minorities actually profit from being allocated to classes dominated by other top achievers with no minority status. Otherwise, negative peer pressure within minorities is likely to lead minorities' top achievers to underperform. Hanushek et al. (2003), in turn, found evidence contrary to the belief that larger shares of low-income peers adversely affect achievement of a given student. Nevertheless, they doubt their results as being robust justifying that eligibility to a reduced-price lunch may not be a reliable measure of income differences.

## 3 Institutional Context and Data

### *3.1 Institutional Context*

The Portuguese public educational system groups the 12 sequential mandatory[2] grades that it contains in cycles across four main types of schools (see Table A.1). The 1st cycle begins for pupils aged 5 or 6 years-old at grade 1 and runs up to grade 4, and the 2nd cycle comprises grades 5 and 6. These first two cycles compose primary education or ISCED 1. Lower secondary education (ISCED 2) is the 3rd cycle – grades 7 to 9 – and upper secondary education (ISCED 3) encompasses grades 10 to 12. In turn, type-1 schools offer only the 1st cycle, whereas type-2 schools offer the 2nd and 3rd cycles, type-3 schools the 3rd cycle jointly with upper secondary, and type-4 merely upper secondary.

Classes are usually kept unchanged within a given cycle by schools, which means that purposeful sorting of students within-schools across classes should occur (if such procedure is indeed a reality in any given school) at the beginning of each cycle. Especially if the beginning of the cycle coincides with moving to a new school. These are the cases of students enrolling in grade 1 (always in a type-1 school), in grade 5 (always in a type-2 school), in grade 7 if moving from a type-2 to a type-3 school (otherwise the student usually remains in the type-2 school with the previous class from grades 5 and 6), and in grade 10 if moving from a type-2 to a type-4 school (otherwise the student usually remains in the type-3 school with the previous class from grades 7, 8 and 9).

Students repeating a given grade are usually assigned to one of the newly promoted classes of that grade the following academic year. Although these students may go again through another moment of within-school sorting, they are a minority. Even less frequent, but nonetheless possible, is the event of students moving from one school to another within a given cycle, or even within a given grade during a given academic year (e.g. due to behavioral issues on current school, parents moving to another place within the country, or the arrival of immigrant families with school-aged children). These students theoretically may also be subject to within-school sorting at the destination school at different moments compared to the majority of the students of the system.

[2] Grades 1 to 9 were mandatory until academic year 2011/12. From 2012/13 onwards the last mandatory grade has been the 12th.

Within-school sorting of students across classes, during the time period spanned in the dataset, has been vaguely regulated by the relevant public authorities. In general, schools have been requested to express each grade-school-year specific heterogeneity of students across the respective classes, however with no concrete quantitative bounds on what means to comply with that heterogeneity. The most precise regulation states that no class should be uniquely composed of students that have just been retained in a given grade the previous academic year (they may have however experienced retentions in the past). From this it is expectable to observe variability of class composition in every compositional dimension and this is exactly shown by the corresponding standard deviations depicted in the descriptive statistics mentioned below.

Across the years that the dataset spans there has been some variation of which cycles end with a national exam or not, and whether the national exam score is a requirement for students' progression to the next cycle – high-stakes – or not – low-stakes – in which case the exams' results are a mere tool to assess the overall state of the system with no specific consequences to any given student. Table A.1 also summarizes these features and the respective scale of each type of exam which will feed the educational outcome of interest as well as an important baseline score.

Finally, the Portuguese education system contains a regular academic track and other tracks related to arts' education, vocational education and educational programs for those that failed to complete mandatory schooling levels at the expected age. The regular one is by far the most demanded. For example, DGEEC (2013, page 28) reports that in 2011-12, the percentage of students in public schools enrolled in the regular academic track in the 2nd and 3rd cycles were 99% and 90%, respectively.

### *3.2 Data*

In this paper I make use of an administrative dataset compiled by the Portuguese Ministry of Education. I use information on four cohorts of 6th grade students enrolled in public schools[3] in regular academic track classes, in continental Portugal. The cohorts refer to those with 6th grade high-stakes national exams of both mathematics and reading taken at the end of 2011/12 through 2014/15 academic years.[4] The education outcome I use is at the individual level and is the average of the standardized scores of mathematics and reading of a given student, where score standardization is calculated for each (first round) subject-year exam score distribution. The correlations between the average score with each subjects' scores are about 0.9 in both cases. I use the end of 4th grade exam scores of both mathematics and reading of each student as baseline scores.[5] Given that those students that, simultaneously, belonged to the cohort that took the 6th grade exam at end of 2014/15 *and* went through 5th grade only once, had a high-stakes national exam in their 4th grade taken at the end of 2012/13 (i.e. in a 0-100 scale), whereas all other cases of students had a low-stakes baseline score (i.e. in a 1-5 scale), I converted the 0-100 baseline scores of that particular cohort into the 1-5 scale using the stipulated conversion bands used by any grader when grading low-stakes exams.[6]

[3] According to DGEEC (2013, page 28) around 87% of the 6th grade students in continental Portugal in 2011/12 were enrolled in public schools.
[4] For students observed taking more than one 6th grade exam consecutively due to grade repetition I only use the score of the first exam. Also, high-stakes exams typically offer a second-round exam for those that could have not been physically present in the exam at the stipulated hour (e.g. due to medical reasons) or as a chance to improve the score obtained in the first-round. I use only scores obtained in the first-round of both subjects' exams as these are comparable for everyone, at the cost of dropping merely a few hundred students that had to be examined in one or both subjects at the second round.
[5] For those with two or more consecutive 4th grade exam scores (due to repeating this grade) I use only their last 4th grade score.
[6] The conversion bands are: 0-19 to 1 (lowest score); 20-49 to 2; 50-69 to 3; 70-89 to 4; and 90-100 to 5 (highest score).

The variables of interest for this study, related to the composition of the class, refer to the average class compositions students were exposed to during the 2nd cycle, that is during grades 5 and 6.[7] Knowing the 5th and 6th grade school and class membership I am able to compute, for each student at each grade's class, the percentage of classmates who: are top achievers; have no previous retentions; are male; have home access to internet; are foreigners; and live in a low-income household. Given the two values (one for each grade's class of a given student) for each class compositional dimension I then compute the average value between them. A student is considered top achiever if the baseline scores of mathematics and reading are jointly high. Specifically, if the student presents one of the following (math, reading) combinations of scores: (3, 5); (4, 5); (4, 4); (5, 4); (5, 3); or (5, 5). The group of top achievers varies somewhat across cohorts (see Table A.2) given that 4th grade exams across years also vary in type (see Table A.1) and in intrinsic difficulty. The status of having no previous retentions is given to students with no retentions recorded until the beginning of grade 5. This is accomplished by checking whether the students' age (birthdate present in the dataset), at the beginning of their 5th grade academic year (by middle September), is equal or lower than exactly 11 years-old. This threshold age marks the limiting age a student is expected to have without having repeated none of the previous grades (1st through 4th)[8]. Moreover, students are considered foreigners if either they or at least one of their parents were born in one of the Portuguese speaking countries (excluding Portugal).[9] Therefore the baseline case is composed of students born in Portugal (by far the majority) or in countries not belonging to Portuguese speaking countries.[10] Finally, a student is categorized as living in a low-income family if he received social support.[11] The percentages of male and with internet at home classmates have a straightforward computation given that the dataset supplies information on student gender and internet at home possession.

[7] At this stage I use only students with no retentions during grade 5, though they may have been retained on grade 4 or earlier. Hence, for all students in the sample there are only two values for each compositional dimension from which I compute the average compositional dimension they were exposed to: one from their unique 5th grade and another from their first 6th grade.

[8] This procedure is not perfectly accurate. Students enrolling grade 1 aged less than 6 years-old, say 5.8, and with one retention during primary schooling will still be below the threshold age of 11 (for the 5.8 example they will be 10.8 years-old) in spite of their retention, hence incorrectly be given the status of no previous retentions. The reason for choosing this way to assign the no retention status has to do with the fact that the dataset starts only reporting data from 2006/07 which allows to only follow those in the oldest cohort (taking the 6th grade exam at the end of 2011/12) with no retentions during primary schooling. And for the second oldest cohort (6th grade exam at the end of 2012/13) those with up to only one such retention. Since the retention rates in Portugal are somewhat considerable, it is relatively frequent to find students with 1 and 2 accumulated retentions during primary schooling. Using only students whose educational path is completely observed in the dataset, i.e. since their enrolment in grade 1, to perfectly pinpoint those with at least one retention between grades 1 and the beginning of grade 5 would mean to drop from the sample those with a number of retentions such that their enrolment in grade 1 is unobservable. This would constitute a mechanical sample selection capable of generating the corresponding (and unwanted) selection bias as the retention history of students is very likely to correlate with their 6th grade outcomes and it is also likely to predict to what class (and consequently to what class composition) the student is assigned at the beginning of grade 5. However, the comparison of both methods for the youngest cohort (6th grade exam at the end of 2014/15) for which students with up to 3 accumulated retentions during primary schooling can be followed up to entering the system at grade 1 reveals that the method chosen assigns a false no past retention status to less than 5% of the possible cases. This is reassuring as the level of noise introduced seems very low. Possible attenuation bias stemming from the introduced noise is likely to be much smaller and unimportant than the selection bias that would have been introduced by disregarding important fractions of students from the oldest cohorts. Another important benefit is the possibility to use four instead of just three cohorts in the econometric analysis below, not only in terms of adding to the external validity of the empirical results, but especially in terms of facilitating the estimation of teacher fixed effects which requires time to observe teachers moving across schools and/or students across teachers.

[9] The Portuguese speaking countries (excluding Portugal) are: Brazil, Angola, Cape Verde, Guinea-Bissau, Mozambique, Sao Tome and Principe, and East Timor.

[10] A third category differentiating those students born in neither Portugal nor a Portuguese speaking country would be of very small size and too heterogeneous (it would group students born in countries that are usually targeted by Portuguese emigration, e.g. France and Switzerland, hence likely to have parents born in Portugal, and students born in regions that have recently come to constitute suppliers of migrants into Portugal, such as Eastern Europe and Asia).

[11] It conveys quite similar information to students benefiting from a reduced-price lunch in the USA.

I emphasize that all class compositional measures were computed as "leave-out-percentages". The contribution of student *i* was purposefully excluded from the percentage of classmates with a given characteristic that that individual was exposed to in his class. This way the compositional measures are closer to be interpreted as peer measures with respect to student *i*. Moreover, given the objective of estimating potentially different class composition effects depending on own characteristics of the students, by proceeding with a leave-out computation of the variables of interest one avoids problematic cases where student *i* might be the sole classmate with a given characteristic, say male, which implies a strictly positive percentage of overall classmates who have that characteristic (males in the example), whereas from his perspective the percentage of such students that he faces in his class is actually zero.

Besides the class composition variables of interest and the corresponding individual level categorical variables stated above, the list of variables also contains: a categorical variable describing the highest level of completed education across both parents to which each student is exposed in the household[12]; another categorical variable stating the current job status of the parents[13]; and class size (only classes with 10 or more students are considered for analysis). Note that for the categorical individual level variables missing values were coded as another category in itself (the computation of the class compositional measures does take in account missing information, see next paragraph). Since the data characterizing students and their parents is mainly obtained via filling out forms and questionnaires at the school level directly by them, a non-response may still convey information regarding their individual heterogeneity. For example, unwillingness to show up at school or to correctly fill the entire form may be related with students or parents' cognitive ability or time preferences which, in turn, may predict how students are sorted between and within schools and students' education performance.

To circumvent the fact that some of the abovementioned categorical variables describing observed students' heterogeneity (from which the class composition ones were derived) present missing values (especially regarding baseline scores – about 25% of the original population of 6th graders per cohort), the computation of the class composition measures was allowed to be done in classes (of both grades 5 and 6) with up to one third of classmates missing information on the relevant characteristic. Classes with a larger share of missing information in a given compositional dimension were excluded from the analysis.[14] This way several students, for whom there is the full set of information, belonging to classes with relatively small fractions of missing information, i.e. classes for which the compositional measurement can still be seen as sufficiently close to what it would be under full information, are prevented to be excluded from the analysis. Nevertheless, for each compositional measure that is allowed to be computed using classes with a minority of missing information – the cases of the percentages of top achievers and with no past retentions (due to missing info on birthdate) – another corresponding binary variable was created taking

[12] Nine categories: missing value; no formal education; 1st cycle – 4th grade; 2nd cycle – 6th grade; 3rd cycle – 9th grade; 4th cycle – 12th grade; bachelor or similar; master; PhD.

