Does class size influence student achievement?

Adrien Bouguen, Julien Grenet, Marc Gurgand

Institut des Politiques Publiques

DOES CLASS SIZE INFLUENCE STUDENT ACHIEVEMENT?

IPP Policy Brief

No. 28

September 2017

Adrien Bouguen
Julien Grenet
Marc Gurgand

www.ipp.eu

Summary

This policy brief presents the results of the most methodologically robust studies that estimate the impact of a reduction in class size on the student achievement and longer-run outcomes. Contrary to an idea that persisted for a long time, the effects are high, if they are compared with those of other educational policies conducted on a large scale and assessed rigorously. The absolute cost of such a policy justifies targeting it on the most disadvantaged pupils, but the investment is profitable whenever the salary gain for two years spent in a half-size class is greater than 1%. However many unknown quantities remain, in particular about how reduction in class size links up with teaching practices. ■

- A certain number of international research studies, most of which are recent, estimate the effects of a reduction in class size on the progress of the pupils by convincingly neutralising selection bias.
- Those studies show that halving the size of a class of 24 pupils by having two classes instead of one improves the mean performance levels of the pupils significantly, and it is even possible to see the long-term effects on academic pathways and occupational integration.
- These effects are observed even when the teachers are not given specific assistance or support with their teaching.
- The high cost of such a policy justifies targeting it on the most socially disadvantaged children, thereby working towards the goal of reducing inequalities.
- Cost-benefit calculations indicate that such a policy is profitable whenever it enables the future salaries of its beneficiaries to increase by 1%, through its benefits on academic pathways.

The Institut des Politiques Publiques (IPP) has been developed through a scientific partnership between the Paris School of Economics (PSE) and the Centre for Research in Economics and Statistics (CREST). IPP's aim is to promote quantitative analysis and evaluation of public policy using cutting-edge research methods in economics.

CREST
PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

www.crest.fr www.parisschoolofeconomics.eu

The impact of class size on pupils' scholastic performance has been the subject of a considerable number of scientific articles and reports. Interest in this subject is fuelled both by the popularity of such policies with various stakeholders (teachers, and parents), and by the relative comparability of the effects measured in very different educational contexts. As the French Government starts to significantly reduce class sizes in the most disadvantaged schools, the results show us that although this policy is costly, it can reduce scholastic performance gaps when it is implemented in a targeted and intensive manner. The work and reports on France's "priority education" policy generally emphasise that, since it was created, that policy has always failed to concentrate more resources per pupil in priority education establishments than in the rest of the education system (Bénabou *et al.*, 2005; Cour des Comptes, 2016). By targeting the networks of schools classified as REP (Priority Education Network) and as REP+ (Priority Education Network Plus), the reductions in class sizes are making it possible, doubtless for the first time, to contribute much more to the most disadvantaged pupils.

Measuring the effects of class size: an empirical challenge

What is the effect of a reduction in class size on pupil performance. Paradoxically, it is difficult to answer this simple question. Imagine we are to observe the performance of school pupils in classes of various sizes. Merely by comparing the performance of pupils in the "large" and the "small" classes, we would often observe that the results of the pupils are better in the large classes.

Can we conclude from this that the reduction in class size affects performance negatively? Of course not. This misleading result is due to what researchers call "selection bias": head teachers tend to place the least able pupils in the smallest classes whenever possible. Mere comparison of different class sizes thus teaches us absolutely nothing about the causal impact of reducing the number of pupils per class.

Testing at the start and end of the school year makes it possible to reduce this bias to some extent. It is then possible, instead of comparing the achievement or ability levels, to compare the progress observed in the large and small classes over the year (using the "value added" method, see box 1). This approach makes it possible to remove only some of the selection bias because the start-of-year tests will only measure a small part of the dimensions that determine the progress of the pupils, and that might have also influenced how they are distributed across the various classes. The vast majority of the studies on the effects of class size are based on this uncertain approach. Most of the studies taken into account in the classic summaries (meta-analyses) by Glass and Smith (1979, 1980) in the 1980s are of that nature. This also applies to the albeit very restrictive (eight studies) meta-analysis by Slavin (1989), and to the meta-analyses by Hanushek (1997) and by Hattie (2005). Such exercises are sensitive to the choice of the studies selected and to the weight given to their methodological soundness (Meuret, 2001; Krueger, 2003).

Box 1: Methods used for measuring the effect of class size on pupil performance

It is not sufficient to compare pupils in "small" and "large" classes in order to measure the impact of class size on scholastic performance. Such a comparison is distorted by multiple selection biases. Below, we briefly describe the methods that make it possible to reduce such biases, or indeed to correct them totally.

