Getting Parents Involved: A Field Experiment in Deprived Schools.[1]

Francesco Avvisati[*], Marc Gurgand[*], Nina Guyon[†], Eric Maurin[*]

Abstract:

This article presents a randomized field experiment conducted in French middle schools located in a deprived educational district. Parents in treatment groups were invited to participate in a program of parent-school meetings on how to get better involved in their children's education. The experiment follows a "partial population" design, whereby volunteer parents are listed at the beginning of the year and treatment classes are then randomized. As only volunteer parents attend meetings, this provides a robust identification of both the overall effect of the program on the beneficiaries and the spillover effect on the untreated.

At the end of the school-year, we find that treated families effectively increased their school- and home-based involvement activities. Children of families who were directly targeted by the program developed more positive behavior and attitudes in school, and received better marks from their teachers, especially in French. Importantly, we also find large spillover effects of the program on the behavior of classmates of treated families. This experiment proves that schools are able to increase parents' awareness and that parental inputs have strong effects on pupil behavior. Our results on spillovers demonstrate that similar initiatives can be effective even in case of low parental take-up of the program.

Keywords: Parental Involvement, Cluster Randomized Trial, Classroom Peer-Effects, Child support

JEL Classification: I21, J13, J18

[1] This research was supported by a grant from the French Experimental Fund for the Youth. We are very grateful for the support of the schools and administrative teams from the *rectorat* de Créteil, and particularly to Bénédicte Robert. We thank the many J-Pal Europe research assistants that worked on this project. We also thank seminar participants at LSE (London), TSE (Toulouse), PSE (Paris) GREQAM (Marseille), Science Po (Paris), IZA and the University of Mannheim for useful comments, as well as discussants and participants at SOLE/EALE conference (London), ESPE conference (Essen) and EEA congress (Glasgow).

[*] Paris School of Economics

[†]LIEPP, Sciences Po

I. Introduction

Middle schools in modern societies face the challenge of providing foundation skills to very heterogeneous populations. The problems of truancy, violence, and pupil indiscipline are epidemic, especially in deprived urban areas. After spending three or four years in middle schools, many pupils are still far from reaching the basic requirements of curriculum (OECD, 2010). In this context, the view that better informed and more involved parents could contribute to overcome many difficulties enjoys a large consensus. Local initiatives abound, and plans to foster parental involvement are already eligible for federal funding in the US ("No Child left Behind" Act) and part of the national education policy in the UK ("Every Child Matters" Green Paper).

Yet, there is still very little evidence on whether such policies make a difference. It is not clear whether involvement policies conducted by schools can effectively increase parents' participation in education-related activities, especially among the most disadvantaged. It is not even clear whether improved parental involvement has any positive effect on pupils' behavior. The most involved parents differ from the less involved across many observed and unobserved dimensions and it is far from obvious that the empirical correlation between parental involvement and pupils' outcomes represents any causal effect at all.

This article proposes entirely novel evidence on these issues, using a field experiment undertaken in Paris area middle schools. We show that a simple program of parent-school meetings increased the level and quality of school-related parental care; and that this improved involvement of parents translated into a significant reduction of truancy and misbehavior in test classes, as well as improved motivation for school work. Impacts are strong: their order of magnitude is that of average differences between white collar and non-white collar families in the control group. Moreover, while only a limited share of parents took part in the program, we find that the behavior of all pupils was affected, including those whose parents did not participate. We provide evidence that those spillover effects result from direct interactions between pupils. Robust evidence on spillover effects comes from the "partial population" (Moffitt, 2001) design that was implemented: volunteer parents were listed in all classes prior to class randomization. Non-volunteer parents were not invited to participate to the program, but a random selection of their children was exposed to treated classmates.

The experiment started in September 2008, at the beginning of the 2008-2009 school year, in 34 middle schools of a relatively deprived educational district. In 183 classes, almost 1000 parents (22.5%) of sixth graders agreed to enroll in a program of three meetings with the school staff on how to successfully manage the transition from primary school to middle school. The program offered information on the functioning of schools and advice on how to help children with homework. The transition between primary and middle school represents one of the most critical stages of an educational career and this is why we chose to focus on 6^{th} grade, i.e., the first year of secondary education in France.

During the enrolment period, schools made it clear to parents that agreement to enroll would not necessarily result in participation and that only a random selection of enrolled parents would be effectively invited to participate in the program. By early October 2008, the enrolment period was closed. Of the 183 classes, 96 were randomly chosen to effectively run the program in November and December 2008. In each school and each class we are therefore able to clearly identify volunteer and non-volunteer families prior to the random decision of running the program. By comparing volunteer families in test classes to volunteer families in control classes, we capture the direct effect on the treated. By comparing their children's classmates, we capture the indirect effect of having treated families within the class. Finally, by comparing all pupils and families in test classes with all pupils and families in control classes, this randomization design is able to capture in a simple way the equilibrium impact of this program, under the assumption that there are no spillovers across classes.

We have collected data from a variety of sources and respondents, so as to come up with comprehensive and intersecting analysis of outcomes. The program is first found to have a positive impact on school-related involvement activities of enrolled families. For instance, the proportion of enrolled parents who actively participate in parents' organization at their school is 37% in test classes, whereas it is only of 25% in control classes. Our results therefore show that schools have a critical ability to influence parental attitudes and behaviors.

As a consequence, the behavior of pupils was affected along many dimensions. We find that the program is associated, by the end of the school year, with a decline by about 0.09 standard deviations in truancy in test classes, and a decline of the same magnitude in the probability of being sanctioned. At the same time, indicators of positive attitudes and behavior improved, thus signaling that the impact is not limited to the lower end of the distribution. Everything indicates the existence of a causal link between parents' involvement and their pupils' behaviors. Furthermore, changes in pupils' behavior are not limited to children of enrolled families: improvements spread out to their classmates. In particular, children from non-volunteer families are significantly less absent and receive fewer disciplinary sanctions when they are in treatment classes. These facts contradict the view that involvement policies are bound to benefit to a small fraction of volunteer families only.