[13] Three categories: missing value; employed; unemployed.

[14] To clarify, let a class be constituted by 10 classmates from which 7 report the required information regarding their baseline scores (4 coded as low achievers, 3 as top achievers). Hence, there is 30% of classmates with missing information which is below the 1/3 threshold. The (leave-out) percentage of top achievers given to each of the 4 low achievers was 50% (= 3/6) since each low achiever observes 3 top achievers in the class and this is divided by the number of students (excepting himself) with the required information, whereas the percentage given to each of the 3 top achievers was 33.3% (=2/6). To each of the 3 students not reporting baseline score information was attributed a missing value in terms of percentage of top achievers. Had the share of classmates with missing information with respect to baseline scores been larger than 1/3 and all 10 classmates would have been attributed a missing value in terms of percentage of top achievers.

value one if the student indeed belongs to such a class and zero if he belongs to a full information class. These binary variables will be used as control variables throughout the econometric analysis.

Table A.3 describes information with respect to the variables and their main descriptive statistics (under the regression sample stemming from the most preferred and demanding specification; in a nutshell this specification requires each 6th grader to have non-missing information across all relevant individual, class, teacher and school level variables; to belong to the connected set; and not being a singleton, all this to be detailed in the next section). The fact that, in many cases, opposing types of students are exposed to different mean percentages of the respective compositional dimension suggests that indeed within-school sorting of students is at work, at least to some extent. The following section will detail the identification strategy to overcome this and other likely sorting mechanisms in order to identify credible causal effects of class composition on individual student general cognitive skills.

# 4 Methodology

## *4.1 Econometric Model*

I estimate equation (1) as a benchmark model of the education production function

$$Y^{6}_{iCTSa} = \boldsymbol{CC}^{\boldsymbol{6\&5}}_{\boldsymbol{i(-i),C}}\,\boldsymbol{\beta_1} + \boldsymbol{CX}^{\boldsymbol{6\&5}}_{\boldsymbol{i,C}}\,\boldsymbol{\beta_2} + \boldsymbol{TX}^{\boldsymbol{6\&5}}_{\boldsymbol{i,T}}\,\boldsymbol{\beta_3} + \textstyle\sum_{(j,G)} \phi^{G}_{t(j)} + \sum_{G} \phi_{s^{G}a} + \boldsymbol{X}^{\boldsymbol{6\&5}}_{\boldsymbol{i}}\,\boldsymbol{\beta_4} + \boldsymbol{Y}^{\boldsymbol{4}}_{\boldsymbol{i}}\,\boldsymbol{\beta_6} + \boldsymbol{W_i}\,\boldsymbol{\beta_5} + + \varepsilon^{6}_{iCTSa} \qquad (1)$$

where $Y^{6}_{iCTSa}$ is the average of the standardized mathematics and reading scores (taken at the end of grade 6, hence the superscript) of a given student *i* who, during the 2nd cycle (grades 5 and 6), was placed in a combination of two sequential classes *C*, with a combination of two to four math and reading teachers *T*, across up to two potentially different schools *S*, and belonging to cohort *a*. The education outcome is chosen to be the average between the particular outcomes of mathematics and reading in order to capture average class composition effects that may affect the acquisition of general cognitive skills (of the sort transmitted by schools' curriculums) in opposite to compositional effects specific to the acquisition of particular types of cognitive skills.

In turn, $\boldsymbol{CC}^{\boldsymbol{6\&5}}_{\boldsymbol{i(-i),C}}$ is the vector containing the variables of interest: six distinct class compositional dimensions (described in section 3.2) that a given student *i* was exposed to, across his two (potentially different) classes *C* from his passage through the 2nd cycle (i.e. through grades 5 and 6, hence the superscript). Note that the subscript *(-i)* means that each class compositional measure was computed, as described above, in a leave-out fashion, at each of his two classes *C*; and, as well, that the compositional measures were based on predetermined characteristics of the students which is intended to avoid reflexivity bias, see Hanushek et al. (2003).[15]

A few remarks are worth regarding the choice to compute the class composition variables of interest as averages between the respective values observed across the two classes (one at each grade) of a given student. First, since the system is divided in cycles with exams assessing the evolution of students throughout each cycle arch (contrary to assessing annual evolution) it is natural to study the impacts of cycle level inputs rather than at other frequencies, say annual. Second, the correlations between the 6th and 5th grade classes' compositions are larger than 0.7 which, on one hand, is indicative that each student faces, on average, very similar classmates across both classes

[15] The class (leave-out) percentages of pupils with internet at home and with low-income status, although based on characteristics determined during the 2nd cycle, should not be subject to reflexivity bias since they depend on factors not driven by the education performance of the students.

(not necessarily the exact same classmates as similar class compositions do not strictly depend on such requirement, but on the existence of classmates sharing the same characteristics), and, on the other hand, that the inclusion of separate 6th and 5th grade class compositions on the regression would produce estimates of class compositional effects likely to be subject to multicollinearity problems. Third, from a policy perspective that acknowledges that classes across grades 5 and 6 are kept quite similar in composition, it is more straightforward to simulate reallocations of students across a representative class of both grades rather than across two "different" classes across two sequential grades.

Consequently, $\boldsymbol{\beta_1}$ is the parameter vector of interest in this benchmark model. It contains the marginal effects that map *ceteris paribus* changes in each of the class compositional measures to average changes in general cognitive skills. Correct identification of $\boldsymbol{\beta_1}$ hinges on using proper variation of the class compositional measures, that is on class compositional variation that can be confidently assumed to be independent of variation of any factor capable to correlate with both class composition and individual student general cognitive achievement. All the remaining controls in equation (1) are meant to capture such portions of class composition variation that are likely to be misleading.

The vector $\boldsymbol{CX_{i,C}^{6\&5}}$ contains class level controls. Namely, class size (the average of the class sizes that student $i$ was exposed to in his 6th and 5th grade classes) to account for the possibility that differing sizes of classes may contain systematically different compositions of classmates (West & Wößmann, 2006), and binary variables (one for each compositional dimension with missing information) stating whether student $i$ class composition measure bypassed missing information when it was computed (see details in section 3.2) to control for the possibility that classes with at least one missing value in a given compositional dimension may be classes that systematically present a tendential composition in that dimension.

In turn, the assignment of the two to four teachers each student was exposed to during cycle 2 (at least a different teacher per subject across grades 5 and 6, at most a different teacher per subject per grade) may correlate with the composition of his class. It is likely that (i) different teachers may have different preferences regarding which class to teach, i.e. regarding which class compositions they prefer to face; and (ii) different teachers may also have different levels of bargaining power within the school to get their most preferred class, say due to age, seniority or current position in the school hierarchy. To account for these confounding factors and improve the interpretation of $\boldsymbol{\beta_1}$ as if teachers were randomly allocated across classes I include the vector $\boldsymbol{TX_{i,T}^{6\&5}}$ containing binary variables decomposing a categorical variable consisting of the average age of the teachers each student was exposed to during cycle 2 divided in 5 years bands' categories[16] and teacher fixed effects $\phi_{t(j)}^{G}$ from the teachers of each grade ($G = 5, 6$) and of each subject ($j = math, reading$). The first term is meant to capture systematic teacher allocation to classes based on their age or any other characteristic correlated with it, say seniority or school hierarchical position

[16] This vector also contains a binary variable similar to those included in $\boldsymbol{CX_{i,C}^{6\&5}}$ to account for missing information relative to teacher age that was bypassed when computing the average age of the teachers to which each student was exposed to. Again, the threshold chosen was one third meaning that the average teacher age was computed only for students possessing information on two thirds of the count of teachers across reading and math and across grades 5 and 6. Since it is only possible for each student to have been exposed to 2, 3 or 4 teachers, then the average teacher age measure was only computed for students recording either 0 missings, or 1 missing out of 3, or 1 missing out of 4, in teacher age. The binary variable then takes value one if the student falls in one of the last two cases, zero if the student falls in the first case, and missing otherwise (keeping away from the regression students with too many teachers with unknown age).

(it is time-varying across four different cohorts and student-varying across different combinations of teachers that each student is exposed to, even across students of the same class) and the second captures unobserved (time-invariant) teacher specific heterogeneity that may also explain possible non-random sorting of teachers across classes.

Inclusion of current ($G = 6$) and previous ($G = 5$) school-by-cohort fixed effects $\phi_{s^G a}$ is based on the possibility that the student composition at the school level experienced by a given student from cohort $a$ may determine to some extent the class level student composition that he experienced too. If school level composition of students impacts on individual student cognitive achievement, then omitting it may be a source of bias to the correct estimation of $\boldsymbol{\beta_1}$. Moreover, given that school characteristics (such as school level composition of students) may change over time it is appropriate to interact school effects with cohort effects to capture those school-by-cohort specific factors. A student belonging to the cohort examined at the end of the 6th grade of 2011/12 might have been exposed to a school-level composition of students at school $s$ during that grade that differs from another student that went through 6th grade at that same school $s$ but during the 2014/15 academic year (e.g. school $s$ may be improving or deteriorating its position in the rank of schools of the local area – either a publicized one by the media, as it happened in Portugal during the period studied, or an informal one produced by word of mouth between local parents – which may induce different types of parents enrolling or unenrolling their children from that school across the years). A parallel argument can be made for both the school and class level student compositions each student was exposed to during grade 5 which justifies the inclusion of the 5th grade school-by-cohort fixed effects. Inclusion of $\phi_{s^G a}$ helps then to interpret the estimates of $\boldsymbol{\beta_1}$ as if students of any given cohort had been randomly allocated across schools. Similarly, it also helps to interpret the estimates as if teachers had been randomly allocated across schools, as specific schools in specific years may attract teachers with different preferences regarding the overall school environment. Other benefits of including school-by-cohort effects in the model is to control for school time-invariant factors (e.g. its geographic position coupled with possible time-invariant geographical segregation of parents may induce students with similar socio-economic characteristics to constantly flock to a given school and compose an important fraction of its student population each year), as well as cohort specific factors such as different difficulty levels of the baseline and end-line exams (the fact that baseline scores vary somewhat from cohort to cohort could compromise the comparability of what is considered a top or low baseline achiever and, consequently, could compromise the comparability of the respective class compositional measures related with the percentage of top achieving classmates, thus it is, *a priori*, important to control for such cohort specific factors as I do by including cohort effects), as these are embedded in the interaction of the fixed effects.

Another important source of bias is the likely non-random sorting of students across classes within-schools (as suggested by the descriptive statistics) and the inclusion of $\boldsymbol{X_i^{6\&5}}$, $\boldsymbol{Y_i^4}$, and of $\boldsymbol{W_i}$ is intended to alleviate it. The first term is a vector containing individual socio-economic characteristics of the students that are potentially time-varying: parents' education level and job status, low-income status, and possession of internet at home. Each categorical variable referring to what is observed by grade 6 is interacted with what is observed by grade 5 in order to capture different combinations of the categories over the two years (hence the superscript)[17]. This way a detailed socio-

[17] So, for example, one effect to be estimated refers to students presenting in both years a low-income status, another if students present a combination (low-income; non-low-income) in grades 6 and 5, and so on. The same pattern applies to all the other variables within $\boldsymbol{X_i^{6\&5}}$.

economic characterization of the students is fed into the econometric model to control for within-school purposeful allocation of students to classes based on such individual characteristics. The fact that $\boldsymbol{X_i^{6\&5}}$ embeds information of both grades 5 and 6 improves the chances to control for purposeful allocation of students across classes throughout both grades. Whether individual characteristics observed by grade 5 may predict to which 5th grade class the student was assigned to (school authorities may take those characteristics into account when setting up the classes that will tend to persist during the 2nd cycle), characteristics observed by grade 6 may predict, in turn, to which class students that happen to have to change classes during cycle 2 (due to retention, social incompatibility, or a school transfer) will end up in, or, conversely, which type of incoming classmates will students (that are placed in the receiver class) get. The second vector contains three individual level education outcomes recorded by the end of grade 4: the number of accumulated retentions during cycle 1 and the scores (1-5 scale) obtained by each student in both mathematics and reading at end of grade 4 national exams. It also contains a binary variable taking value one if the student entered cycle 1 (primary schooling) aged 6 years-old or more – the most common case – and zero otherwise (i.e. aged between 5 and 6 years-old). The three outcome baseline variables and the entering age binary variable enter the model in equation (1) in a four-way interaction. Again, the interaction should capture with more detail different sorting behaviors, from the part of school authorities, applied to students with different recorded combinations of previous achievement (for a given level of maturity as proxied by their entering age). It is not hard to conceive that school authorities possessing information regarding the past of the student, namely past achievement (which they possess in a form quite similar to the actual dataset used in this study) do indeed take it also into account when forming classes in a way that, at least, mirror their priors regarding what may be an optimal classroom composition in terms of student ability distribution. In other words, it is conceivable that a student recording top scores in both exams at the end of grade 4, never experiencing retention and entering primary school with less than 6 years-old might be subject to a type of within-school sorting that differs from that of another student recording low baseline scores on both subjects, recording one or more retentions during primary schooling while having entered that cycle aged 6 years-old or more. If we add that students with different combinations of past outcomes tend to persist in those outcomes afterwards, then omitting such combinations could be a source of bias to the class composition effects' estimation.[18] Finally, the third term contains two individual time-invariant observable characteristics of the students: gender and place of birth (see section 3.2 for details on how the place of birth binary variable was created), which close the individual level characteristics fed into the econometric model aimed at controlling, jointly, for within-school sorting that they may have been subject.