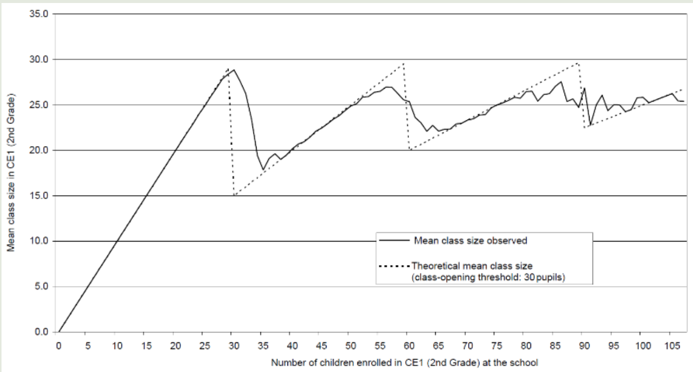
"Value added" method. This method consists in comparing progress in the performance of the pupils in small and large classes, on the basis of standardised tests given at the start and end of the year. The main limitation with this approach is that it does not make it possible to correct the differences in progress that are not related to the dimensions of pupil initial level that are measured by the tests, but rather to other differences not observable between the classes of different sizes (such as the characteristics of the teachers or the characteristics of the pupils that are not picked up by the tests).

Randomised experiment. In order to neutralise the selection biases that distort comparison between small and large classes, the most convincing method consists in allocating the pupils and the teachers randomly to classes of different sizes. Comparing the results of the pupils assigned to the classes constituted in this way makes it possible to isolate the "pure" effect of class size on scholastic performance. To our knowledge, the only example of application of this method is the STAR experiment that was conducted in Tennessee in 1986.

Quasi-experiments. When a random assessment cannot be conducted, it is sometimes possible to use events or rules that naturally produce random variations, beyond the control of the researcher. For example, this might be the variation in class size that is induced by the size of the school cohort each year (Hoxby, 2000), or it might be teacher recruitment forecasting errors that cause classes to be of larger or smaller size (Bressoux *et al.*, 2009). We also include in this category the French non-randomised experiment in its most rigorous analysis (Bressoux and Lima, 2011).

Class-opening thresholds. This method can be considered to be a quasi-experiment or a "life-size" experiment insofar as it uses an administrative rule that is encountered in many countries to determine whether a new class should be opened in a school. Most education systems have a maximum number of pupils per class (that number is about 30 in France). When that number is exceeded, the establishment requires that a new class be created, inducing an "exogenous" reduction in the average size of the classes of any given grade or year depending on the total number of children enrolled in that grade or year.
Graph 1 opposite, taken from the study by Piketty and Valdenaire (2006), illustrates this phenomenon for the classes of CE1 (Year 3 or 2nd Grade) during the year 1998-1999. By comparing the pupils at schools whose class sizes are just above and just below the thresholds at which new classes must be opened, this quasi-random variation makes it possible to measure the causal effect of class size on scholastic performance.

Graph 1 – Mean size of CE1 (Year 3 or 2nd Grade) classes as a function of number of children enrolled in CE1 in the schools in 1998-1999



Source: Calculations by the authors based on the administrative files for primary schools (survey No. 19)

This policy brief is based on studies, most of which are recent, that have adopted strategies making it possible to neutralise this selection bias much more convincingly. They constitute the most reliable information currently available to researchers for estimating the impact of a reduction in class size (see Table 1). The most famous study is the randomised experiment "STAR" (which stands for "Student-Teacher Achievement Ratio") conducted in Tennessee in 1986: the pupils and the teachers who took part in the experiment were assigned randomly to classes of small size (about 15 pupils) or large size (about 23 pupils), in a total sample of 325 classes. This method (cf. box 1, "randomised experiment") makes it possible to neutralise the selection bias manifestly and transparently. Another category of methods consists in using "quasi-experiments" making it possible to cause class size to vary for reasons that are unrelated to the initial achievement or ability levels of the pupils in the various classes (cf. box 1, "quasi-experiments"). Finally, a third method consists in using the maximum class size thresholds that exist in many countries (cf. box 1, "class-opening thresholds") as an exogenous source of variation of the number of pupils per class.

Effects on performance are established at primary level, but many unknown quantities remain

In Table 1, we present the main results of the selected studies, which cover a small sample of developed countries (USA, France, Israel, Norway, and Sweden). In this table, we look at the impact of one less pupil in the class on the mean scholastic performance levels of the pupils in that class, as measured by standardised tests (1). These tests are defined on variable and arbitrary scales (for example, the scales depend on the difficulty of the test). In order for the orders of magnitude to be comparable between the studies, it is common practice to normalise such tests by using their standard deviations in the population. Table 1 thus presents the effects of spending a year in a class with one less pupil, as expressed in a percentage of the standard deviation (box 2 explains how to interpret these measurements).