Finally, we find that pupils in test classes receive better marks from their teachers, especially in French. By contrast, we do not find any impact of the program on pupils' results at tests that are externally set and marked. We interpret the effect on teachers' marks as signaling a change in children's attitude and motivation rather than stronger cognitive skills. Interestingly enough, effects on teachers' marks remain perceptible and significant at the end of 7th grade, 18 months after the end of the program, despite the fact that most pupils have moved to different classes with different teachers.

Our article lies within the scope of several different strands of the literature. First, we contribute to the economic literature on the importance of parental inputs for children's education. The few existing studies in this field within the economics discipline all adopt a structural approach, using survey data.[3] Most studies make use of NLSY panel data: Todd and Wolpin (2007) emphasize the preeminent role of "home inputs" relative to "school inputs"; Cunha and Heckman (2008) extend the analysis to distinguish the effect of parents' involvement on cognitive and non-cognitive skills. With a narrower focus on parents' supervision after school, Aizer (2004) or Welsch and Zimmer (2008) quantify its impact with different fixed-effect strategies. The identification of causal links and impacts in these articles has to rely on model assumptions about the form of the education production function. The education literature has also produced a number of estimates of the relation between parental involvement and student achievement but they are based on simple correlations.[4] To the best of our knowledge, our article is one of the first to provide large scale experimental evidence on the potential benefits of parental involvement for children's success up to late childhood.

We also contribute to the debate over the policy levers that can actually be used to improve pupils' behavior and performance. Parental attitude and involvement at school are widely perceived as key inputs for educational success, but little is known on whether these inputs can effectively be manipulated through simple policy initiatives. As it turns out, they are rooted in parents' own past and belong to the private sphere. Our study constitutes one of

[3] For a recent survey, see Avvisati, Besbas and Guyon (2010).

[4] Desforges and Abouchaar (2003, p.5) noted that "evaluations of interventions [in the area of parental involvement] are so technically weak that it is impossible on the basis of publicly available evidence to describe the scale of the impact on pupils' achievement". In a recent survey of educational studies, Hill and Tyson (2009) report only very few randomized interventions but they are very small scale and targeted to specific programs limited to involving parents into homework. In a different context, a rigorous experimental evaluation has been conducted by psychiatrists in four London elementary schools (Scott *et al.*, 2006). Its focus is on a parenting program for five- and six-years old.

the very few social experiments demonstrating that parental attitudes and school involvement can be significantly upgraded through simple participation programs and that such initiatives have a strong potential for reducing disciplinary problems in young teenagers. It is also the only study to provide evidence on spillovers, which is of considerable importance for public policy.

Our article is also related to the large and still growing literature on social interactions and spillover effects in education (see e.g., Hoxby 2000; Angrist and Lang, 2004; Carrel and Hoekstra, 2010). Specifically we provide new insights on how the school context can influence the behavior of pupils. We show that an early intervention targeting parents at the beginning of the school year translates into a progressive improvement of their own children's behavior and performance, with a maximum improvement at the end of the year. This result suggests that pupils have indirectly benefited from the intervention all over the year (through repeated interactions with their parents) and that this increasingly large "dose" of family interactions has contributed to a progressive modification of their behavior at school. Also, our specific cluster-randomized design makes it possible to identify spillover effects on non-volunteer parents (and on non-volunteer pupils) in addition to direct effects on volunteer populations and without the need for additional parametric assumptions. Most interestingly, we find no spillover effects on non-volunteer parents, but strong spillovers on their children, especially on behavioral outcomes and, again, with peak improvement by the end of the school year. Overall, our results are consistent with the assumption that the initial treatment has first influenced the attitude of children of volunteer parents through repeated family interactions, which has in turn progressively influenced the attitude of children of non-volunteer parents through repeated classroom interactions all over the school year. To the best of our knowledge, this social experiment is the first to provide such a decomposition of the working of social interactions between parents and pupils in the same school. It improves our understanding of how educational policies can exploit spillover effects and social interactions to enlarge the number of beneficiaries beyond the small share willing to participate.

Finally, our article contributes to the ongoing debate on community and user empowerment policies in western societies. As it happens, many developed countries are faced with the problem of an increasingly fragmented urban landscape with increasing disparities between poor and rich neighborhoods. Within this context, it is often argued that enclaves of social exclusion deserve special policies, relying on much greater involvement of local communities. Our article provides new experimental evidence on the outcome of increasing local residents' involvement in one key public service (education), in the context of a poor urban district of a western country.[5]

The remainder of this article proceeds as follows. Section 2 provides background information on the context in which the program took place and describes the interventions and its objectives. Section 3 introduces outcome measures, experimental design and estimation strategy, performs balancing tests on baseline data and discusses take-up and attrition issues. Section 4 presents the main results of the study. Section 5 provides robustness analysis to ensure that results are not driven by reporting biases. Section 6 concludes with implications for policy and future research.

II. Program and experimental protocol

A. French middle schools

The French state-run educational system is highly centralized with schools having limited autonomy. All schools are required to complete the same national curriculum and

[5] In developing countries, there is a similar debate on whether community empowerment policies would also result in reducing corruption and better public services, with mixed empirical evidence (Banerjee et al. 2010, Bjorkman & Svensson, 2009).

teachers are civil servants, selected through national examinations. After 5 years of elementary school, children enter middle schools at the age of 11 until the age of 14 (or more if they repeat classes). There is no streaming by ability across schools, and French parents are not free to choose the State school that their children will attend (otherwise than through residential location). In middle schools each subject is taught by a different teacher. For sixth graders, a typical week consists of 29 school hours, distributed across 9 different subjects, and, hence, different teachers.

Pupils stay in the same class throughout the school year, and in every subject. The class is therefore a very distinct and closed entity where most of interactions with same age children take place. Classes are groups of 20 to 30 pupils. The school head allocates children to classes before the beginning of the school-year, taking into account the second language chosen by the family (if the school offers a choice) and other preferences regarding elective subjects. By contrast, the school head is not allowed to sort children by ability. Each class has one specific reference teacher (*professeur principal*). A class council (*conseil de classe*) is formed for each class and composed of all teachers of the class, and representatives of parents, pupils and the school administration.