Conditional on all abovementioned control variables I assume that the error term $\varepsilon_{iCTSa}^{6}$ is uncorrelated with the leave-out class compositions experienced by any given student, thus estimation of $\boldsymbol{\beta_1}$ should be, to a great extent, unbiased. Many likely confounding factors that simultaneously may predict class composition and affect individual student achievement are taken into account, namely those related with between and within-school sorting of students and teachers, and confounding treatments such as class size. It is the variations in class compositions that remain

[18] Todd & Wolpin (2003) and Hanushek & Rivkin (2010) provide theoretical frameworks that justify the use of a single baseline score as a summary of past factors under some technical assumptions. The fact that I include three interacted baseline education outcomes facilitates then the assumption that those past factors have indeed been controlled for.

within specific values of the confounding factors the ones that can credibly deliver causal estimates of their effects on individual student general cognitive achievement.

The final and most preferred specification is presented in equation (2) which allows the class composition effects to be heterogeneous across the respective individual characteristic.

$$Y_{iCTSa}^{6} = [\boldsymbol{CC}_{i(-i),C}^{\mathbf{6\&5}} \times \mathbb{I}_{\boldsymbol{z_i=1}}]\boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{z_i=1}} + [\boldsymbol{CC}_{i(-i),C}^{\mathbf{6\&5}} \times \mathbb{I}_{\boldsymbol{z_i=0}}]\boldsymbol{\beta}_{\mathbf{1}}^{\boldsymbol{z_i=0}} + \boldsymbol{CX}_{\boldsymbol{i,C}}^{\mathbf{6\&5}}\, \boldsymbol{\beta}_{\mathbf{2}} + \boldsymbol{TX}_{\boldsymbol{i,T}}^{\mathbf{6\&5}}\, \boldsymbol{\beta}_{\mathbf{3}} + \sum_{(j,G)} \phi_{t(j)}^{G} + \\ + \sum_{G} \phi_{s^G a} + \boldsymbol{X}_{\boldsymbol{i}}^{\mathbf{6\&5}}\, \boldsymbol{\beta}_{\mathbf{4}} + \boldsymbol{Y}_{\boldsymbol{i}}^{\mathbf{4}}\, \boldsymbol{\beta}_{\mathbf{6}} + \boldsymbol{W}_{\boldsymbol{i}}\, \boldsymbol{\beta}_{\mathbf{5}} + \varepsilon_{iCTSa}^{6} \tag{2}$$

The term $\mathbb{I}_{\boldsymbol{z_i=1}}$ is an indicator function taking value one only when the student has a *given* characteristic (from the relevant ones used to compute the compositional measures), say when the student is top achiever, and $\mathbb{I}_{\boldsymbol{z_i=0}}$ is another indicator function taking value one only when the student has the *opposing* characteristic, say when the student is not a top achiever. Since all students can only have one characteristic at a time (they cannot be simultaneously top and low achievers, or male and female, or foreigner and not foreigner, and so on) the model in equation (2) is not perfectly collinear, hence estimable. The advantage of this specification is that the compositional effects may differ for opposing types of students and are directly comparable.

Given the large number of fixed effects to be estimated it is necessary to firstly identify the largest connected set which, in turn, permits to compute the greatest number of the former, see Abowd, Creecy & Kramarz (2002). I use Stata command *reghdfe* by Correia (2017) throughout the regression analysis as it incorporates a first routine that identifies the connected set. On top of that, the command also drops singleton observations to improve the estimation of the variance-covariance matrix, see Correia (2015).

### *4.2 School Planner Problem*

Endowed with credibly causal estimates of the class composition effects, i.e. with $\widehat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{z_i=1}}$ and $\widehat{\boldsymbol{\beta}}_{\mathbf{1}}^{\boldsymbol{z_i=0}}$, I set up the (school level) social planner problem in this subsection. The aim is to contrast whether, at a given school, the actual allocation of students across classes can be considered to be socially optimal or if there is room to improve overall education performance of the students through a hypothetical better reallocation of pupils across the existing classes in that school. The planner's problem is to maximize aggregate end of 6th grade student general cognitive achievement in school *s* for a given cohort of students *a*:

$$\underset{l \in L}{\text{Max}}\, \hat{Y}_{sa}^{6}(l) \tag{3}$$

by choosing an allocation of students *l* from all possible allocations *L* that respect the following requirements: (i) each student is assigned to a given class, but not more than to one single class; (ii) the number of existing classes is kept constant (as compared to what is observed *ex ante* in that school for that cohort); and (iii) the size of the classes is equal or larger than the *ex ante* average class size registered in school *s* for cohort *a*. Each element belonging to *L* can be seen as an allocation matrix with number of columns equal to the fixed number of existing classes, number of rows equal to the fixed number of students of cohort *a* enrolled in school *s*, and each and all entries of the matrix either equal to one – row(student) *i* placed in column(class) *c* – or equal to zero – row(student) *i not* placed in column(class) *c*. Constraint (i) is straightforward – any given student can only be placed in one single class – which means that the sum of the elements of each row of the allocation matrix is constrained to be always one. Constraint (ii) fixes the number of columns (i.e. of classes) of the allocation matrices in order to improve the *ceteris paribus* interpretation of the reallocation exercise, i.e. to shut down possible school planner's endogenous reaction to increase or decrease the number of classes (and that of hired teachers) as the reallocation of students is processed. I add constraint (iii) in order to control, to some extent, the class size that each student will experience *ex post* the reallocation in comparison to the class size they were *ex ante* exposed to[19]. Under this constraint all *ex post* classes will end up having sizes equal (or marginally above) to the *ex ante* mean class size, meaning no individual student will be subject to any tremendous change in the class size he experiences (not larger than the difference between the smallest and largest class size to the mean class size). This also reinforces the *ceteris paribus* interpretation of the reallocation exercise, as class size is bound to not change dramatically for any given student (the fact that the estimated class size effects are actually not significant in the econometric models alleviate concerns from allowing even some class size variation to occur for each student across his *ex ante* and *ex post* classes). It is worth mentioning that the optimization problem in itself is simplified by reducing the number of possible allocations found in $L$ with constraint (iii).[20]

The objective function $Y_{sa}^{6}(l)$ can take many forms as long as it produces one single metric describing overall cognitive achievement across all students of a given cohort in a given school. Nevertheless, I impose a basic structure to it to facilitate the interpretation of the results. No matter what the form of $Y_{sa}^{6}(\cdot)$ is, its arguments will always be the differences in individual exam scores' differentials between *ex post* and *ex ante* allocations. This way the optimal allocation matrix will be the one that improves scores the most *relative* to the scores obtained with the *ex ante*, observed, allocation of students.[21] In turn, the differentials are the portions of the exam scores attributable only to the (leave-out) class compositions which may affect in different ways students with different characteristics as in equation (2). Given that the reallocation exercise assumes that all other education inputs (other than class composition) are fixed, then the exercise can only impact individual achievement – thus producing a differential – through changes of class compositions. These changes depend, in turn, on the reallocation of students, i.e. depend on $l$. Equations (4), (5), and (6) describe, respectively, the differentials and the difference in differentials:

$$Ex\ Ante\ Differential_i \equiv \hat{Y}_i^{6} = \left[\boldsymbol{CC}_{\boldsymbol{i(-i),C}}^{\boldsymbol{6\&5}} \times \mathbb{I}_{\boldsymbol{z_i=1}}\right]\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{z_i=1}} + \left[\boldsymbol{CC}_{\boldsymbol{i(-i),C}}^{\boldsymbol{6\&5}} \times \mathbb{I}_{\boldsymbol{z_i=0}}\right]\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{z_i=0}} \tag{4}$$

$$Ex\ Post\ Differential_i \equiv \hat{Y}_i^{6}(l) = \left[\boldsymbol{CC}_{\boldsymbol{i(-i),C}}^{\boldsymbol{6\&5}}(l) \times \mathbb{I}_{\boldsymbol{z_i=1}}\right]\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{z_i=1}} + \left[\boldsymbol{CC}_{\boldsymbol{i(-i),C}}^{\boldsymbol{6\&5}}(l) \times \mathbb{I}_{\boldsymbol{z_i=0}}\right]\widehat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{z_i=0}} \tag{5}$$

$$Difference\ in\ Differentials_i \equiv \hat{Y}_i^{6}(l) - \hat{Y}_i^{6} \tag{6}$$

Hence, the objective function is

$$\hat{Y}_{sa}^{6}(l) \equiv \hat{Y}_{sa}^{6}\left(\left(\hat{Y}_{i=1}^{6}(l) - \hat{Y}_{i=1}^{6}\right), \dots, \left(\hat{Y}_i^{6}(l) - \hat{Y}_i^{6}\right), \dots, \left(\hat{Y}_{i=N_{sa}}^{6}(l) - \hat{Y}_{i=N_{sa}}^{6}\right)\right) \tag{7}$$

that is, with as many arguments as the number of students $N_{sa}$ belonging to the relevant cohort *a* in a given school *s*.

---

[19] Imagine that a given school has distributed a given cohort of students throughout 6 classes, each with varying dimensions between 10 to 32 students. Not imposing any restriction on the class size each student may end up in after the reallocation exercise would make possible that students from the sized 10 class could be experiencing *ex post* a class size of, say, 30, and vice-versa for students initially placed in the sized 32 class.

[20] Even constraining the *ex post* classes to have specific sizes the number of possible allocation matrices in L is enormous. Using the previous example and assuming the total number of students dispersed across the 6 classes is 126, constraining all *ex post* classes to have a size of 21 renders a total number of possible allocation matrices that equals the sum of possible combinations of students across each class (the ordering of students within each class being irrelevant): $\overbrace{\binom{126}{21}}^{class\ 1\ combinations} + \overbrace{\binom{105}{21}}^{class\ 2\ combinations} + \cdots + \overbrace{\binom{21}{21}}^{class\ 6\ combinations} \approx 4.29 \times 10^{23} + 6.39 \times 10^{21} + \cdots + 1$. Letting each class size vary freely would increase even more the possible number of allocation matrices describing students' allocations across classes.

[21] Note that the problem is exactly the same as one that merely maximizes *ex post* aggregate achievement differentials since only these may be affected by the reallocation exercise. *Ex ante* allocations are predetermined as any outcome stemming from them.

The actual functional forms that the objective function takes are three:

$$Utilitarian: Y_{sa}^{6}(l) = \sum_{i=1}^{N_{sa}} \left(Y_i^6(l) - Y_i^6\right) \tag{8}$$

$$Quasi - Utilitarian: Y_{sa}^{6}(l) = \sum_{i=1}^{N_{sa}} w_i \left(Y_i^6(l) - Y_i^6\right) \tag{9}$$

$$Rawlsian: Y_{sa}^{6}(l) = \min_{i \in \{1,\ldots,N_{sa}\}} \left(Y_i^6(l) - Y_i^6\right) \tag{10}$$

with the first – equation (8) – meaning that the planner weights equally all students. The second – equation (9) – makes the planner to weight differently each student i, specifically assigning twice the weight to low baseline achievers:

$$w_i = \begin{cases} 2 \; if \; student \; i \; is \; low \; baseline \; achiever \\ 1 \; if \; student \; i \; is \; top \; baseline \; achiever \end{cases}$$

and the third – equation 10 – makes the planner to pursue the reallocation *l* that makes the largest increment between *ex post* and *ex ante* differentials in exam scores for the student with the smallest such increment.