(1) A standardised test is an exam that homogeneously measures pupils coming from different classes or schools. Thus, the final exams for the *Diplôme National du Brevet* (a sort of general certificate taken in the 9th Grade (year 10) in France) is a standardised test, whereas the continuous assessment for that certificate (that depends on the teacher and on the school) is not. Using standardised tests enables results coming from different educational contexts to be compared better.

Box 2: Interpreting an effect measured in standard deviation per pupil less in the class

The effect of an education policy on the performance of the pupils is often measured on the basis of testing of the pupils. Such tests do not have a natural scale, they can be more or less difficult, they can have a larger or smaller number of questions, etc., and their absolute value has no precise meaning. It is common practice in the research literature to make them comparable by normalising them by using the standard deviation of the test in the population. Technically, this involves dividing the score obtained by each pupil by the standard deviation of the sample. Having two tests that are more or less difficult is like having two ladders with the rungs spaced apart to greater or lesser extents: normalising a test by its standard deviation is like looking at the ladder rung on which a pupil is standing, without taking account of the gaps between the rungs.

Thus, when the results are normalised by their standard deviation, what they express is more a relative level than an absolute performance. Once they are normalised, all of the tests have a standard deviation brought to 1. So, if the effect of an action is to increase the test by one (i.e. by one standard deviation, or by 100% of a standard deviation), that would be equivalent to bringing a median pupil (the 12th out of 24 pupils, for example) to the level of the pupil who is ranked 4th out of 24, i.e. progress of 8 positions in the class. An effect of 50% of a standard deviation corresponds, for the median pupil, to progress of 4 positions in a class of 24.

Halving the class size (from 24 to 12 pupils per class) by doubling the number of classes would, according to the studies identified in this policy brief, lead to an improvement in scholastic performance lying in the range 20% to 30% of a standard deviation. Such an effect is considerable: in a class of 24 pupils, it corresponds, for the median pupil, to progress of from 2 to 3 positions and it is of an order of magnitude comparable to one half of the mean performance gap that is observed in France at the start of CP ("*Cours Préparatoire*", i.e. 1st Grade (Year 2) between children from advantaged backgrounds (with parents who have managerial and professional occupations) and children from disadvantaged backgrounds (with parents who are workers, or who are not in work). In the United States, this effect corresponds to 60% to 80% of the average gap between white pupils and black pupils.

Table 1: Short-term effect of class size on pupil performance in standardised skills tests

Reference of study	Country	Impact of one less pupil on scholastic performance (in percentage of the standard deviation of the score in question)		Empirical method
A. Elementary (Primary) School				
Krueger (1999)	United States	**2,66**	Reading and maths (kindergarten)	Randomised experiment
		3,31	Reading and maths (1st Grade)	
		1,85	Reading and maths (2nd Grade)	
		2,44	Reading and maths (3rd Grade)	
Hoxby (2000)	United States	0,00	Maths (3rd Grade)	Quasi-experiment
		0,01	Reading (3rd Grade)	
		0,00	Writing (3rd Grade)	
		0,00	Maths (5th Grade)	
		0,00	Reading (5th Grade)	
		0,01	Writing (5th Grade)	
Angrist and Lavy (1999)	Israel	1,18	Maths (4th Grade)	Class-opening thresholds
		2,34	Reading (4th Grade)	
		3,95	Maths (5th Grade)	
		5,89	Reading (5th Grade)	
Angrist *et al.* (2017)	Israel	0,02	Maths (5th Grade)	Class-opening thresholds
		0,05	Hebrew (5th Grade)	
Piketty and Valdenaire (2006)	France	**2,79**	Maths (2nd Grade)	Class-opening thresholds
		2,19	French (2nd Grade)	
Bressoux and *al.* (2009)	France	**2,70**	Maths (3rd Grade)	Quasi-experiment
		2,60	Reading (3rd Grade)	
Bressoux and Lima (2011)	France	**2,20**	Maths and French (1st Grade)	Non randomised experiment
Fredriksson and *al.* (2013)	Sweden	**3,30**	Logical and verbal skills (4th to 6th Grades)	Class-opening thresholds
Iversen and Bonesrønning (2013)	Norway	**0,46**	Maths and reading (4th Grade)	Class-opening thresholds
B. Lower Secondary School				
Piketty and Valdenaire (2006)	France	**0,80**	Maths (9th Grade)	Class-opening thresholds
		1,14	French (9th Grade)	
		1,63	History & Geography (9th Grade)	
Leuven and *al.* (2008)	Norway	0,01	Maths (7th Grade)	Class-opening thresholds
		0,01	Norwegian (7th Grade)	

Notes : The references of each of the studies are indicated in the bibliography. The effects are expressed in percentage points of a standard deviation of the score in question per pupil less in a class for a school year. A positive coefficient should therefore be interpreted as a beneficial effect of the reduction in class size on pupil performance. The coefficients indicated in bold are statistically significant at the threshold of 5%. The other coefficients are non-significant at the same threshold.