The year is divided into three terms. At the end of each term, the class council meets to discuss each student's work, achievement, and behavior. The council bestows honors and warnings that are transmitted to families on the report card, together with teacher grades; at the end of the year, the council decides on grade repetition, and on the elective courses that each pupil will be allowed to take in the future. Only the best students are allowed to take the optional courses which are considered prestigious (Latin, Greek, additional hours in Chinese, German or English, etc…). Through these decisions, teacher assessments have a lasting influence on later tracking decisions (general versus occupational tracks) which are taken at the end of middle school (9th grade).

B. The intervention

The experiment took place in the educational district of *Créteil*, which includes all suburbs located to the east of Paris. The district covers an area of approximately 6400 km² and 4 millions inhabitants. This mostly urban and suburban area has the highest density of immigrant populations in France (according to the 1999 census, 20.9% of the population were born outside Metropolitan France) and includes some of the most deprived areas of the Paris region.

During the school year 2008-2009, this district wanted to test out a program to improve parental involvement in school, because it was strongly perceived that disadvantaged parents have inadequate knowledge of and confidence with schools, and that this situation could be potentially improved by a simple intervention. The program was targeted at families of 6th graders (11 years old), the entry grade in middle school. The district asked the authors to design an evaluation of the program.

The experimental program consisted mainly of a sequence of three meetings which took place every two to three weeks, between November and December (early January in some cases).[6] Only parents were invited to those meetings and not their children. Sessions started at 6pm at the school and lasted typically two hours. Most of the time, they were conducted by the school head. He or she could draw on precise guidelines, designed by the districts' educational experts, and show excerpts from a specially conceived DVD introducing the main actors in middle schools, and what is at stake in this stage of education. Both local and district-level documents were distributed at these meetings, explaining the functioning and opportunities of the school attended by the children.

[6] The program was named "*La Mallette des Parents*" (the parents' schoolbag). An official description of the program can be found at http://www.ac-creteil.fr/equite-participation-mallettedesparents.html (in French; accessed in September 2011).

The two initial sessions of the program focused on how parents can help their children and participate at school and at home in their education. The last session took place after the first class council and end-of-term report card. It offered parents advice on how to adapt to the results of the first term. Parents were encouraged to ask questions, explain their problems and share their own experience. The meetings were framed as "discussions" – both between school representatives and parents and between parents – rather than as "information sessions".

District-level guidelines insisted that the facilitator of the discussion should develop the following arguments in discussions. (a) All parents can help their child, no matter what their own school record was and how familiar they are with the institution: what matters most is that children feel that their parents are interested in their school experience, and feel encouraged to talk often about it. (b) To do well, work in the classroom is not sufficient; homework and regular exercise are extremely important. (c) Parents should regularly scrutinize homework diaries and notebooks, and stay close to children while they repeat their lessons or do exercise. (d) To develop the best attitudes, children must feel that their parents have a good perception and knowledge of the school and that they adhere to the demands of teachers and administration.

At the end of the third session, the facilitator asked whether participants would like to attend additional sessions: (a) on parenting issues (following up with the first three meetings/discussions) or (b) on the use of (school-related) internet or (c) specifically designed for those who are not fluent in French. These additional sessions included more training elements, and were led by qualified adult trainers. Overall, the program cost between 1000 and 1500 euros per school (i.e. 8-12 € per student)

Generally speaking, the program and its materials were developed by educational experts at the district level in accordance with state-of-the-art psychological literature on parental involvement (see e.g., Hoover-Dempsey and Sandler 1995, 1997). According to this literature, parental involvement depends on three basic ingredients: (1) Parents become involved in schools if they hold the belief that they should be involved; (2) if they believe that their involvement can exert a positive influence on children's educational outcomes; and (3) if they perceive that the child and the school want them to be involved. As it happens, the program explicitly increased the level of invitations, and simultaneously raised the opportunities offered by the school to parents in the treatment group. Also, the topics developed at the meeting insisted on arguments (drawn from role-model and efficacy theories in psychology) about the ways in which involved parents can exert a positive influence on children's achievement. According to these theories parents can increase the quality of the effort exerted by children by giving them interest, attention, praise and rewards related to behaviors that lead to school success.

C. Experimental protocol

Over the summer before the start of academic year 2008-2009, school heads of the 350 State-run middle schools from the district were invited to volunteer for participating in the experimental study. About 10% did so. As a result, the universe to which the program was offered, and baseline and follow-up data were collected, consists of 34 volunteer schools, representing 183 classes, and the families of some 4300 pupils.[7] Almost two thirds of the schools that entered the study were located in an "educational priority zone" – a label that distinguishes historically deprived areas. The district average is about one third, which implies that disadvantaged schools were more likely to consider the program.[8] Experimental

[7] 37 schools initially volunteered to participate in the program, but 3 of them did not implement the experimental protocol properly (they did not conform to the randomization). As will appear below, randomization is within schools, so that this exclusion does not affect internal consistency.
[8] The national rate in turn is 17%, three times smaller than the rate observed in our experimental set of schools.

schools also have lower than average pass-rates at the national examination that takes place at the end of middle school ("*Brevet des collèges*"): in 2008, the pass rate was 72% against a district average of 79%, and a national average of 83%. Many families attending these schools are relatively poor.

Just after the start of academic year, during September 2008, the schools advertised the program to the families of their 6th-graders. School heads mainly used a standardized leaflet to inform families of the program. They were also encouraged to contact the families who are the least familiar with the school system through directed phone calls.

The program was presented as an outreach effort, distinct from usual parent-teacher meetings. Parents of 6th graders were told that the school would organize a series of three evening meetings to help them understand the role of each member of the educational community, the schools' organization, and to help them develop positive involvement attitudes towards their children's school education. Given this information, parents were invited to volunteer for participation in the program. It was always explicit that actual eligibility to the program would occur only conditional on a random selection of eligible classes. The leaflet explicitly stated that the experimental nature of the program implied a limitation on the number of beneficiary classes. By mid-October each school listed all families who signed up, and closed the registration phase. This list defines the population that we call "volunteer families", and has not been amended thereafter. Volunteer families constitute approximately one fifth of the total population (970 out of 4308).