A few final notes regarding the thought experience of the reallocation exercise are needed. First, the assumption that only class composition is allowed to change, i.e. that all education inputs are fixed, imply other assumptions when interpreting the corresponding results from the exercise. As discussed above, one such implication is that principals (or equivalent school authorities) are not allowed to endogenously change the number of classes, nor the number of hired teachers, in face of the reallocation of students. Another implication is that the mechanism by which teachers are assigned to classes is kept unchanged. It is also necessary to assume that parents do not strategically respond to the reallocation exercise, either by pressuring within the school to revert the exercise or by taking their children to other school. This can be conceived if parents are unaware of the exercise, e.g. if it is private knowledge to school authorities.

Second, the fact that the planners' problem is the same either working with *ex post* differentials or differences in differentials (these last quantities depend on both *ex post* and *ex ante* differentials, but the reallocation matrix *l* can only impact *ex post* differentials as *ex ante* ones are predetermined) means that the optimal allocation of students *l** can be interpreted not only as the allocation that most improves upon a realized and observed allocation – valuable from an evaluation viewpoint – but also as the allocation that is expected to deliver the best aggregate outcome when no allocation has yet been decided – valuable as an important input and tool for schools when setting up classes each year.[22]

Third, given the complexity of the school planner problem described here (stemming from the enormous amount of possible allocation choices available to him), and given that the problem cannot be solved with usual optimization techniques (such as solving first order conditions) – the objective function has a *discrete* domain consisting of different combinations of students across classes which makes it non differentiable – it is necessary to employ an algorithm that is able to deliver an acceptable result (within the context of students' allocation across classes) and that is simple enough to be employed in a decentralized manner, i.e. at the school level. The algorithm

---

[22] This claim does not hold if the objective function is specifically the Rawlsian one. In that case, the optimal allocation *l** derived from using as arguments the differences in differentials delivers either a weakly Pareto improving allocation (the minimum difference in differentials is null or strictly positive, thus everyone was not harmed by the reallocation), or the *ex ante* one (all other allocations harm at least one student, so it is better to not reallocate at all and the corresponding maximum minimal difference is zero as no reallocation occurs). If the starting point is to set up classes optimally and not to compare an actual realized allocation with an optimal allocation, then the arguments to be used are the *ex post* differentials and the resulting allocation will be the one that maximizes the minimal *ex post* differential which can be a negative value.

then used can be described as follows (for a given school, for a given choice of the objective function): (1st) generate $k$ random allocation matrices $l$, each respecting constraints (i), (ii), and (iii), and compute the implied *ex post* differential and difference in differentials (for all students individually) using equations (4), (5), and (6) for each generated matrix $l$; (2nd) for each generated matrix $l$ compute the value of the objective function – one of those in equations (8), (9), or (10) – using as inputs the individual students' differences in differentials which were calculated in the previous step; (3rd) from the sample distribution formed by the $k$ objective function values computed in the previous step store its sample maximum; (4th) keep generating allocation matrices $l$ and computing their implied objective function value until a first matrix $l^*$ is generated such that it yields an equal or larger objective function value than the one stored in the previous step; (5th – if the goal is to evaluate a predetermined *ex ante* allocation of students) state $l^*$ as a solution if its implied objective function value is larger than the value implied by the *ex ante* allocation, otherwise state the *ex ante* allocation as the solution. This algorithm is executable with usual spreadsheet software that is expectably available and used by local school authorities in many countries. In spite of, with great probability, it does not deliver the global maximizer(s) of the objective function[23] – thus it does not solve, with great probability, the planners' problem in a strict sense – it may deliver an allocation of students $l^*$ that improves upon the registered *ex ante* allocation. If so, finding such $l^*$ is sufficient to make the point that there might be room for improving overall general cognitive achievement through class composition rearrangements and that $l^*$ itself is an improving allocation of students within a given school.

### *4.3 Choice of Representative School*

The dataset presents around 3000 schools-year (around 800 distinct schools contributing students to the regression samples of models in equations (1) and (2) across the four years of data), each potentially a target for the school planner problem outlined in the previous subsection. In order to simplify the analysis a representative school was chosen according to the following procedure: (i) for each of the six class level compositions used in the econometric analysis the same statistic was computed but at the school-year level (using its student wise average) and the median across all schools-year stored (in addition to these six school level statistics it was also computed the number of students in the 6th grade in each school); (ii) using the actual school-year level compositional statistics (including 6th grade size) and their medians the Euclidian distance was computed as a metric describing how "close" each school-year is to the median school-year; (iii) sorting all schools-year from smallest to largest distance I picked the one with smallest distance that was composed of up to three 6th grade classes (this last restriction to ensure the generation of matrices $l$ was less time consuming).

Note that for the reallocation exercise one needs actual classes which, given the context being that of 2nd cycle general cognitive achievement, should be either of grade 5 or 6. However, whichever the grade of the class to be picked it is necessary to assume that that class is representative of both grades. This is so since the reallocation exercise uses cycle level estimated effects of average cycle level class compositional inputs as described in this subsection and in the previous one. Since the class compositions experienced by a given student between 5th and 6th grade are quite positively correlated (which justified in part to average them across the cycle) it seems an acceptable assumption that the 6th grade class of student $i$ is representative of what his 5th grade class was. Consequently, a

[23] There are, at least, one maximizer and one minimizer allocations as the number of possible allocations $l$ in $L$ is finite (in spite of being large) and each one produces an objective function value that is a real number.

reallocation exercise performed to a 6th grade class can also be seen as being performed to the class that has been persisting from grade 5.

# 5 Empirical Results

## *5.1 Education Production Function Estimates*

Table B.1 presents the estimation results of the education production functions outlined in equations (1) and (2) across columns (1) and (2), respectively. The former refers to a scenario where the effects of the class compositional measures are invariant with respect to the individual students' characteristics, whether the latter allows those marginal effects to be heterogeneous across students' types. Each coefficient represents the average number of standard deviations (SD) by which the average of mathematics and reading 6th grade national exams' scores changes given a one percentage point (p.p.) change of the respective 2nd cycle average class composition measure, all else constant. In what follows regarding the interpretation of the results I *always* consider a *20 p.p.* variation of the compositional dimensions and state the corresponding effect as a percentage of a SD. The largest significance level permitted to consider a given estimate to be statistically significant is the 5% as the number of observations used is relatively large.

The effect of increasing the (leave-out) percentage of top achieving classmates is estimated to be negative and statistically significant in both specifications. Both the top and the low achiever student is harmed by being placed in 6th and 5th grade classes that are increasingly populated by top achieving classmates. On average, their performance deteriorates by 4.2% to 4.8% of a SD, depending on their type.

Regarding the impact of increasing the share of classmates with no previous retentions it is positive and significant to any student in general, but of modest magnitude. Any general student performs about 3% of a SD better. Interestingly, decoupling the general effect to effects specific to students recording no past retentions and to students with at least one past retention the picture changes somewhat. It is the students with no previous retentions that drive the general result as they are the only ones with a statistically significant positive effect. Its magnitude is more than double than the one estimated for a generic student: this particular type of student increases his general cognitive performance by about 8% of a SD, whereas increasing the class share of students with no past retentions seems to have no effect on those recording at least one such retention.

The percentage of low-income classmates provides contrasting results when comparing the models without and with heterogeneous compositional effects. On one hand, a general student is estimated to be harmed by being exposed to an increasing percentage of low-income classmates as per column (1). However, on the other hand, allowing the effect to be specific to either low-income or non-low-income students reveals that low-income students seem indifferent to the share of low-income classmates, while non-low-income students are estimated to actually benefit from exposition to larger shares of low-income classmates – about 4.6% of a SD.

In turn, increments in the percentage of students with internet at home is estimated to deliver non-significant impacts in general, and only negatively impacting students without internet at home themselves. Nevertheless, the latter effect is small in magnitude, about 2.6% of a SD.

As in the previous case increases in the percentage of male classmates yield, in general, non-significant effects. Only male students seem to be somewhat hurt in their cognitive performance levels if exposed, during grades 5 and 6, to larger shares of other male classmates. They record a small performance decrease of about 2% of a SD.

Finally, the estimated effects stemming from changes in the percentage of foreign classmates also differ across the two specifications. Whereas a generic student is estimated to benefit in face of an increasing proportion of foreign classmates – a sizeable boost in performance of about 5.6% of a SD – the two particular types of students (foreigner and non-foreigner) do not present statistical significant coefficients which diminishes the robustness of the estimated effect of the former.

### *5.1.1 Discussion of Education Production Function Estimates*

The results concerning the class composition effect of the percentage of top achieving classmates challenges previous studies, e.g. Hoxby (2000a), Hanushek et al. (2003), Sund (2009), and partially Burke & Sass (2013). Whereas these studies point to the hypothesis that sharing the grade or the class with more able peers is a positive input to the average student's education production, I find evidence supporting the opposite, i.e. that increasing class average predetermined achievement (through increasing the share of top achieving classmates) yields a negative effect on the average student (which can be seen as more in line with Burke & Sass, 2013), as well as on the average top and average low achieving students. Noteworthy, these results should not be driven by a large set of possible confounding factors stemming from non-random allocation of both students and teachers between-schools and, within-schools, across classes, as these were explicitly taken in consideration. I interpret this result as supporting the hypothesis that teachers, at any given specific age and with any given specific time-invariant combination of unobserved attributes, may tailor the teaching practices differently when facing different audiences in the classroom as posited in Duflo, Dupas, & Kremer (2011). Namely, classrooms which are increasingly academically oriented may lead the teacher to either show more advanced contents which may not necessarily be present on the national exams, or simply let the classes flow more rapidly.[24] Either way, both top and low achievers may have less opportunities during classes that have an increasing proportion of top achievers to entirely grasp the specific contents that will actually be present on the exams at the end of the schooling year. On the other hand, this also implies that the (negative) effects stemming from an increasing predominance of top performers or academically oriented classmates dominate over the (positive) effects that interactions with better achieving peers yield.[25]

In turn, the results concerning the impact of the percentages of classmates who have low-income status and internet at home can be interpreted as providing evidence supporting the previous hypothesis. As the proportions of these types of resource endowed classmates (that are likely to enable them to be more academically proficient) increase (i.e. as the percentage of classmates that have internet at home or that are non-low-income increase),

[24] Although not within the scope of this study it would be interesting to test whether the proportion of top achievers in a class can actually have a direct positive effect on later education outcomes, say on end of 9th grade exams. If indeed students placed in 6th and 5th grade classes majorly composed by top achieving classmates induces teachers to show more advanced contents (hence at the cost of confirming whether the programmed contents were indeed absorbed through repeating exercises or similar practices), then this may constitute an advantage when facing the next cycle, especially for top performers.

[25] The latter effect is observed in Kimbrough, McGee & Shigeoka (2017) which take away the presence of a teacher to identify peer-to-peer teaching positive effects.

achievement decreases. Conversely, classes increasingly composed by students from poorer backgrounds may induce teachers to momentarily exert extra effort in terms of exam related contents when facing such classrooms.

Finally, the results related with class gender and foreign background compositions align partially with Hoxby (2000a), namely that by increasing the proportion of females the education performance of males increases (although not by much). Less aligned is the result that the effect of increasing the share of foreign students has no significant effect, especially not a positive one for those that are foreigners themselves (although it might be questionable whether it is accurate to compare foreign background students – the ones under analysis here – with minority ones in the USA context as in her study).

### *5.2 School Planner Problem Results*

Table C.1 shows information regarding the outcomes of allocations $l^*$ found using the algorithm described in subsection 4.2 for each objective function[26]. It also presents statistics regarding the *ex ante* and *ex post* distributions of the class compositions across the classes of the chosen representative school[27].