Sources : Krueger (1999): table VII normalised by the standard deviations indicated in the appendix; Angrist and Lavy (1999): table VI; Angrist et al. (2017): table 6; Piketty and Valdenaire (2006): table 5, normalised by the standard deviations indicated in tables 1 and 10; Fredriksson et al. (2013): Table V; Iversen and Bonesrønning (2013): table 4; Bressoux and Lima (2011): table 3; Bressoux et al. (2009): table 4; Hoxby (2000): table IV (since the coefficients are estimated as percentages, they are converted into percentage points by estimating their values at the mean class size in the data, i.e. 21 pupils); Leuven et al. (2008): table 7.

The estimated effects of class size in elementary or primary schools are generally positive: of the nine studies identified, seven found statistically significant effects. These effects generally lie in the range 2% to 3% of a standard deviation of the test per pupil less in the class, which means, for example, that halving a class of 24 pupils (i.e. a reduction of 12 pupils per class) improves the mean performance of the pupils in question by from 20% to 30% of a standard deviation (2) at the end of the year (see box 2). Those conclusions are consistent with the recent review by Monso (2014).

The second part of Table 1 presents the few studies that are available on the impact of class size at lower secondary (junior high) schools. The effects appear less marked than at primary school level. In France, for example, the class-opening thresholds method results in effects that are half those found at primary school level (3). One possibility that is often discussed is that class size measurement is less reliable at lower secondary level. Another possible interpretation lies in the differences of teaching context. However, it is very difficult to speculate on interpretation of these differences with such a limited sample of lower secondary school studies.

Beyond the effects on mean performance levels of pupils, a few studies have endeavoured to measure the effects of class size for various different sub-populations of pupils. Piketty and Valdenaire (2006), for example, obtained effects that were twice as high as those presented in Table 1 when they considered children of disadvantaged social origin (children whose parents are workers or are not in work), thus suggesting that a targeted policy of class size reduction in the priority education schools could contribute to significantly narrowing performance differences as a function of social origins.

(2) It can be noted that the order of magnitude of the effects measured by the studies cited in Table 1 is twice as high as the one obtained in the recent review by Hattie (2005), which is 13% of a standard deviation. That review incorporates a large of number of studies, many of which are based neither on randomised experiments nor on quasi-experiments.

Similarly, Krueger (1999) found higher effects for the poorer pupils and for the black pupils, and Iversen and Bonesrønning (2013) identified effects that were twice as high for children from single-parent families. However, the other studies mentioned in Table 1 either do not examine this heterogeneity of effects, or do not find significant differences relative to the mean effect. Overall, the benefit of a reduction in class size could be particularly high for the most socially disadvantaged pupils, but the literature as a whole does not give a very firm demonstration of that.

Another unknown quantity concerns the non-linearity of the effect of class size. The above-mentioned studies are, in general, based on quite large variations in class sizes, of about ten pupils relative to classes numbering from 20 to 25 pupils on average. But we do not know whether the impact per pupil more or per pupil less would be lower or higher for class size variations of larger or smaller amplitude. We are not aware of any studies capable of establishing such fine distinctions convincingly (4).

The long-term effects of class size

Beyond its short-term effects on end-of-year scholastic performance, some studies suggest that class size influences academic pathways and labour market outcomes in the longer term (Table 2).

(3) The same uncertainty weighs on the effects of class sizes at lower secondary (*collège*) level on the numbers of pupils having to do the same year again: Gary-Bobo and Mahjoub (2013) show, for example, that in France smaller class size at lower secondary (*collège*) level significantly reduces the probability of repeating the year in Year 7 (6th Grade) and in Year 8 (7th Grade) but has no significant impact on the probability of repeating the year in Years 9 and 10 (8th and 9th Grade).

(4) In their review of the literature, Glass and Smith (1979) assert that lower effects (per pupil) are obtained by going from 40 to 30 pupils by class than by going from 20 to 10. However, that conclusion is based on comparing studies that were conducted in very different contexts, and that were based on methodologies that are difficult to compare.