Overall, the initial information campaign defined very clearly two distinct populations within each school and each class: volunteers and non-volunteers. In substance, volunteers are the fraction of parents who are the most receptive to the policy under consideration. Appendix Table A1 (based on data that will be detailed below) shows the basic pre-treatment differences between our samples of volunteer and non-volunteer families.

Volunteers have slightly more often white-collar occupation than non-volunteers (21% vs 18%) and belong more often to two-parent families (78% vs 73%). Generally speaking, there are no strong observable pre-treatment differences between the two populations. This fact may be a consequence of the principals' efforts to inform and attract all categories of parents, even those whose involvement is usually very weak. It may also be a combination of higher background parents having a stronger interest, but facing a higher opportunity cost of attending the meetings.

Randomization took place as soon as the registration phase was closed (mid-October). We randomized classes within each school at a uniform rate of: $m/2$ if the number of eligible classes in the school, m, was even; or $(m+1)/2$ if m was odd.[9] This procedure was carried out separately within each school, in the presence of the school head.[10] It resulted in the selection of 96 classes in the treatment group and 87 classes in the control group. Following randomization, volunteer families belonging to test classes were informed by the administration of the exact dates at which the three meetings would take place.

Appendix Table A1 shows that observable characteristics, as measured at the beginning of the year before treatment starts, are balanced across treatment and control groups: differences are weak and we can never reject the null that they occur by chance at standard levels of significance within subpopulations of volunteers and non-volunteers.

The randomization procedure defines four basic groups of families within each school: volunteers in test classes; non-volunteers in test classes; volunteers in control classes; non-volunteers in control classes. Of these four groups, only volunteers in test classes are invited to the program. The ability to evaluate separately the effect of the program on volunteers and

[9] 15 classes had 0 volunteer families: they were excluded from the whole process. In schools with an uneven number of classes, when a class had less than 4 volunteers, it was grouped with the class just above in terms of number of volunteers and the two formed a single randomization unit: the procedure for an even number of units was then applied. All empirical analyses use weights defined as the inverse of the probability of assignment.

[10] The publicity of the random allocation was intended to ensure trust in the impartiality and transparency of researchers, as was the fact that the "random sequence" was actually based on externally verifiable numbers: the landline number of the school and the school head's month and day of birth.

non-volunteers is one of the very attractive features of our experimental design: it corresponds to a "partial population experiment" (Moffitt, 2001). Within this framework, any difference in outcomes between volunteers in test and control groups will capture the causal effect of eligibility to the program on the population of volunteers. In contrast, any difference in outcomes between non-volunteers in test and control groups will capture the causal effect of having eligible peers on the population of non-volunteers. This will be interpreted as a treatment spillover. Finally, we will also provide estimates of the average equilibrium effect of the program by comparing all pupils in test classes to all pupils in control classes.

D. Program take-up

At the beginning of each session, we asked participants to sign in and collected the attendance lists. Table 1 shows the effective take-up rates across the four basic categories of families: volunteers in test classes, non-volunteers in test classes, volunteers in control classes, non-volunteers in control classes. Comfortingly, take-up is large for volunteers in test classes only, even though it remains far from 100%. Specifically, about 50% of families in this group participated in at least one session, and only about 12% attended all the three basic meetings. As a result of imperfect compliance, any significant difference between treatment and control groups, even among volunteers, will be driven by a relatively small proportion of actually treated families. School heads were not to invite other parents, nor even to inform them on session times. As a result, only a very small share in the control group did attend some meetings.

As discussed above, families attending the last meetings in the initial program could determine whether, and in what form, to continue with additional sessions. "Parenting sessions" ended-up being organized in only 17 schools out of 34; 14 schools offered sessions for parents on the use of the internet, and only a handful ran additional sessions for non-

French speakers. Overall, the number of families which participated in at least one additional session beyond the initial three meetings makes up about 16% of eligible parents. It must be understood, therefore, that additional sessions only have a marginal turnout. For most eligible parents, the program consisted in receiving additional invitations to the school over the school year and in participating to the initial three meetings only.

III. Outcome Measures

We are interested in three types of outcomes: (1) parental involvement attitudes and behaviors; (2) children's non-cognitive skills, as reflected by their disciplinary record and work effort; and (3) children's cognitive achievement, such as academic results. The latter two should be considered as final objectives from a policy point of view, whereas the first type of outcome should be assessed in order to understand the workings of the program. Non-cognitive skills are no less important than cognitive ones, with regards to social and economic implications for the individual. In particular, the effects of non-cognitive characteristics on the labor market are now well documented (see e.g. Chetty et al. 2011, Lindqvist and Vestman, 2011, or Heckman, Stixrud and Urzua, 2006).

Attitudes and behavior are hard to measure and are subject to reporting biases when subjectively assessed. In order to circumvent this issue, we have produced a number of different measures, either based on administrative data or questionnaires. This allows us to provide measures from different sources for each category of outcome and thus ensure that different points of view converge to the same conclusion. In addition, in the last section of this article, we will provide evidence that subjective reporting has not been systematically influenced by possible knowledge of who is treated and who is not.

In this section, we describe our instruments for data collection in turn and explain how they contribute to the measure of each of the outcomes.

A. Parent questionnaire

To assess the impact of the program on parental involvement attitudes, we distributed a questionnaire to all families at the end of the school year. It was distributed in all schools to each family via their children; parents were asked to send it back within a week. The parent questionnaire is a self-administered short questionnaire, with 12 questions on school-based involvement, home-based involvement as well as on parents' perception of the school. Specifically, the questionnaire consisted of 3 questions on school-based involvement behaviors (participation in parents' organization – a necessary condition for being a representative in the class council – participation in parents/teachers information sessions, individual appointments with teachers), 4 questions on home-based involvement and parental control (help with homework, knowledge of grades, control over time spent watching TV, control over time spent on videogames) and 4 questions on understanding and general perception (knowledge of available optional courses, plans about child's future, satisfaction with school, anxiety about violence). Finally, one question asks whether parents have been summoned by the school administration to discuss their child's behavior or academic results. Being summoned is not only a symptom of the child's insufficient discipline or work but also, under this coercive form, a consequence of a lack of regular contact with the school.