Apart from the Rawlsian objective function, the other two specifications of the objective function allowed (re)allocation matrices $l^*$ to be found that are welfare improving (in comparison to the *ex ante* allocation matrix), as shown by the positive values attained by the objective function (first row) and by the corresponding average aggregate change in achievement per student[28]. The fact that using the Rawlsian objective function no welfare improving allocation matrix was found is not surprising. Finding such an allocation would imply that either the same differentials had been obtained for each student (so that the largest minimum difference between *ex post* and *ex ante* differentials is zero, i.e. every student was as better off with the *ex post* allocation as with the *ex ante* one), or a Pareto move could be produced since the new allocation would be capable to provide a positive achievement increment to the student with the worst difference of differentials. This is not expectable s changing the composition of the classes will with great probability benefit some individuals and harm a few others given the fact that removing or reinforcing a given type of student at a given class will likely 1) affect different types of students in *that* class in different directions and 2) affect different types of students in the *other* classes in different directions too (the multidimensionality of class composition embedded in the education production function and its heterogenous effects across different types of students justify this observation). In other words, it is necessary to assume that letting a few students to be harmed with a given hypothetical reallocation exercise could be justified with the performance boosts that it could offer to the many other students (or to many others of a particular type as per the spirit of the quasi-utilitarian function).[29]

---

[26] The parameter $k$ of the optimization algorithm was set to 30 for the current version of the paper. It is a small number that offers computational speed, however its cost is a higher chance to find an allocation matrix $l^*$ that delivers a not so much large extreme value for the given objective function at use. Given that, on top of this, other restrictions are at work for computational speed reasons, namely using a school with only 3 classes and constraining the *ex post* classes to have a specific average size, the interpretation of the school planner problem results should be qualitative rather than quantitative. Future versions of the paper will relax these constraints in order to provide a better description of the quantitative potential of class composition rearrangement.

[27] This school reports three classes of grade 6 from academic year 2013/14. Note that the class statistics presented in Table C.1 refer to the non-leave compositions of the three classes.

[28] Obtained by dividing the sum of the individual changes in achievement by the total number of $6^{th}$ grade students spread across the three $6^{th}$ grade classes.

[29] Note that although from an evaluation viewpoint using the Rawlsian objective function may be of little use (the *ex ante* allocation is very likely to always dominate over any possible *ex post* one since someone in some class may end up being harmed by the new allocation), from the viewpoint of a first iteration in setting up classes in the beginning of a given school year it is still valuable (the problem would then be to maximize the minimum differential instead of maximizing the minimum difference

Both the utilitarian and quasi-utilitarian induced better allocation matrices change the distribution of class compositions within the school compared to the distribution under the *ex ante* allocation. Whether the mean (non-leave out) percentages of classmates with each characteristic is almost unchanged (which is expectable given the *ex post* school population is exactly the same as the *ex ante* one), the corresponding levels of dispersion present observable patterns.

The proportion of top achieving classmates is less evenly distributed across the three classes under both welfare improving *ex post* allocations rather than under the *ex ante* one. The (leave-out) percentage of top achievers was estimated to impact negatively both top and low achievers which could lead one to conclude that it could be welfare improving to evenly spread top achievers across classes – making the classes having similar percentages of this sort of students – within a given school (so that this "bad" input could be diluted the most). However, the welfare improving allocations show larger dispersion levels of this class compositional dimension than the *ex ante* one, meaning that this type of student is less evenly distributed than before – the classes have more dissimilar percentages. This illustrates that credible estimated causal effects of class composition might be limited by its *ceteris paribus* nature when directly applied to policy. Top achieving classmates also possess other individual characteristics that have to also be taken into account when reallocating them across classes. Furthermore, the final effect of reallocating top achieving classmates also depend on the actual characteristics of all the other students within the school that may be affected by such reallocation.

Students with no previous retentions were estimated to be the only ones affected – positively – by the share of such classmates, hence the expected reallocation should group this type of students so that they may benefit from this more beneficial class composition. Again, the reallocation exercise goes in the other direction: the dispersion of students with no past retentions slightly decreases in both welfare improving allocations which signals that *ex post* classes have a more even distribution of this type of students instead of a more pronounced tracking.

The reallocation of low-income students followed a more expectable pattern. The share of low-income classmates, being estimated to improve education performance of non-low-income ones, could be expected to be evenly distributed across classes so to spread the most this "good" input. Indeed, for both welfare improving allocations the *ex post* dispersion levels were severely reduced.

The reallocation of students with internet at home does not seem to follow a clear pattern as its dispersion across classes increases and decreases depending on the welfare improving allocation. However, the associated estimated effect is small, hence less likely to drive any given reallocation exercise. The same goes for the percentage of male classmates as its estimated effect is also small in magnitude, nevertheless the pattern is clearer and is that of a more even distribution of males in the *ex post* welfare improving allocations.

Another possible explanation for the small increments (per student) yielded by the welfare improving allocations (beyond those related with computational speed constraints) is the relatively contained changes in class composition that students have experienced from the *ex ante* to the welfare improving *ex post* ones. Table C.2 documents that on average students experienced class compositional changes varying mostly between 5 to 15 p.p. (either as a decrease or as an increase as per the values on the column of standard deviation; the mean change in the

in differentials) as the best allocation matrix it may induce does not have to be the same as the ones induced by other objective functions that do not take in consideration merely the worst off student.

percentage of class compositions cancels out to zero). These small figures (lower than the hypothetical 20 p.p. variations performed in the previous subsection) imply small variations in performance differentials.

Finally, even if on average the general student is estimated to be benefited by the welfare improving allocations it is important to gain a sense of which specific types of students are made better or worse off. Table C.3 provides that information. Its first row replicates the values found in Table C.1 relative to the average aggregate change in achievement per student induced by each objective function. The remaining rows show the same value but now for each specific type of student. Given that using the Rawlsian objective function no welfare improving allocation was found the *ex ante* one is the *ex post* one, hence its column present zeros as there are no changes in achievement. The utilitarian and quasi-utilitarian cases are more interesting because although both are welfare improving for the general student they present specific types that were "sacrificed". Namely, the top achievers who are hurt by both of these *ex post* allocations. In contrast, the low achievers are, in both improving allocations, clear "winners" as the gains for them are above those for the general student. Moreover, under the quasi-utilitarian improving allocation, there are other types of "sacrificed" students, namely, those with at least one past retention and those with internet at home. In turn, those with no previous retentions, the low-income, those without internet at home and females are the types that come as gainers across both welfare improving allocations.

## 6 Conclusion

Using a rich micro dataset that permits to follow students through grades 5 and 6 from four different cohorts across schools, classes and teachers, I find evidence that (i) changes in class composition cause variations in general cognitive achievement (as measured by the average of the standardized scores of mathematics and reading national exams); (ii) these effects are heterogeneous across several of the compositional dimensions depending on students' own characteristics; (iii) using an utilitarian and a quasi-utilitarian social welfare functions there exist welfare improving allocations of students across the classes of a representative school (though the improvements are small in magnitude due to the restrictions imposed to improve computational speed); and (iv) the movements of students across classes induced by the welfare improving reallocations not always follow the pattern one could assume to be taken by just looking at the estimated *ceteris paribus* compositional effects. This last point suggests that indeed it is necessary to take in account all the students and all of their particular characteristics as any reallocation exercise is very likely to not be a *ceteris paribus* one (it will affect those that are reallocated as well as those that were *and* will be exposed to the reallocated students).

Future work should improve the robustness of the welfare improving allocations by alleviating the restrictions that were imposed to increase computational speed. A robust estimation of causal (heterogeneous) class composition effects and a robust algorithm that delivers non-marginal welfare improving allocations of students across classes (within a given school) are important ingredients to establish class composition rearrangement as a valid policy for improving general cognitive achievement (with the additional benefit that it should be financially costless to implement).

**Acknowledgments**

I thank Luís Catela Nunes, Ana Balcão Reis, Carmo Seabra, participants at the QED Jamboree 2017 meeting in Paris and at the Research Group of Nova SBE for helpful comments. I thank DGEEC - Ministry of Education of Portugal for providing the data and Diogo Pereira for excellent assistance with the dataset. I acknowledge financial support from Fundação para a Ciência e a Tecnologia [grant SFRH/BD/122573/2016] which had no involvement in any phase of this study. The views expressed here are solely mine. Any errors are my own responsibility.

## References

Abowd, J. M., Creecy, R. H., & Kramarz, F. (2002). *Computing person and firm effects using linked longitudinal employer-employee data* (No. 2002-06). Center for Economic Studies, US Census Bureau.

Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, 14(3), 229-241.

Bosworth, R. (2014). Class size, class composition, and the distribution of student achievement. *Education Economics*, 22(2), 141-165.

Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1), 51-82.

Card, D., & Giuliano, L. (2016). Can Tracking Raise the Test Scores of High-Ability Minority Students? *The American Economic Review*, 106(10), 2783-2816.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. (1966). Equality of educational opportunity (Coleman Report), *U.S. Department of Health, Education, and Welfare*.

Collins, C. A., & Gan, L. (2013). *Does sorting students improve scores? An analysis of class composition*. NBER. Working Paper 18848.

Correia, S. (2015). *Singletons, cluster-robust standard errors and fixed effects: A bad mix.* Technical Note, Duke University.

Correia, Sergio. (2017). *Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator* Working Paper. http://scorreia.com/research/hdfe.pdf

Direção-Geral de Estatísticas da Educação e Ciência (DGEEC) (2013). *Estatísticas da Educação 2011/2012 – Jovens*. Lisboa: Direção-Geral de Estatísticas da Educação e Ciência.

Duflo, E., Dupas, P, & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5): 1739–1774.

Duflo, E., Dupas, P., & Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of Public Economics*, 123, 92-110.

Hanushek, E. (1970). The Production of Education, Teacher Quality, and Efficiency. In *U.S. Office of Education, Do Teachers Make a Difference?*, 79-99. Washington, D.C.: Government Printing Office.

Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *Journal of Human Resources*, 14(3), 351-388.

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

Hanushek, E. A. (2008). Education Production Functions. In *The New Palgrave Dictionary of Economics Online*, ed. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.

Hanushek, E. A., & Rivkin, S. G. (2010). *Constrained Job Matching: Does Teacher Job Search Harm Disadvantaged Urban Schools?* NBER. Working Paper 15816.

Hoxby, C. (2000a). *Peer Effects in the Classroom: Learning from Gender and Race Variation*. NBER. Working Paper 7867.

Hoxby, C. (2000b). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics*, 115(4), 1239-1285.

Jürges, H., & Schneider, K. (2004). International differences in student achievement: An economic perspective. *German Economic Review*, *5*(3), 357-380.

Kimbrough, E. O., McGee, A. D., & Shigeoka, H. (2017). *How Do Peers Impact Learning? An Experimental Investigation of Peer-to-Peer Teaching and Ability Tracking*. NBER. Working Paper 23439.

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, 116(3), 777-803.

Pritchett, L., & Filmer, D. (1999). What education production functions really show: a positive theory of education expenditures. *Economics of Education Review*, 18(2), 223-239.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far. *Handbook of the Economics of Education*, 3(3), 249-277.

Sund, K. (2009). Estimating peer effects in Swedish high school using school, teacher, and student fixed effects. *Economics of Education Review*, 28(3), 329-336.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485).

West, M. R., & Wößmann, L. (2006). Which school systems sort weaker students into smaller classes? International evidence. *European Journal of Political Economy*, 22(4), 944-968.

Wößmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3), 695-736.

Yao, Y., Ohinata, A., & van Ours, J. C. (2016). The educational consequences of language proficiency for young children. *Economics of Education Review*, 54, 1-15.

Appendix

Table A.1 - Organization of the Portuguese public education system and temporal limits of the dataset in terms of national exams (mathematics and reading).