Table 2: Long-term effects of class size

Reference	Country	Classes concerned	Effect of one pupil less in the class for one year	Empirical method
Krueger and Whitmore (2001)	United States (STAR)	Primary	Probability of taking the college (university) entrance examination (SAT test): +0.7 ppt Scores obtained in the SAT test: +0.03 of a standard deviation of the distribution of the scores	Randomised Experiment
Chetty and *al.* (2011)	United States (STAR)	Primary	Probability of going to college (university): +0.12 ppt Salary : n.s.	Randomised Experiment
Browning and Heinesen (2007)	Denmark	Lower Secondary	Number of years of studies: +0.02 years	Class-opening thresholds
Fredriksson and *al.* (2013)	Sweden	Primary	Number of years of studies: 0.02 years Rate of employment: n.s. Salary: +0,2%	Class-opening thresholds
Falch and *al.* (2017)	Norway	Lower Secondary	Number of years of studies: n.s. Salaries : n.s.	Class-opening thresholds
Leuven and Løkken (2017)	Norway	Primary and Lower Secondary	Number of years of studies: n.s. Salaries : n.s.	Class-opening thresholds

Notes: The references of each of the studies are indicated in the bibliography. The effects indicated in column 4 correspond to the impact of one pupil less in the classes concerned by the reduction in class size (column 3). The effects not statistically significant at the threshold of 10% are indicated by the abbreviation "n.s." (not significant). The coefficients expressed in percentage points are indicated by the abbreviation "ppt".

Sources: Krueger and Whitmore (2001): section 5 and table 6 (by dividing the effect by the mean difference between the large and small class sizes, i.e. 4.4 pupils); Chetty *et al.* (2011): table 5 (by dividing the effects by the mean difference between the large and small class sizes, i.e. 7.5 pupils, for a time of 2.14 years); Browning and Heinesen (2007): table 2; Fredriksson *et al.* (2013): table IV (by dividing the effects by the number of years spent in a class of reduced size, i.e. 3 years); Falch *et al.* (2017): table 4; Leuven and Løkken (2017): table 3.

Based on monitoring data from the American STAR programme, Krueger and Whitmore (2001) showed that the pupils who had been randomly assigned to the small classes at primary school were more likely to take college (university) entrance exams and achieved higher marks in them than the pupils who had been assigned to the larger classes. Following on from that study, Chetty *et al.* (2011) showed a positive effect of having been in a small class at school on the probability of going to college (university), without however managing to detect any significant impact on salaries in adulthood. In a very different context, Fredriksson *et al.* (2013) show particularly spectacularly that in Sweden, people who were in smaller classes at primary/elementary school not only attain a higher level of studies, but also enjoy a higher employment rate and higher income in adult life.

There is more uncertainty about the long-term effects of class sizes at lower secondary level. In Denmark, Browning and Heinesen (2007) found that small classes at lower secondary school had positive effects on number of years of studies. Conversely, two recent articles on Norwegian data (Falch *et al.*, 2017; Leuven et Løkken, 2017) did not detect any significant effects of class size at lower secondary school on level of studies or on income, tying in with the absence of effect on performance in the short term shown previously by Leuven *et al.* (2008).

Reduction in class size and teaching practices

It is sometimes said that reduction in class size is effective only if the teachers adjust their teaching practices to take advantage of it, and that such a policy is therefore relevant only if it is accompanied by teacher training to that end. Such a statement has no empirical basis. The experiments and quasi-experiments used in the literature presented above do not include any specific teacher training, and yet they produce significant gains in pupil performance. **In other words, empirical studies show that class size reduction is effective even if the teachers are not given specific support and assistance as regards the way they teach.**

Why? This point is discussed in education science, but it would seem that even if teachers make little change in their practices when they teach in smaller classes, it is observed that the pupils are more committed to their tasks, there are fewer discipline problems, more time is spent on teaching, and there is more prevention than remediation (Meuret, 2001; Ecalle *et al.*, 2006; Blatchford *et al.*, 2007; Blatchford *et al.*, 2016).
It might be legitimately thought that the effect of a reduction in class size would be even more significant if that policy were supplemented by teacher training and by tools or practices that take particular advantage of having smaller groups. But, to our knowledge, no studies exist that rigorously show such complementarity: for that purpose we would need to have an experiment or a quasi-experiment that makes it possible to establish the causal effect of a class size reduction coupled with training, and to compare it with the causal effect of a class size reduction without training.

A costly policy that needs to be appropriately targeted

Taken as a whole, the results of the studies identified in this policy brief indicate that a determined policy for reducing the size of classes in primary/elementary schools would significantly improve the performance of the pupils in question, and could, in the longer term, positively influence their academic pathways and their labour market outcomes.

The benefits of such a policy are however to be compared with its cost for the public purse, it being possible for that cost to be prohibitive if the class size reduction is insufficiently targeted. By taking inspiration from the figures proposed by various studies (Krueger, 2003; Fredriksson *et al.*, 2013), the annual cost of a policy consisting in halving the size of all of the CP (Year 2/1st Grade) and CE1 (Year 3/2nd Grade) classes in France by reducing them on average from 24 to 12 pupils per class (5) can be estimated to be about 5 billion euros, i.e. an amount equivalent to over 7% of the current French National Education budget!