Parents' answers to these questions define our elementary measures for parental involvement. We have also constructed four synthetic scores – a global parenting score, a school-based involvement score, a home-based involvement score and an understanding and perception score. Each synthetic score is defined to be the equally weighted average of the corresponding elementary z-scores, where each elementary z-score is calculated by substracting the control group mean and dividing by the control group standard deviation. The global parenting score applies this scoring technique to the 12 items in the questionnaire; the three other scores apply the same technique to the three distinct subsections of the questionnaire. We have checked that the relative magnitudes of impacts on different outcomes and populations are unchanged if the elementary questions are aggregated using multiple correspondence analysis rather than simple average of z-scores. Average z-scores tend to maximize t-values, but produce somewhat lower effect size. The difference is however small.

B. Administrative records

Pupils' individual outcomes are mostly measured based on administrative registry data. A first category of measures reflects *general behavior and discipline in the school*, and is produced by the administration in different contexts. We first use information on absenteeism: it is defined as the number of half-days the child is not at school without a valid justification from his parents (an occasional hour skipped counts as a half-day). This is a very objective measure that we have for each term. We also collected data on whether pupils were given an official "disciplinary warning" or were temporarily excluded during each term. Temporary exclusions signal violent behaviors or repeated transgression of the rules. They are sentenced by the school head. We define a "disciplinary sanction" dummy that takes value 1 if the child has received any of those during the term. Finally, we can also use a good conduct mark (*note de vie scolaire*) that is given each term in most schools and should reflect assiduity, contribution to the school and respectful behavior.[11] It usually has a much skewed distribution, with about 30% pupils having either the maximum (20/20) or next-to-maximum mark (19/20). Therefore, our variable "good conduct" is a dummy that takes value 1 for those top values and 0 otherwise.

Another type of measure reflects *attitude in class and work involvement*, according to teacher judgment. Every term, the class council meets and comments on pupils' academic

[11] Ministère de l'Education nationale, *Circulaire DGSCO* n°2006-105, June 23, 2006.

results: when its general impression is positive, the pupil gets honors (given to about 30%). They are not just a reflection of marks in each topic, but also a judgment on the general attitude and an encouragement to persevere. They can take 3 values: we code a variable named "honors" as taking value 0 if no honor is granted, 1 for the lower level (*encouragements*), 2 for the second level (*compliments*) and 3 for the highest level (*félicitations*).

Together, this set of variables sheds light on changes taking place at both ends of the distribution of behavior and work attitudes. The "disciplinary sanction" variable separates the small proportion of students with heavy conduct problems, and can therefore measure an improvement occurring at the bottom of the distribution of behaviors. By contrast, "good conduct" and "honors" make it possible to capture improvements in performance among the better students. With respect to absenteeism, we find that it increases sharply from an average of one half-day per term to nine half-days per term, as we move from students with honors and top marks in good conduct to students with a disciplinary sanction. Absenteeism can thus be considered as providing an objective and continuous measure of the quality of pupil behavior over the whole distribution of student types.

Finally, we have data on children's *academic achievement*. We use two sources of information. First, we collected the teacher-given marks from end-of-term report cards. For each subject, they represent the average mark given by the corresponding teacher during the term. We use marks given in French and Math, as well as a weighted average of all subjects (with weights proportional to hours of instruction per week). Second, we can use a national test in Math and French conducted at the beginning of 6th grade in all middle schools. It is uniform across schools and classes, and was externally graded. We complemented that information with a second test in Math and French, specifically designed by the district pedagogical services, which was submitted to all classes at the end of the year, and also externally graded. We use the score in French and in Math separately.

The administrative database also contains baseline information from a registration form collected in July 2008, when parents registered their children for the upcoming school-year: this gives us demographic information on child and family (gender, date of birth, family situation, socioeconomic status…), as well as an exhaustive count of the universe of pupils.

C. Reference teacher questionnaire

In each class a teacher acts as “reference teacher”. He or she can be a teacher of any subject, although main subjects (Math and French) are over-represented. Any teacher can be a reference teacher for one class only. This teacher is thus most knowledgeable about the class, and he is the one most likely to meet with parents. At the end of the year, we submitted a self-administered short questionnaire to the reference teacher of each class.

For every pupil in the class, we first asked reference teachers whether their dialogue with the child’s parents was satisfactory and whether they felt that parents did provide support to the child with school work. We also asked two questions about the child’s attitude in class: how agreeable he/she is in class and whether he/she works seriously. One last question refers to academic outcomes: we asked how much progress the pupil made over the year. For every pupil in the class, the two first questions can be used to complement measures of parental attitudes, whereas the other questions provide additional measures for child cognitive and non-cognitive outcomes.

D. Long-term outcomes

By the end of the following school year (2009-2010), we went back to the schools in order to download from school information systems the available administrative information for those same pupils who had been exposed to the program in the previous school year. Most of them are 7^{th} graders and a few (2%) have repeated 6^{th} grade. Some, however, have left the school. The information that is available on the system and could be easily retrieved is: teacher marks in all subjects, half-days of absenteeism and the good conduct mark. With this information, we define the same variables as in the main analysis to evaluate the impact of the program at the end of 7^{th} grade (i.e., 18 months after the end of the program). As a consequence of grade repetitions and school transfers, attrition rate between 6^{th} and 7^{th} grade is relatively high, but reassuringly, attrition is not related to experimental arms (see Table A2). As will be discussed below, this data can serve two objectives. One is to assess the existence of long-lasting impacts of the intervention; the other is to use the fact that pupils are grouped in different classes with different teachers to neutralize any possible reporting bias in favor of treated classes that may have affected estimated effects at the end of the 6^{th} grade.

E. Response Rates

The number of observations included in each analysis is only limited by non-response, or more generally unavailability of the information. Response rates are generally high, but they vary for each type of source and, within each source, for each variable. Non-response acts as a filter on the information flow: it is thus very important to check that remaining observations are balanced across treatment arms; otherwise we could suspect that treatment-control comparison might be biased. As it will appear, there is no evidence of differential response rates for any of the sources. Also, by performing baseline comparisons again on the sample with observed response, we have checked that the initial balancing properties are still valid even after attrition.