<table>
<tr><td rowspan="2">Year of Exams (scale in integers)</td><td>Low-Stakes</td><td colspan="3"></td><td>2007 to 2012<br>(1-5)</td><td></td><td>2007 to 2011<br>(1-5)</td><td colspan="6"></td></tr>
<tr><td>High-Stakes</td><td colspan="3"></td><td>2013 to 2015<br>(0-100)</td><td></td><td>2012 to 2015<br>(0-100)</td><td colspan="2"></td><td>2007 to 2016<br>(0-100)</td><td colspan="2"></td><td>2007 to 2016<br>(0-200)</td></tr>
<tr><td colspan="2">Grades</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td></tr>
<tr><td colspan="2">Cycles</td><td colspan="4">1st Cycle</td><td colspan="2">2nd Cycle</td><td colspan="3">3rd Cycle</td><td colspan="3">4th Cycle</td></tr>
<tr><td colspan="2">ISCED (2011)</td><td colspan="6">Level 1 - Primary Education</td><td colspan="3">Level 2 - Lower Secondary Education</td><td colspan="3">Level 3 - Upper Secondary Education</td></tr>
<tr><td colspan="2" rowspan="4">Types of Schools</td><td colspan="4">Type-1</td><td colspan="8"></td></tr>
<tr><td colspan="4"></td><td colspan="5">Type-2</td><td colspan="3"></td></tr>
<tr><td colspan="6"></td><td colspan="6">Type-3</td></tr>
<tr><td colspan="9"></td><td colspan="3">Type-4</td></tr>
</table>

Table A.2 – Decomposition of students by top baseline achiever status.

| Academic Year 2011/12 | | | Academic Year 2012/13 | | | Academic Year 2013/14 | | | Academic Year 2014/15 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Percent | | N | Percent | | N | Percent | | N | Percent |
| Top Achievers | 24,443 | 41.5 | Top Achievers | 26,889 | 45.3 | Top Achievers | 13,914 | 24.6 | Top Achievers | 8,101 | 13.5 |
| Low Achievers | 34,500 | 58.5 | Low Achievers | 32,435 | 54.7 | Low Achievers | 42,570 | 75.4 | Low Achievers | 52,118 | 86.6 |
| All | 58,943 | 100 | All | 59,324 | 100 | All | 56,484 | 100 | All | 60,219 | 100 |

Table A.3 – Descriptive statistics of the (leave out) class composition variables and of their respective individual level categorical variables

| | N | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Average 6th Grade Score | 234,970 | 55.4 | 17.4 | 1.5 | 100 |
| Top Baseline Achiever | 234,970 | 0.31 | 0.46 | 0 | 1 |
| Leave-Out % Top Baseline Achievers | 234,970 | 27.2 | 19.3 | 0 | 100 |
| Leave-Out % Top Baseline Achievers (if Top Baseline Achiever) | 73,347 | 37.3 | 19.7 | 0 | 96.2 |
| Leave-Out % Top Baseline Achievers (if Low Baseline Achiever) | 161,623 | 22.6 | 17.2 | 0 | 100 |
| No Previous Retentions | 234,970 | 0.97 | 0.18 | 0 | 1 |
| Leave-Out % No Previous Retentions | 234,970 | 82.2 | 11.9 | 0 | 100 |
| Leave-Out % No Previous Retentions (if No Previous Retentions) | 226,890 | 82.4 | 11.7 | 0 | 100 |
| Leave-Out % No Previous Retentions (if at Least 1 Previous Retention) | 8,080 | 75.2 | 14.6 | 0 | 100 |
| Low-Income | 234,970 | 0.43 | 0.50 | 0 | 1 |
| Leave-Out % Low-Income | 234,970 | 47.8 | 18.3 | 0 | 100 |
| Leave-Out % Low-Income (if Low-Income) | 102,163 | 53.8 | 17.4 | 0 | 100 |
| Leave-Out % Low-Income (if Non-Low-Income) | 132,807 | 43.2 | 17.7 | 0 | 100 |
| Internet | 234,970 | 0.57 | 0.50 | 0 | 1 |
| Leave-Out % Internet | 234,970 | 55.6 | 24.7 | 0 | 100 |
| Leave-Out % Internet (if Internet) | 133,308 | 65.3 | 19.9 | 0 | 100 |
| Leave-Out % Internet (if No Internet) | 101,662 | 42.8 | 24.5 | 0 | 100 |
| Male | 234,970 | 0.50 | 0.50 | 0 | 1 |
| Leave-Out % Males | 234,970 | 52.8 | 10.5 | 0 | 95.5 |
| Leave-Out % Males (if Male) | 116,455 | 52.9 | 10.4 | 0 | 91.2 |
| Leave-Out % Males (if Female) | 118,515 | 52.7 | 10.6 | 3.6 | 95.5 |
| Foreigner | 234,970 | 0.09 | 0.28 | 0 | 1 |
| Leave-Out % Foreigners | 234,970 | 9.7 | 11.5 | 0 | 91.4 |
| Leave-Out % Foreigners (if Foreigner) | 20,374 | 19.5 | 15.6 | 0 | 85.4 |
| Leave-Out % Foreigners (if Non-Foreigner) | 214,596 | 8.8 | 10.5 | 0 | 91.4 |

Table B.1 – Estimates of the education production functions with and without heterogeneous (leave-out) class composition effects.

| | | Model Without Heterogeneous (leave-out) Class Composition Effects (1) | Model With Heterogeneous (leave-out) Class Composition Effects (2) |
|---|---|---|---|
| % Top Baseline Achievers | | -0.0035*** | |
| % Top Baseline Achievers | × Top Baseline Achiever | | -0.0024*** |
| | × Low Baseline Achiever | | -0.0021*** |
| % No Previous Retentions | | 0.0015* | |
| % No Previous Retentions | × No Previous Retentions | | 0.0040*** |
| | × At Least One Previous Retention | | 0.0010 |
| % Low-Income | | -0.0026*** | |
| % Low-Income | × Low-Income | | -0.0010 |
| | × Non-Low-Income | | 0.0023*** |
| % Internet | | 0.0002 | |
| % Internet | × Internet | | 0.0004 |
| | × No Internet | | -0.0013*** |
| % Males | | -0.0007 | |
| % Males | × Male | | -0.0010* |
| | × Female | | 0.0008 |
| % Foreigners | | 0.0028*** | |
| % Foreigners | × Foreigner | | 0.0006 |
| | × Non-Foreigner | | 0.0007 |
| Average Class Size | | ✓ | ✓ |
| Average Teachers' Age Fixed Effects | | ✓ | ✓ |
| Grade 6 Math Teacher Fixed Effects | | ✓ | ✓ |
| Grade 6 Reading Teacher Fixed Effects | | ✓ | ✓ |
| Grade 5 Math Teacher Fixed Effects | | ✓ | ✓ |
| Grade 5 Reading Teacher Fixed Effects | | ✓ | ✓ |
| Grade 6 School-by-Cohort Fixed Effects | | ✓ | ✓ |
| Grade 5 School-by-Cohort Fixed Effects | | ✓ | ✓ |
| Grade 6 Parents' Job Status-by-Grade 5 Parents' Job Status Fixed Effects | | ✓ | ✓ |
| Grade 6 Parents' Education Level-by-Grade 5 Parents' Education Level Fixed Effects | | ✓ | ✓ |
| Grade 6 Low-Income Status-by-Grade 5 Low-Income Status Fixed Effects | | ✓ | ✓ |
| Grade 6 Internet At Home Status-by-Grade 5 Internet At Home Status Fixed Effects | | ✓ | ✓ |
| Male Fixed Effect | | ✓ | ✓ |
| Foreigner Fixed Effect | | ✓ | ✓ |
| Grade 1 Entering Age-by-Grade 4 Number Accumulated Retentions-by-Grade 4 Math Score-by-Grade 4 Reading Score Fixed Effects | | ✓ | ✓ |
| Adjusted R2 | | 61.9% | 60.2% |
| N | | 234,970 | 234,970 |

Notes: Significance levels: * 5%, ** 1%, *** 0.1%. Robust standard errors clustered at the class level. The dependent variable is the individual student average between math and reading standardized scores of end of grade 6 national exams. Each model also contains dummies equal to 1 if student *i* peers' measures were computed using partial class information, i.e. if up to 1/3 classmates of *i* had missing information about their baseline scores or their birthdate (necessary to compute the percentage of classmates of *i* with at least one past retention), as well as a dummy equal to 1 if student *i* teacher average age measure was computed using partial teacher age information, i.e. if 1 (out of a combination of 3 or 4) teacher had missing information with respect to age. The class composition variables (i.e. the percentages of classmates of student *i* with a given characteristic) were computed in a leave-out fashion, i.e. excluding student *i*, and averaged across grades 6 and 5 classes of student *i*. Models of columns (1) and (2) refer to equations (1) and (2), respectively (see subsection 4.1). Only classes with 10 or more students were used. Given the large number of fixed effects both models were computed using Stata command *reghdfe* which identifies the connected set (necessary for model identification) and drops singleton observations (to improve correct estimation of the variance-covariance matrix), see Correia (2017).

Table C.1 – "Better" allocations' results and within-school distributional changes

| | Ex Post | | | | | | Ex Ante | |
|---|---|---|---|---|---|---|---|---|
| | Utilitarian | | Quasi-Utilitarian | | Rawlsian | | | |
| Aggregate Change in Achievement (in SD) | 0.46 | | 0.50 | | 0.00 | | -- | |
| Average Aggregate Change in Achievement per Student (in SD) | 0.006 | | 0.002 | | 0.000 | | -- | |
| Class Composition (not leave out) at School Level | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| % Top Baseline Achievers | 19.4 | 8.6 | 20.0 | 7.5 | a | a | 18.6 | 6.2 |
| % No Previous Retentions | 93.1 | 5.2 | 92.4 | 5.3 | a | a | 92.7 | 6.1 |
| % Low-Income | 37.5 | 0.0 | 37.1 | 4.2 | a | a | 38.5 | 12.1 |
| % Internet | 48.6 | 22.1 | 49.6 | 14.1 | a | a | 48.1 | 16.3 |
| % Males | 56.9 | 7.1 | 56.1 | 7.9 | a | a | 57.7 | 9.2 |
| % Foreigners | 1.4 | 2.0 | 1.4 | 2.0 | a | a | 1.4 | 2.0 |

a - exactly same values as in the *ex ante* allocation

Table C.2 – Changes of (leave out) class compositions experienced by students.

| Utilitarian | | | | |
|---|---|---|---|---|
| Changes in (leave out) Class Composition | Mean | S.D. | Min | Max |
| Δ % Top Baseline Achievers | 0.0 | 7.9 | -16.3 | 11.6 |
| Δ % No Previous Retentions | 0.0 | 5.6 | -13.0 | 7.1 |
| Δ % Low-Income | 0.0 | 11.9 | -18.8 | 12.1 |
| Δ % Internet | 0.0 | 20.6 | -51.0 | 31.8 |
| Δ % Males | 0.0 | 9.1 | -21.5 | 19.8 |
| Δ % Foreigners | 0.0 | 2.4 | -4.5 | 4.3 |
| Quasi-Utilitarian | | | | |
| Changes in (leave out) Class Composition | Mean | S.D. | Min | Max |
| Δ % Top Baseline Achievers | 0.0 | 8.4 | -14.0 | 11.1 |
| Δ % No Previous Retentions | 0.0 | 7.1 | -15.8 | 12.1 |
| Δ % Low-Income | 0.0 | 11.7 | -21.1 | 18.0 |
| Δ % Internet | 0.0 | 19.2 | -39.4 | 23.1 |
| Δ % Males | 0.0 | 13.5 | -26.3 | 15.4 |
| Δ % Foreigners | 0.0 | 2.6 | -4.5 | 4.3 |
| Rawlsian | | | | |
| Changes in (leave out) Class Composition | Mean | S.D. | Min | Max |
| Δ % Top Baseline Achievers | 0 | 0 | 0 | 0 |
| Δ % No Previous Retentions | 0 | 0 | 0 | 0 |
| Δ % Low-Income | 0 | 0 | 0 | 0 |
| Δ % Internet | 0 | 0 | 0 | 0 |
| Δ % Males | 0 | 0 | 0 | 0 |
| Δ % Foreigners | 0 | 0 | 0 | 0 |

Table C.3 – Decomposition of the welfare gains and losses per type of student.

| | Average Aggregate Change in Achievement per Student (in SD) | | |
|---|---|---|---|
| Type of Student | Utilitarian | Quasi-Utilitarian | Rawlsian |
| Overall | 0.006 | 0.002 | 0.000 |
| Top Baseline Achiever | -0.001 | -0.016 | 0.000 |
| Low Baseline Achiever | 0.008 | 0.006 | 0.000 |
| No Previous Retentions | 0.007 | 0.005 | 0.000 |
| At Least 1 Previous Retention | 0.004 | -0.037 | 0.000 |
| Low-Income | 0.009 | 0.005 | 0.000 |
| Non-Low-Income | 0.005 | 0.000 | 0.000 |
| Internet At Home | 0.004 | -0.009 | 0.000 |
| No Internet At Home | 0.009 | 0.012 | 0.000 |
| Male | 0.005 | 0.000 | 0.000 |
| Female | 0.008 | 0.005 | 0.000 |
| Foreigner | 0.011 | -0.016 | 0.000 |
| Non-Foreigner | 0.006 | 0.002 | 0.000 |