In addition to its high cost for the public finances, a general reduction in the size of CP and CE1 classes might give rise to a reduction in the quality of recruitment of the teachers assigned to them, as has been observed in certain American States that have implemented policies of generalised reduction in class size (Jepsen and Rivkin, 2009; Dieterle, 2015). Under these conditions, it would seem more opportune, at least in the short term, to limit the reduction in class size to primary schools having priority education status only, especially in a context in France that is already marked by recruitment difficulties (Cnesco, 2016).

Targeting the scheme on the pupils who are the most disadvantaged socially also has the aim of redistributing the narrowing of performance gaps according to social origin: based on a mean value for performance improvement of 2% of a standard deviation per pupil less in a class, the benefit of halving CP and CE1 classes would represent almost one half of the performance gap between children from disadvantaged backgrounds (with parents who are workers, or who are not in work) and children from advantaged backgrounds (with parents who have managerial or professional occupations) at the start of the CP year (see box 2). Targeting the scheme on schools with priority education status could also be more effective than a general reduction in class size if, as suggested by certain studies, the socially disadvantaged pupils benefit to a greater extent from reduced class size than their classmates from better-off backgrounds.

Based on the same calculation hypotheses as above, the cost of halving the size of all CP and CE1 classes of REP (priority education network) schools and of REP+ (priority education network plus) schools can be estimated to be a little under 700 million euros per year. The annual cost of the scheme would be reduced to 300 million euros if it were limited to REP+ schools only (6).

(5) This calculation is made by assuming, like Fredriksson *et al.* (2013), that the fixed costs induced by a class size reduction policy (building new classrooms, additional training expenses, etc.) represent 30% of the variable costs. The variable costs correspond to the wage bill that would have to be paid to recruit the teachers of the new CP and CE1 classes. In 2016, the number of CP and CE1 classes (not combined or grouped together) in public primary (elementary) schools in France can be estimated to be about 60,000 (source: Indicateurs APAE). Doubling the number of those classes (so as to halve class size) would require about 60,000 teachers to be hired, with the corresponding annual wage bill being estimated at about 50,000 euros.

(6) in 2016, there were about 6,000 CP and CE1 classes (not combined or grouped together) in the REP schools, and 4,000 classes in the REP+ schools.

Would the benefits of such a policy for the pupils in question be greater than its cost for the public finances? To try and answer this question, it is possible to calculate the salary gain that should be enjoyed, on average, by the pupils of the schools concerned by the halving in size of the CP and CE1 classes to compensate for the cost of the measure. Assuming that the mean salary that these pupils will be able to earn in their careers is close to the present mean salary of the teachers, and choosing a conventional discount rate of 5%, **halving the sizes of CP and CE1 classes in schools having priority education status would be "profitable" if the salary gain for two years spent in a half-size class were greater than 1%** (7)**.** The results obtained by Fredriksson *et al.* (2013) on Swedish data suggest that this threshold could be exceeded, insofar as their estimations imply that a reduction in class size by 12 pupils for two years at primary (elementary) school would increase the salaries of the pupils in question by 5% on average.

Conclusion

While certain authors think that it is pointless to assign additional resources to the education system and recommend systemic reforms (such as introducing competition between schools), **recent research on class size reduction makes it possible, on the contrary, to justify a determined policy of public spending on education.** The soundest studies make it possible today to consider that the effects of class size reduction are relatively high, going against an idea that persisted for a long time in the education world. The cost-benefit calculation that we have proposed would suggest that the budgetary effort imposed by such a policy is very likely a profitable investment for the future, especially if it is appropriately targeted.

Another approach to education policy that seems less costly and more consensual consists in changing the teaching practices implemented, through training actions that follow the recommendations from the research. The order of magnitude obtained for a reduction in class size of about twelve pupils, i.e. 20% to 30% of a score standard deviation, is, admittedly less than the effects of teaching actions conducted in extremely controlled contexts, such effects easily approaching 50% to 100% of a standard deviation.