The parent questionnaire was returned back in due time by approximately two thirds of volunteer and non-volunteer families (Appendix Table A3). On the subset of volunteer families, we made a special effort to improve the response rate: all volunteer families which did not return the questionnaire after a week were called during the following week, to answer the same questions during a short phone interview. This has increased the response rate for volunteers to about 80%. The table shows that non-response is balanced across treatment arms.

For outcomes measured from administrative records related to pupils' behavior, availability of information varies between 64% (good conduct) and 91% (honors) of the baseline sample (Table A4). Attrition here does not stem from intentional behavior, but rather from varying school or, sometimes, class-level practices. Indeed, for all outcomes most attrition is at the school-level (with entire schools missing from our data), or at the class level, and the residual individual-level attrition is small and similar (about 6%) for all outcomes. School-level attrition does not have the potential to introduce biases in estimation, as randomization was stratified by school. In principle, class-level attrition, and individual-level attrition, might cause more trouble, but we have checked for each outcome that resulting samples remained balanced with respect to baseline characteristics. The residual individual-level attrition can with high probability be attributed to school-migration during the school year, or, in some cases, over the summer preceding the school-year.

End-of-the-year standardized tests in French and Math are an exception to the extent that individual attrition, due to absenteeism, contributes more (13-15%). However, the overall response rate is still above 80%, and balanced.

Finally, the reference teacher questionnaire was filled and returned for about 75% of pupils (Table A5). Most of the attrition is at the school and class level, reflecting the fact that some schools or some teachers objected to answering this questionnaire. The individual attrition rate is small and similar to the rate for administrative measures, which indicates that, whenever returned, information on almost every pupil was provided. Again, non-response is balanced.

IV. Results

In this section, we analyze in turn the effect of the program on parental involvement and on pupils' outcomes. When focusing on eligible families, our basic finding is a joint improvement in both parental involvement and pupils' behavior (truancy, discipline). Because meetings were not designed for the children –who only very occasionally attended– impacts found at the pupil level can safely be interpreted as the consequence of constant exposure to their altered parents, between the beginning of the program (October 2008) and the end of the academic year (June 2009). We also find significant improvement in teachers' marks, especially in French, but no significant effects on the cognitive tests that are externally set and marked. As discussed below, the combination of strong effects on marks and small effects on test scores suggests that the effect on marks derives from a better attitude in class and motivation rather than from stronger cognitive skills. Finally, we show that the behavior of pupils from non-eligible families is affected when they stand in a treated class. As there is no such spillover at the parent level, this is interpreted as the result of interaction with affected peers.

A. Statistical model

In most specifications, we estimate the two following models on the volunteer and on the non-volunteer populations separately:

Volunteers:

$$Y_{\mathrm{ics}} = \alpha^{\mathrm{V}} T_c + X_i \beta_{\mathrm{V}} + \mathrm{u_s} + \nu_{ics}$$

Non- volunteers:

$$Y_{\mathrm{ics}} = \alpha^{\mathrm{NV}} T_c + X_i \beta_{\mathrm{NV}} + \mathrm{n_s} + \eta_{ics}$$

where, for each individual i in class c and school s, the variable T_c is a dummy indicating whether class c is a treatment class, and X_i is a vector of control variables that includes demographic characteristics of children (dummies for gender and birth rank), indicators for high or low SES (dummies for white collar families and for scholarship recipients), and measures of the child's academic track record at the beginning of the school year (a dummy for early grade repetition and controls for first-term marks in French and Maths). Parameters u_s and n_s represent two potentially distinct sets of school effects (for which school dummies are included), whereas variables v_{ics} and η_{ics} represent unobserved individual random effects which may have positive correlation for pupils from the same class. The parameters of interest are α^V and α^{NV}. The identifying assumption is that the treatment variable T_c is uncorrelated with the other observed and unobserved explanatory variables conditional on the volunteer status. This is a direct consequence of the experimental nature of the treatment assignment variable T_c, randomization having taken place after the information campaign was closed. Under this assumption, the estimated impact α^{NV} on non-volunteers can be interpreted as providing evidence of spillover effects within classes whereas the estimated impact α^V on volunteers encompasses direct effects of being invited to the meetings and possible feedbacks from treated and untreated peers, which cannot be disentangled.

As indicated above, observed predetermined characteristics are not statistically different across treated and control groups, so including the vector of controls has no significant effect on the results but increases precision. Parameter α^V (resp. α^{NV}) is identified as the difference in average outcomes between volunteers (resp. non-volunteer) pupils in treated and untreated classes, conditional on control variables. Both models are estimated with standard errors clustered at the class level.

In this evaluation, volunteers and non-volunteers from control classes are used to estimate the 'no program' counterfactual for treatment classes. This assumes that the outcomes in a given class are not influenced by the treatment status of pupils from other classes. This exclusion restriction is justified by the fact that the class being the constant structure of daylong school life in the French system, most peer interactions do take place within the class. Should also control classes be partly affected, we would most likely underestimate the true impacts.

B. Increases in Parental Involvement

The experimental evidence suggests that the program was successful at significantly improving parental attitudes. Based on both parents' questionnaires and reference teachers' assessments, Table 2 reveals higher levels of parental involvement by volunteer parents in test classes, as well as a better perception and understanding of the school. Volunteer families in test classes also declare having less often been summoned to the school for disciplinary reasons. Most estimated impacts on volunteer parents' are statistically significant at standard level and the 'global parenting score', averaging all questions in the parent questionnaire, increases by 15% of a standard-deviation. Appendix Table A6 further describes the observed differences in parents' behavior between treatment and control volunteers, across the 12 original dimensions measured by the questionnaire (from which the synthetic scores are computed).