**2013**

**2013/1, Sánchez-Vidal, M.; González-Val, R.; Viladecans-Marsal, E.:** "Sequential city growth in the US: does age matter?"
**2013/2, Hortas Rico, M.:** "Sprawl, blight and the role of urban containment policies. Evidence from US cities"
**2013/3, Lampón, J.F.; Cabanelas-Lorenzo, P-; Lago-Peñas, S.:** "Why firms relocate their production overseas? The answer lies inside: corporate, logistic and technological determinants"
**2013/4, Montolio, D.; Planells, S.:** "Does tourism boost criminal activity? Evidence from a top touristic country"
**2013/5, Garcia-López, M.A.; Holl, A.; Viladecans-Marsal, E.:** "Suburbanization and highways: when the Romans, the Bourbons and the first cars still shape Spanish cities"
**2013/6, Bosch, N.; Espasa, M.; Montolio, D.:** "Should large Spanish municipalities be financially compensated? Costs and benefits of being a capital/central municipality"
**2013/7, Escardíbul, J.O.; Mora, T.:** "Teacher gender and student performance in mathematics. Evidence from Catalonia"
**2013/8, Arqué-Castells, P.; Viladecans-Marsal, E.:** "Banking towards development: evidence from the Spanish banking expansion plan"
**2013/9, Asensio, J.; Gómez-Lobo, A.; Matas, A.:** "How effective are policies to reduce gasoline consumption? Evaluating a quasi-natural experiment in Spain"
**2013/10, Jofre-Monseny, J.:** "The effects of unemployment benefits on migration in lagging regions"
**2013/11, Segarra, A.; García-Quevedo, J.; Teruel, M.:** "Financial constraints and the failure of innovation projects"
**2013/12, Jerrim, J.; Choi, A.:** "The mathematics skills of school children: How does England compare to the high performing East Asian jurisdictions?"
**2013/13, González-Val, R.; Tirado-Fabregat, D.A.; Viladecans-Marsal, E.:** "Market potential and city growth: Spain 1860-1960"
**2013/14, Lundqvist, H.:** "Is it worth it? On the returns to holding political office"
**2013/15, Ahlfeldt, G.M.; Maennig, W.:** "Homevoters vs. leasevoters: a spatial analysis of airport effects"
**2013/16, Lampón, J.F.; Lago-Peñas, S.:** "Factors behind international relocation and changes in production geography in the European automobile components industry"
**2013/17, Guío, J.M.; Choi, A.:** "Evolution of the school failure risk during the 2000 decade in Spain: analysis of Pisa results with a two-level logistic mode"
**2013/18, Dahlby, B.; Rodden, J.:** "A political economy model of the vertical fiscal gap and vertical fiscal imbalances in a federation"
**2013/19, Acacia, F.; Cubel, M.:** "Strategic voting and happiness"
**2013/20, Hellerstein, J.K.; Kutzbach, M.J.; Neumark, D.:** "Do labor market networks have an important spatial dimension?"
**2013/21, Pellegrino, G.; Savona, M.:** "Is money all? Financing versus knowledge and demand constraints to innovation"
**2013/22, Lin, J.:** "Regional resilience"
**2013/23, Costa-Campi, M.T.; Duch-Brown, N.; García-Quevedo, J.:** "R&D drivers and obstacles to innovation in the energy industry"
**2013/24, Huisman, R.; Stradnic, V.; Westgaard, S.:** "Renewable energy and electricity prices: indirect empirical evidence from hydro power"
**2013/25, Dargaud, E.; Mantovani, A.; Reggiani, C.:** "The fight against cartels: a transatlantic perspective"
**2013/26, Lambertini, L.; Mantovani, A.:** "Feedback equilibria in a dynamic renewable resource oligopoly: pre-emption, voracity and exhaustion"
**2013/27, Feld, L.P.; Kalb, A.; Moessinger, M.D.; Osterloh, S.:** "Sovereign bond market reactions to fiscal rules and no-bailout clauses – the Swiss experience"
**2013/28, Hilber, C.A.L.; Vermeulen, W.:** "The impact of supply constraints on house prices in England"
**2013/29, Revelli, F.:** "Tax limits and local democracy"
**2013/30, Wang, R.; Wang, W.:** "Dress-up contest: a dark side of fiscal decentralization"
**2013/31, Dargaud, E.; Mantovani, A.; Reggiani, C.:** "The fight against cartels: a transatlantic perspective"
**2013/32, Saarimaa, T.; Tukiainen, J.:** "Local representation and strategic voting: evidence from electoral boundary reforms"
**2013/33, Agasisti, T.; Murtinu, S.:** "Are we wasting public money? No! The effects of grants on Italian university students' performances"
**2013/34, Flacher, D.; Harari-Kermadec, H.; Moulin, L.:** "Financing higher education: a contributory scheme"
**2013/35, Carozzi, F.; Repetto, L.:** "Sending the pork home: birth town bias in transfers to Italian municipalities"
**2013/36, Coad, A.; Frankish, J.S.; Roberts, R.G.; Storey, D.J.:** "New venture survival and growth: Does the fog lift?"
**2013/37, Giulietti, M.; Grossi, L.; Waterson, M.:** "Revenues from storage in a competitive electricity market: Empirical evidence from Great Britain"

**2014**

**2014/1, Montolio, D.; Planells-Struse, S.:** "When police patrols matter. The effect of police proximity on citizens' crime risk perception"
**2014/2, Garcia-López, M.A.; Solé-Ollé, A.; Viladecans-Marsal, E.:** "Do land use policies follow road construction?"
**2014/3, Piolatto, A.; Rablen, M.D.:** "Prospect theory and tax evasion: a reconsideration of the Yitzhaki puzzle"
**2014/4, Cuberes, D.; González-Val, R.:** "The effect of the Spanish Reconquest on Iberian Cities"
**2014/5, Durán-Cabré, J.M.; Esteller-Moré, E.:** "Tax professionals' view of the Spanish tax system: efficiency, equity and tax planning"
**2014/6, Cubel, M.; Sanchez-Pages, S.:** "Difference-form group contests"
**2014/7, Del Rey, E.; Racionero, M.:** "Choosing the type of income-contingent loan: risk-sharing versus risk-pooling"
**2014/8, Torregrosa Hetland, S.:** "A fiscal revolution? Progressivity in the Spanish tax system, 1960-1990"
**2014/9, Piolatto, A.:** "Itemised deductions: a device to reduce tax evasion"
**2014/10, Costa, M.T.; García-Quevedo, J.; Segarra, A.:** "Energy efficiency determinants: an empirical analysis of Spanish innovative firms"
**2014/11, García-Quevedo, J.; Pellegrino, G.; Savona, M.:** "Reviving demand-pull perspectives: the effect of demand uncertainty and stagnancy on R&D strategy"
**2014/12, Calero, J.; Escardíbul, J.O.:** "Barriers to non-formal professional training in Spain in periods of economic growth and crisis. An analysis with special attention to the effect of the previous human capital of workers"
**2014/13, Cubel, M.; Sanchez-Pages, S.:** "Gender differences and stereotypes in the beauty"
**2014/14, Piolatto, A.; Schuett, F.:** "Media competition and electoral politics"
**2014/15, Montolio, D.; Trillas, F.; Trujillo-Baute, E.:** "Regulatory environment and firm performance in EU telecommunications services"
**2014/16, Lopez-Rodriguez, J.; Martinez, D.:** "Beyond the R&D effects on innovation: the contribution of non-R&D activities to TFP growth in the EU"
**2014/17, González-Val, R.:** "Cross-sectional growth in US cities from 1990 to 2000"
**2014/18, Vona, F.; Nicolli, F.:** "Energy market liberalization and renewable energy policies in OECD countries"
**2014/19, Curto-Grau, M.:** "Voters' responsiveness to public employment policies"
**2014/20, Duro, J.A.; Teixidó-Figueras, J.; Padilla, E.:** "The causal factors of international inequality in co2 emissions per capita: a regression-based inequality decomposition analysis"
**2014/21, Fleten, S.E.; Huisman, R.; Kilic, M.; Pennings, E.; Westgaard, S.:** "Electricity futures prices: time varying sensitivity to fundamentals"
**2014/22, Afcha, S.; García-Quevedo, J,:** "The impact of R&D subsidies on R&D employment composition"
**2014/23, Mir-Artigues, P.; del Río, P.:** "Combining tariffs, investment subsidies and soft loans in a renewable electricity deployment policy"
**2014/24, Romero-Jordán, D.; del Río, P.; Peñasco, C.:** "Household electricity demand in Spanish regions. Public policy implications"
**2014/25, Salinas, P.:** "The effect of decentralization on educational outcomes: real autonomy matters!"
**2014/26, Solé-Ollé, A.; Sorribas-Navarro, P.:** "Does corruption erode trust in government? Evidence from a recent surge of local scandals in Spain"
**2014/27, Costas-Pérez, E.:** "Political corruption and voter turnout: mobilization or disaffection?"
**2014/28, Cubel, M.; Nuevo-Chiquero, A.; Sanchez-Pages, S.; Vidal-Fernandez, M.:** "Do personality traits affect productivity? Evidence from the LAB"
**2014/29, Teresa Costa, M.T.; Trujillo-Baute, E.:** "Retail price effects of feed-in tariff regulation"
**2014/30, Kilic, M.; Trujillo-Baute, E.:** "The stabilizing effect of hydro reservoir levels on intraday power prices under wind forecast errors"
**2014/31, Costa-Campi, M.T.; Duch-Brown, N.:** "The diffusion of patented oil and gas technology with environmental uses: a forward patent citation analysis"
**2014/32, Ramos, R.; Sanromá, E.; Simón, H.:** "Public-private sector wage differentials by type of contract: evidence from Spain"
**2014/33, Backus, P.; Esteller-Moré, A.:** "Is income redistribution a form of insurance, a public good or both?"
**2014/34, Huisman, R.; Trujillo-Baute, E.:** "Costs of power supply flexibility: the indirect impact of a Spanish policy change"
**2014/35, Jerrim, J.; Choi, A.; Simancas Rodríguez, R.:** "Two-sample two-stage least squares (TSTSLS) estimates of earnings mobility: how consistent are they?"
**2014/36, Mantovani, A.; Tarola, O.; Vergari, C.:** "Hedonic quality, social norms, and environmental campaigns"
**2014/37, Ferraresi, M.; Galmarini, U.; Rizzo, L.:** "Local infrastructures and externalities: Does the size matter?"
**2014/38, Ferraresi, M.; Rizzo, L.; Zanardi, A.:** "Policy outcomes of single and double-ballot elections"

**2015**

**2015/1, Foremny, D.; Freier, R.; Moessinger, M-D.; Yeter, M.:** "Overlapping political budget cycles in the legislative and the executive"
**2015/2, Colombo, L.; Galmarini, U.:** "Optimality and distortionary lobbying: regulating tobacco consumption"
**2015/3, Pellegrino, G.:** "Barriers to innovation: Can firm age help lower them?"
**2015/4, Hémet, C.:** "Diversity and employment prospects: neighbors matter!"
**2015/5, Cubel, M.; Sanchez-Pages, S.:** "An axiomatization of difference-form contest success functions"
**2015/6, Choi, A.; Jerrim, J.:** "The use (and misuse) of Pisa in guiding policy reform: the case of Spain"
**2015/7, Durán-Cabré, J.M.; Esteller-Moré, A.; Salvadori, L.:** "Empirical evidence on tax cooperation between sub-central administrations"
**2015/8, Batalla-Bejerano, J.; Trujillo-Baute, E.:** "Analysing the sensitivity of electricity system operational costs to deviations in supply and demand"
**2015/9, Salvadori, L.:** "Does tax enforcement counteract the negative effects of terrorism? A case study of the Basque Country"
**2015/10, Montolio, D.; Planells-Struse, S.:** "How time shapes crime: the temporal impacts of football matches on crime"
**2015/11, Piolatto, A.:** "Online booking and information: competition and welfare consequences of review aggregators"
**2015/12, Boffa, F.; Pingali, V.; Sala, F.:** "Strategic investment in merchant transmission: the impact of capacity utilization rules"
**2015/13, Slemrod, J.:** "Tax administration and tax systems"
**2015/14, Arqué-Castells, P.; Cartaxo, R.M.; García-Quevedo, J.; Mira Godinho, M.:** "How inventor royalty shares affect patenting and income in Portugal and Spain"
**2015/15, Montolio, D.; Planells-Struse, S.:** "Measuring the negative externalities of a private leisure activity: hooligans and pickpockets around the stadium"
**2015/16, Batalla-Bejerano, J.; Costa-Campi, M.T.; Trujillo-Baute, E.:** "Unexpected consequences of liberalisation: metering, losses, load profiles and cost settlement in Spain's electricity system"
**2015/17, Batalla-Bejerano, J.; Trujillo-Baute, E.:** "Impacts of intermittent renewable generation on electricity system costs"
**2015/18, Costa-Campi, M.T.; Paniagua, J.; Trujillo-Baute, E.:** "Are energy market integrations a green light for FDI?"
**2015/19, Jofre-Monseny, J.; Sánchez-Vidal, M.; Viladecans-Marsal, E.:** "Big plant closures and agglomeration economies"
**2015/20, Garcia-López, M.A.; Hémet, C.; Viladecans-Marsal, E.:** "How does transportation shape intrametropolitan growth? An answer from the regional express rail"
**2015/21, Esteller-Moré, A.; Galmarini, U.; Rizzo, L.:** "Fiscal equalization under political pressures"
**2015/22, Escardíbul, J.O.; Afcha, S.:** "Determinants of doctorate holders' job satisfaction. An analysis by employment sector and type of satisfaction in Spain"
**2015/23, Aidt, T.; Asatryan, Z.; Badalyan, L.; Heinemann, F.:** "Vote buying or (political) business (cycles) as usual?"
**2015/24, Albæk, K.:** "A test of the 'lose it or use it' hypothesis in labour markets around the world"
**2015/25, Angelucci, C.; Russo, A.:** "Petty corruption and citizen feedback"
**2015/26, Moriconi, S.; Picard, P.M.; Zanaj, S.:** "Commodity taxation and regulatory competition"
**2015/27, Brekke, K.R.; Garcia Pires, A.J.; Schindler, D.; Schjelderup, G.:** "Capital taxation and imperfect competition: ACE vs. CBIT"
**2015/28, Redonda, A.:** "Market structure, the functional form of demand and the sensitivity of the vertical reaction function"
**2015/29, Ramos, R.; Sanromá, E.; Simón, H.:** "An analysis of wage differentials between full-and part-time workers in Spain"
**2015/30, Garcia-López, M.A.; Pasidis, I.; Viladecans-Marsal, E.:** "Express delivery to the suburbs the effects of transportation in Europe's heterogeneous cities"
**2015/31, Torregrosa, S.:** "Bypassing progressive taxation: fraud and base erosion in the Spanish income tax (1970-2001)"
**2015/32, Choi, H.; Choi, A.:** "When one door closes: the impact of the hagwon curfew on the consumption of private tutoring in the republic of Korea"
**2015/33, Escardíbul, J.O.; Helmy, N.:** "Decentralisation and school autonomy impact on the quality of education: the case of two MENA countries"
**2015/34, González-Val, R.; Marcén, M.:** "Divorce and the business cycle: a cross-country analysis"