(7) With w denoting the annual cost of a teacher at an REP+ school and assuming that the fixed costs associated with the reduction in class size represent one-third of the variable costs, going from 24 to 12 pupils per class would result in an increase in the cost per pupil of 1.3 × w/12 – 1.3 × w/24, i.e. 5.4%. The present (discounted) value of the cost of doubling the number of CP and CE1 classes per pupil in question (measured at the age of entry into the CP year (Year 7, or 6th Grade) can be written $\sum_{t=0}^{1} \frac{0,054 * w}{(1+r)^t}$ where r designates the discount rate (assumed to be equal to 5%). Assuming, for simplifying the calculations, that the future mean salary of the pupils concerned by the scheme is equal to the present mean salary of the teachers (w), and that they work from the age of 20 (i.e. 14 years after they entered the CP year) to the age of 64 (i.e. 58 years after they entered the CP year), the discounted (present) value of the benefits of the policy can be calculated as the discounted sum of the salary gains induced by the years spent in classes of reduced size, i.e. $\sum_{t=14}^{58} \frac{\Delta * w}{(1+r)^t}$ where Δ designates the salary gain (expressed as a proportion of the mean salary) induced by two years spent in a class whose size has been halved. On these assumptions, the salary gain Δ that would equalise the discounted values of the benefits and of the costs of halving the size of the CP and CE1 classes in REP+ schools is equal to 1.1%.

Authors: Adrien Bouguen is a post-doctoral student at the University of Mannheim. Julien Grenet is a researcher at CNRS, an associate professor at the Paris School of Economics, and Assistant Director of the Institut des Politiques Publiques. Marc Gurgand is a research director at CNRS, associate chair at the Paris School of Economics, and scientific director of J-PAL Europe.

But **the actions or policies that are implemented on a large scale, or in "ecological" situations, and that are rigorously assessed, rarely present such high effects.** For example, in France, highly structured teacher training on preparing for reading have effects assessed at 15% of a standard deviation (Bouguen, 2016), while an action that is implemented outside the class, that is very intensive, and that is implemented in small groups, such as the "*Coup de pouce clé*" ("Key helping hand") scheme, does not have any demonstrated added value (Goux *et al.* 2017). In the United States, abundant literature estimates the effect of Charter Schools: most of those schools have no impact; only the best of them, those that adopt a specific teaching approach, have demonstrated effects, but they are then comparable with the effects of halving class size. Along similar lines, the Internat d'Excellence de Sourdun (Boarding School of Excellence of Sourdun) also has mean effects of 20% of a standard deviation, for a cost per pupil that is comparable to that of halving class size (Behaghel *et al.*, 2017).

There are too few rigorously assessed education policies, in particular in France, for it to be possible to perform systematic comparison, but it remains to be proved whether many other actions are available whose demonstrated effects, on a large scale, would be higher. Actually, many studies teach us that training often fails to change the practices of teachers in depth, through a lack of intensity and of sufficient support (8). Changing practices is not something that can be decreed. Admittedly, the teacher training should be pursued and reinforced but, by contrast, a class size reduction policy can be decided and implemented by the public authorities in quite a definite manner if it devotes the necessary resources to it. This is what gives it its force and its advantageousness. The literature shows that this impulse suffices to produce significant effects.

This review also shows that our knowledge is actually limited to a few orders of magnitude, while numerous details – are there any non-linear effects? To what extent do the effects differ according to sub-population? What role could teaching support play? What differences are there between the scholastic achievement levels? What context effects are more favourable? etc. – remain unexplored in the context of very robust methods. It is surprising that, on a topic as important and as debated, there has been so little experimental research making it possible to understand more clearly what is at stake.

(8) See Jacob (2017) for a recent example and illustrations.

Full bibliographic references: see following page

Bibliographic References

Angrist, J., and Lavy, V. (1999). "Using Maimonides' Rule to Estimate the Effects of Class Size on Scholastic Achievement", The Quarterly Journal of Economics, vol. 114, No. 2.

Angrist, J., Lavy, V., Leder-Luis, J., and Shany, A. (2017). "Maimonides Rules Redux", NBER Working Paper No. 23486.

Behaghel, L., de Chaisemartin, C., and Gurgand, M. (2017). "Ready for Boarding? The Effects of a Boarding School for Disadvantaged Students", American Economic Journal: Applied Economics, vol. 9, No. 1.

Bénabou, R., Kramarz, F., and Prost, C. (2005). "Zones d'éducation prioritaire : quels moyens pour quels résultats ?", Économie et Statistique, No. 380.

Blatchford, P., Russell, A., Bassett, P. Brown, P., and Martin, C. (2007). "The Effect of Class Size on the Teaching of Pupils Aged 7-11 Years", School Effectiveness and School Improvement, vol. 18, No. 2.

Blatchford, P., Chan, K., Galton, M., Lai, K-C., and Chi-Kin Lee, J. (Eds.) (2016). Class size: Eastern and Western perspectives, London-New York: Routledge.

Bouguen, A. (2016). "Adjusting Content to Individual Student Needs: Further Evidence from an in-service Teacher Training Program", Economics of Education Review, vol. 50(C).

Bressoux, P. Kramarz, F., and Prost, C. (2009). "Teachers' Training, Class Size and Students' Outcomes: Learning from Administrative Forecasting Mistakes", The Economic Journal, vol. 119, No. 536.