A metric in which the magnitude of the results can be assessed is given by the difference across these same dimensions between white-collar families and non white-collar families in control classes (see Table 3).[12] As it turns out, the impact of the program on home-based involvement is smaller than the corresponding difference across white-collar and non white-collar families (6% vs. 18% of a SD), but the impact of the program on school-based involvement is even larger than the corresponding difference across white-collar and non

[12] Roughly speaking, white-collar families (managers, professors, engineers…) represent the top 20% of the population in term of social status.

white-collar families (21% vs. 13% of a SD). Overall, the invitation of the school head created, among volunteer families in different experimental arms, differences in parental involvement of about the same order of magnitude as those which pre-existed between the 20% of families with higher socio-economic status and the rest of families.

By contrast, having volunteer families in the same class does not affect the involvement of non-volunteer parents. When we restrict the analysis to non-volunteers, we find much smaller and equivocal differences between treatment and control groups. Estimated effects on non-volunteer parents are never significant at standard levels. This suggests that the effects of the principal's invitation do not spill over from volunteer to non-volunteer parents of treated classes. In contrast to their children, parents of same class children have no specific reason to interact or even know each other.

C. Improvements in pupils' non-cognitive outcomes

Turning to children, the measures taken at the end of the last school-term unanimously point to a better quality in children's relation to school in treatment than in control classes, across the complete range of available measures on behavior and attitudes (Table 4). With respect to discipline, children of volunteer parents in treatment classes skip fewer classes (absenteeism is lower by 1.1 half-day, where the average is 4.3 half-days during the last term), are less likely to be punished for disciplinary reasons (6.4% against 11.0% in control classes), and are more likely to earn the top marks for their conduct (35.5% against 29.0%). Attitude in class also improves: honors are more often given by the class council and more often of the upper kind, whereas the reference teacher answers more often that the pupil is agreeable to deal with in class (61.3% instead of 53.0%) and works seriously (56.5% against 54.2%), although this latter effect is not significant. As for parents, the observed differences are comparable to the difference between children of white-collar families and other children.

Overall, estimated improvements are perceptible at all levels of the distribution of behavior: very bad behavior is less frequent and very good one more frequent. Truancy, that was shown above to form a continuous, independent and objective measure, is also strongly affected.

Parents' better involvement and interaction with schools thus convert into more positive attitude and behavior of their children, both in terms of respecting the general rules of life in the school and of taking work more seriously. In a context where violence is spreading, even at young ages, and classes are difficult to hold, these are important results. They suggest that the program affects individual behaviors in a socially desirable direction. Overall, those findings imply both that it is possible to change parental involvement towards schools *and* that more involved parents do influence their children's attitude in the interest of the school as a whole and of children themselves.

D. Peer-effects on non-cognitive outcomes

Such strong improvement in behavior happens to also influence non-treated children. The lower panel of Table 4 shows that children from non-volunteer families do behave better when they are in treatment classes. As a result of being in the same class, peers of treated children are less absent (by about 0.6 half-days), receive fewer disciplinary sanctions (the proportion of sanctioned children is 2.4 percentage points lower) and are more likely to get the top mark for their conduct (by 4.6 percentage points). All these effects on general discipline are smaller than for volunteers (by about one half) but remain statistically significant. Attitude in class as assessed by the pedagogical team or the reference teacher shows similar improvements of about one half the size of impacts on volunteer children, but none of the measured differences is statistically significant at standard levels.

As there is no evidence of an interaction between treated and untreated parents, we exclude that spillovers originate from non-volunteer parents' influence on their own children.

Effects on children of non-treated parents could result from progressive modification of the context of teaching and learning within the classroom, made possible by an improvement in the attitude of the volunteer share of the class. However, the fact that spillovers are clearer on absenteeism and discipline than on attitude in class rather points to direct interactions between pupils, namely spillovers resulting from pupils imitating each other or taking action together.

To further test for the nature of peer effects, it is possible to compare classes with different number of volunteers. The test relies on the idea that peer effects are likely to increase with increased level of interaction between volunteer and non-volunteer pupils. If interactions with peers are the explanation for observed spillovers on pupils' discipline, we should observe some dose-response relationship between the magnitude of these spillovers and the number of treated peers.

As it turns out, the prediction that more interactions with treated peers lead to larger impacts on non-treated peers is borne out by the data: on the subsample of non-volunteers, the impact is indeed larger, for the three outcomes for which we find spillovers, when there are many volunteers in a treated class[13] than in classes with few volunteers (Table 5). Since we did not randomize treatment intensity across classes, we cannot exclude that these differences in spillovers across classes with high and low numbers of volunteers reflect (at least to some extent) the heterogeneity of spillover effects across classes which are *ex-ante* different. The exercise which is presented in Table 5 is, however, clearly suggestive of a dose-response relationship between the quantities of interactions with treated peers to which non-volunteers were exposed over the school year and the quality of their behavior at the end of the year.

To sum up, inviting parents to the meetings has produced a net improvement over the year not just in the behavior of children whose families were effectively invited, but also on

[13] In this analysis, "many volunteers" corresponds to a proportion of volunteers above the first tercile of the distribution of volunteers across classes (i.e., above 16% volunteers). Results are robust to change in this threshold.

all other children belonging to the same classes. The existence of spillovers on non-treated children provides an additional argument in favor of this kind of policy, since it extends at least some of the benefits to all children, despite directly involving only a fraction of all parents. Thereby, it provides a tool to increase average outcomes without necessarily generating a gap between the small fraction of children whose parents volunteer to attend evening meetings at school and all other children.

E. Improvements in pupils' cognitive achievement

Through its influence on the perceptions and attitudes of families, or through its effects on the behavior and motivation of pupils, the program could be expected to extend its benefits to academic achievement measures. We have three sets of measures of achievement to test this claim: teacher marks, reference teachers' assessment of progress, and standardized test scores. Teacher marks are essential in shaping pupils' opportunities; they influence grade retention decisions, future high-school plans, and, in the mid-term, the choice of optional subjects. One issue with this outcome, however, is that teachers can adjust their grading practice to the average level of their pupils. In such a case, the comparison between marks given in treated and control classes provides an estimate of the effect of the program which may be downward biased. For this reason, we also conducted externally set and marked tests, in French and Mathematics, which were taken at the end of the school year; these tests supposedly deliver a more objective measure of academic abilities.