**2015/35, Calero, J.; Choi, A.:** "The distribution of skills among the European adult population and unemployment: a comparative approach"
**2015/36, Mediavilla, M.; Zancajo, A.:** "Is there real freedom of school choice? An analysis from Chile"
**2015/37, Daniele, G.:** "Strike one to educate one hundred: organized crime, political selection and politicians' ability"
**2015/38, González-Val, R.; Marcén, M.:** "Regional unemployment, marriage, and divorce"
**2015/39, Foremny, D.; Jofre-Monseny, J.; Solé-Ollé, A.:** "'Hold that ghost': using notches to identify manipulation of population-based grants"
**2015/40, Mancebón, M.J.; Ximénez-de-Embún, D.P.; Mediavilla, M.; Gómez-Sancho, J.M.:** "Does educational management model matter? New evidence for Spain by a quasiexperimental approach"
**2015/41, Daniele, G.; Geys, B.:** "Exposing politicians' ties to criminal organizations: the effects of local government dissolutions on electoral outcomes in Southern Italian municipalities"
**2015/42, Ooghe, E.:** "Wage policies, employment, and redistributive efficiency"

**2016**

**2016/1, Galletta, S.:** "Law enforcement, municipal budgets and spillover effects: evidence from a quasi-experiment in Italy"
**2016/2, Flatley, L.; Giulietti, M.; Grossi, L.; Trujillo-Baute, E.; Waterson, M.:** "Analysing the potential economic value of energy storage"
**2016/3, Calero, J.; Murillo Huertas, I.P.; Raymond Bara, J.L.:** "Education, age and skills: an analysis using the PIAAC survey"
**2016/4, Costa-Campi, M.T.; Daví-Arderius, D.; Trujillo-Baute, E.:** "The economic impact of electricity losses"
**2016/5, Falck, O.; Heimisch, A.; Wiederhold, S.:** "Returns to ICT skills"
**2016/6, Halmenschlager, C.; Mantovani, A.:** "On the private and social desirability of mixed bundling in complementary markets with cost savings"
**2016/7, Choi, A.; Gil, M.; Mediavilla, M.; Valbuena, J.:** "Double toil and trouble: grade retention and academic performance"
**2016/8, González-Val, R.:** "Historical urban growth in Europe (1300–1800)"
**2016/9, Guio, J.; Choi, A.; Escardíbul, J.O.:** "Labor markets, academic performance and the risk of school dropout: evidence for Spain"
**2016/10, Bianchini, S.; Pellegrino, G.; Tamagni, F.:** "Innovation strategies and firm growth"
**2016/11, Jofre-Monseny, J.; Silva, J.I.; Vázquez-Grenno, J.:** "Local labor market effects of public employment"
**2016/12, Sanchez-Vidal, M.:** "Small shops for sale! The effects of big-box openings on grocery stores"
**2016/13, Costa-Campi, M.T.; García-Quevedo, J.; Martínez-Ros, E.:** "What are the determinants of investment in environmental R&D?"
**2016/14**, **García-López, M.A; Hémet, C.; Viladecans-Marsal, E.:** "Next train to the polycentric city: The effect of railroads on subcenter formation"
**2016/15, Matas, A.; Raymond, J.L.; Dominguez, A.:** "Changes in fuel economy: An analysis of the Spanish car market"
**2016/16, Leme, A.; Escardíbul, J.O.:** "The effect of a specialized versus a general upper secondary school curriculum on students' performance and inequality. A difference-in-differences cross country comparison"
**2016/17**, **Scandurra, R.I.; Calero, J**.: "Modelling adult skills in OECD countries"
**2016/18**, **Fernández-Gutiérrez, M.; Calero, J**.: "Leisure and education: insights from a time-use analysis"
**2016/19, Del Rio, P.; Mir-Artigues, P.; Trujillo-Baute, E.**: "Analysing the impact of renewable energy regulation on retail electricity prices"
**2016/20, Taltavull de la Paz, P.; Juárez, F.; Monllor, P.:** "Fuel Poverty: Evidence from housing perspective"
**2016/21, Ferraresi, M.; Galmarini, U.; Rizzo, L.; Zanardi, A.:** "Switch towards tax centralization in Italy: A wake up for the local political budget cycle"
**2016/22, Ferraresi, M.; Migali, G.; Nordi, F.; Rizzo, L.: "**Spatial interaction in local expenditures among Italian municipalities: evidence from Italy 2001-2011"
**2016/23, Daví-Arderius, D.; Sanin, M.E.; Trujillo-Baute, E.: "**CO2 content of electricity losses"
**2016/24, Arqué-Castells, P.; Viladecans-Marsal, E.:** "Banking the unbanked: Evidence from the Spanish banking expansion plan"
**2016/25 Choi, Á.; Gil, M.; Mediavilla, M.; Valbuena, J.:** "The evolution of educational inequalities in Spain: Dynamic evidence from repeated cross-sections"
**2016/26, Brutti, Z.**: "Cities drifting apart: Heterogeneous outcomes of decentralizing public education"
**2016/27, Backus, P.; Cubel, M.; Guid, M.; Sánchez-Pages, S.; Lopez Manas, E.**: "Gender, competition and performance: evidence from real tournaments"
**2016/28, Costa-Campi, M.T.; Duch-Brown, N.; García-Quevedo, J.**: "Innovation strategies of energy firms"
**2016/29, Daniele, G.; Dipoppa, G.**: "Mafia, elections and violence against politicians"

**2016/30, Di Cosmo, V.; Malaguzzi Valeri, L.**: "Wind, storage, interconnection and the cost of electricity"

## 2017

**2017/1, González Pampillón, N.; Jofre-Monseny, J.; Viladecans-Marsal, E.:** "Can urban renewal policies reverse neighborhood ethnic dynamics?"
**2017/2, Gómez San Román, T.:** "Integration of DERs on power systems: challenges and opportunities"
**2017/3, Bianchini, S.; Pellegrino, G.:** "Innovation persistence and employment dynamics"
**2017/4, Curto-Grau, M.; Solé-Ollé, A.; Sorribas-Navarro, P.:** "Does electoral competition curb party favoritism?"
**2017/5, Solé-Ollé, A.; Viladecans-Marsal, E.:** "Housing booms and busts and local fiscal policy"
**2017/6, Esteller, A.; Piolatto, A.; Rablen, M.D.:** "Taxing high-income earners: Tax avoidance and mobility"
**2017/7, Combes, P.P.; Duranton, G.; Gobillon, L.:** "The production function for housing: Evidence from France"
**2017/8, Nepal, R.; Cram, L.; Jamasb, T.; Sen, A.:** "Small systems, big targets: power sector reforms and renewable energy development in small electricity systems"
**2017/9, Carozzi, F.; Repetto, L.:** "Distributive politics inside the city? The political economy of Spain's plan E"
**2017/10, Neisser, C.:** "The elasticity of taxable income: A meta-regression analysis"
**2017/11, Baker, E.; Bosetti, V.; Salo, A.:** "Finding common ground when experts disagree: robust portfolio decision analysis"
**2017/12, Murillo, I.P; Raymond, J.L; Calero, J.:** "Efficiency in the transformation of schooling into competences: A cross-country analysis using PIAAC data"
**2017/13**, **Ferrer-Esteban, G.; Mediavilla, M.:** "The more educated, the more engaged? An analysis of social capital and education"
**2017/14, Sanchis-Guarner, R.:** "Decomposing the impact of immigration on house prices"
**2017/15, Schwab, T.; Todtenhaupt, M.:** "Spillover from the haven: Cross-border externalities of patent box regimes within multinational firms"
**2017/16, Chacón, M.; Jensen, J.:** "The institutional determinants of Southern secession"
**2017/17, Gancia, G.; Ponzetto, G.A.M.; Ventura, J.:** "Globalization and political structure"
**2017/18, González-Val, R.:** "City size distribution and space"
**2017/19, García-Quevedo, J.; Mas-Verdú, F.; Pellegrino, G.:** "What firms don't know can hurt them: Overcoming a lack of information on technology"
**2017/20, Costa-Campi, M.T.; García-Quevedo, J.:** "Why do manufacturing industries invest in energy R&D?"
**2017/21, Costa-Campi, M.T.; García-Quevedo, J.; Trujillo-Baute, E.:** "Electricity regulation and economic growth"

## 2018

**2018/1, Boadway, R.; Pestieau, P.:** "The tenuous case for an annual wealth tax"
**2018/2, Garcia-López, M.À.:** "All roads lead to Rome ... and to sprawl? Evidence from European cities"
**2018/3, Daniele, G.; Galletta, S.; Geys, B.:** "Abandon ship? Party brands and politicians' responses to a political scandal"
**2018/4, Cavalcanti, F.; Daniele, G.; Galletta, S.:** "Popularity shocks and political selection"
**2018/5, Naval, J.; Silva, J. I.; Vázquez-Grenno, J.:** "Employment effects of on-the-job human capital acquisition"
**2018/6, Agrawal, D. R.; Foremny, D.:** "Relocation of the rich: migration in response to top tax rate changes from spanish reforms"
**2018/7, García-Quevedo, J.; Kesidou, E.; Martínez-Ros, E.:** "Inter-industry differences in organisational eco-innovation: a panel data study"
**2018/8, Aastveit, K. A.; Anundsen, A. K.:** "Asymmetric effects of monetary policy in regional housing markets"
**2018/9, Curci, F.; Masera, F.:** "Flight from urban blight: lead poisoning, crime and suburbanization"
**2018/10, Grossi, L.; Nan, F.:** "The influence of renewables on electricity price forecasting: a robust approach"
**2018/11, Fleckinger, P.; Glachant, M.; Tamokoué Kamga, P.-H.:** "Energy performance certificates and investments in building energy efficiency: a theoretical analysis"
**2018/12, van den Bergh, J. C.J.M.; Angelsen, A.; Baranzini, A.; Botzen, W.J. W.; Carattini, S.; Drews, S.; Dunlop, T.; Galbraith, E.; Gsottbauer, E.; Howarth, R. B.; Padilla, E.; Roca, J.; Schmidt, R.:** "Parallel tracks towards a global treaty on carbon pricing"
**2018/13, Ayllón, S.; Nollenberger, N.:** "The unequal opportunity for skills acquisition during the Great Recession in Europe"

ieb@ub.edu
www.ieb.edu

Human Capital

IEB Working Paper