Bressoux, P., and Lima, L. (2011), "La place de l'évaluation dans les politiques éducatives : le cas de la taille des classes à l'école primaire en France", Raisons éducatives, vol. 15.

Browning, M., and Heinesen, E. (2007). "Class Size, Teach Hours and Educational Attainment", Scandinavian Journal of Economics, vol. 102, No. 2.

Chetty, R., Friedman, J., Hilger, N., Saez, E., Schanzenbach, D., and Yagan, D. (2011). "How Does your Kindergarten Classroom Affect your Earnings? Evidence from Project STAR", The Quarterly Journal of Economics, vol. 76, No. 4.

Cnesco (2016), Attractivité du métier d'enseignant : État des lieux et perspectives, Rapport scientifique.

Cour des comptes (2016) Rapport public annuel, La Documentation française.

Dieterle, S. (2015). "Class-size Reduction Policies and the Quality of Entering Teachers", Labour Economics, vol. 36.

Ecalle, J., Magnan, A., and Gibert, F. (2006). "Class Size Effects on Literacy Skills and Literacy Interest in First Grade: A Large-Scale Investigation", Journal of School Psychology, vol. 44.

Falch, T., Sandsør, A., and Strøm, B. (2017). "Do Smaller Classes Always Improve Students' Long-run Outcomes?", Oxford Bulletin of Economics and Statistics, forthcoming.

Fredriksson, P., Öckert, B., and Oosterbeek, H. (2013). "Long-Term Effects of Class Size", The Quarterly Journal of Economics, vol. 128, No. 1.

Gary-Bobo, R., and Mahjoub, M.-B. (2013). "Estimation of Class-Size Effects, Using "Maimonides' Rule" and Other Instruments: The Case of French Junior High Schools", Annals of Economics and Statistics, No. 111-112.

Glass, G. and Smith, M. (1979), "Meta-Analysis of Research of Class Size and Achievement", Educational Evaluation and Policy Analysis, vol. 1, No. 1.

Glass, G. and Smith, M. (1980), "Meta-Analysis of Research of Class Size and Its Relationship to Attitudes and Instruction", American Educational Research Journal, vol. 17, No. 4.

Goux, D., Gurgand, M., and Maurin, E. (2017). "Reading Enjoyment and Reading Skills: Lessons from an Experiment with First Grade Children", Labour Economics, vol. 45.

Hanushek, E. (1997), "Assessing the Effects of School Resources on Student Performance: An Update", Educational Evaluation and Policy Analysis, vol. 19, No. 2.

Hattie, J. (2005). "The Paradox of Reducing Class Size and Improving Learning Outcomes", International Journal of Educational Research, vol. 43.

Hoxby, C. (2000), "The Effects of Class Size on Student Achievement: New Evidence from Population Variation", The Quarterly Journal of Economics, vol. 43, No. 4.

Iversen, J., and Bonesrønning, H. (2013). "Disadvantaged Students in the Early Grades: Will Smaller Classes Help Them?", Education Economics, vol. 21, No. 4.

Jacob, B. (2017). "When Evidence is not Enough: Findings from a Randomized Evaluation of Evidence-Based Literacy Instruction (EBLI)", Labour Economics, vol. 45.

Jepsen Ch., Rivkin S. (2009). "Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size", Journal of Human Resources, vol. 44, No.1.

Krueger, A. (1999). "Experimental Estimates of Education Production Functions", The Quarterly Journal of Economics, vol. 113, No. 2.

Krueger, A. (2003). "Economic Considerations and Class Size", The Economic Journal, vol. 113, No. 485.

Krueger, A., and Whitmore, D. (2001). "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR", The Economic Journal, vol. 111, No. 468.

Leuven, E., Oosterbeek, H., and Rønning, M. (2008). "Quasi-Experimental Estimates of the Effect of Class Size in Norway", Scandinavian Journal of Economics, vol. 110, No. 4.

Leuven, E., and Løkken, S. (2017). "Long Term Impacts of Class Size in Compulsory School", IZA Discussion Paper No. 10594.

Meuret, D. (2001). Les recherches sur la réduction de la taille des classes, Rapport pour le Haut Conseil de l'évaluation de l'Ecole, La documentation française.

Monso, O. (2014). "L'effet d'une réduction de la taille des classes sur la réussite scolaire en France : développements récents", Éducation & Formations, No. 85.

Piketty, T., and Valdenaire, M. (2006), L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français. Estimations à partir du panel primaire 1997 et du panel secondaire 1995, Les Dossiers-Enseignement scolaire, No. 173, MEN-DEP.

Slavin, R. (1989). "Class Size and Student Achievement: Small Effects of Small Classes", Educational Psychologist, vol. 24, No. 1.