Using teacher marks, we do find a significantly higher achievement in French for pupils of the volunteer treatment group, relative to control pupils; the magnitude of the differential is about 12% of a standard deviation (Table 6). We are not able to measure significant differences in Mathematics, but the average of all subjects, weighted by class hours, shows an improvement of 8% of a standard deviation. Confirming these findings, the reference teacher more often indicates progress for volunteer children in treatment classes, and the effect size is similar (12% of a SD).

In contrast, there are no effects in terms of test scores. As they provide objective and comparable measures, this would imply that neither parental involvement, nor pupils' behavior and attitude did translate into higher cognitive abilities. This is not entirely surprising, because cognitive achievement should be less easily altered than attitudes or behavior. But the fact that we find an effect on teachers' marks and no effect on test scores remains to be interpreted.

Test scores provide measures of cognitive achievement that are comparable across classes, but their limit is that pupils do not have any true incentive to succeed at these tests, as they do not have any consequence for their future (this is especially true for end-of-the-year tests). On the contrary, Teachers' assessments measure more largely school achievement which depends on the cognitive potential, but also on pupils' motivations for achievement. An impact of the program on this dimension of teachers' assessments is consistent with the previous observation that the program strongly improves pupils' attitude in class.

Table 7 provides additional evidence in favor of this interpretation in the sense that teachers' marks reflect non-cognitive determinants of school achievement. On the left-hand part of the table, we have regressed (in the control group) teacher marks in French and Mathematics on test scores in the same subjects and on the most objective measures of behavior that we have, namely absenteeism and sanctions. These regressions confirm that teacher marks are strongly correlated with the corresponding test scores. But, even conditional on the test scores, marks are significantly better for children who are less absent and received fewer sanctions. By contrast, the right-hand side of the table shows that, conditional on teacher marks, tests scores are not influenced by children's behavior. These results are consistent with our interpretation of teacher marks as signaling both higher cognitive ability and better attitude, where test scores are instead only informative about cognitive ability.

All in all, the program seems to impact mostly behavior and attitudes, and this is also apparent on teacher marks, to the extent that they reflect children's effort and motivations to succeed in school. Purely cognitive skills, at least when gauged with tests that come with no incentive to do well, are not affected.

Finally, turning to the non-volunteer group (lower panel of Table 6), there is no evidence of spillovers on any of the variables. For one thing, we could not expect spillover impacts on test scores when there is no direct impact on the volunteer population. For the other, spillovers on discipline and attitude in class are not strong enough to translate into significant spillover effects on teacher marks.

F. Class-level effects

Overall, our findings suggest that the involvement of parents modifies the functioning of their children's class both directly and indirectly, i.e., through the parents' direct influence on their own children and through the influence of their own children on other children in the class. In this section, we provide an evaluation of the effect of this combination of direct and indirect influences on the average outcomes of a class. Given the relatively small number of treated in a class, we generally have sufficient precision at class level, only for outcome variables that were hit by significant spillovers.

Specifically, Table 8 shows the results of regressing the average outcomes of a class on a dummy indicating whether the class has been randomly selected for the program or not. This confirms that selecting a class into the program increases significantly the average involvement of parents[14] (as measured by our global score) and improves significantly the average behavior of pupils, with a decrease of about 0.7 half day in average absenteeism during the last term, a decrease of about 3 percentage points in the proportion of sanctioned pupils, and an increase of about 9 percentage points in honors. Also, these reduced-form regressions confirm that the selection into the program generates improvements in average performance in Mathematics and French (as measured by teacher marks) that are positive, although not statistically significant at standard levels.

[14] For the sake of clarity, these regressions use a normalized version of the class-level parental involvement score. After computing the class-level average of individual scores based on available responses, the distance between the first (P25) and third (P75) quartile of the distribution of class scores is set to 1. Given this normalization, the estimated effect on parental involvement (.296) means that selecting a class into the program has an effect on average parental involvement which is equivalent to 29.6% of the difference observed between the 25% of classes where involvement is maximum (high-involvement classes) and the 25% classes where it is minimum (low-involvement classes).

V. Robustness analysis

The evidence presented in the previous section unanimously points to significant effects of the program across a wide range of outcomes. Some of these variables do not depend on subjective judgements, such as absenteeism, which represents an objective, and formalized measure, on which the program is consistently found to have both direct and indirect effects. But other outcomes are based on subjective assessment by teachers or by the school administration and it may be that such judgments were influenced by knowledge of pupils' treatment status. Indeed, telling the staff that some classes were selected and other were control classes could have, as such, an impact on their attitude towards the selected classes. For example, if teachers want the program to be a success (because it provides additional resources), they may tend to better assess selected classes, regardless of pupils' true outcomes. We would hence face a form of Hawthorne effect through the reporting behavior.

The existence of such Hawthorne effect, however, is less plausible than it seems. First, we have multiple view points and data sources on each kind of outcome, so that this sort of bias would need to affect everyone's judgment. Second, effects on both volunteer and non volunteer pupils become perceptible and significant at the end of the school year only, after pupils in selected classes were exposed to repeated interaction with their treated parents and peers. Table 9 compares treated and control samples for all measures that are available by the end of the first trimester. The end of term 1 is in the middle of the sequence of meetings, and approximately one month after the assignment lottery took place. This is the moment where the experimental context of the program is most salient to teachers and school staff. There is nonetheless no detectable impact of the program on any of the outcomes at this point of time, regarding either discipline, attitude in class or teacher marks. This constitutes maybe the most direct set of evidence that teacher assessments are not influenced by class assignment status.

Third, if teacher judgment was influenced by knowing which are the treated classes, this knowledge would impact volunteer and non-volunteer students similarly, thus leading to systematic evidence of peer effects in all dimensions. Furthermore, we should observe peer effects on the most subjective outcomes (such as teachers' marks), but not on the most objective ones (such as absenteeism). This is not what we find. In particular, there are no significant peer effects on attitude in class or marks, which are yet our most subjective outcomes. In contrast, we observe them mostly on objective outcomes. This clearly rules out the reporting bias assumption to explain peer effects.

Therefore, Hawthorne reporting effects provide a possible explanation for the observed impacts on these outcomes only insofar as teachers are able to identify not only treated classes, but also volunteer families within these treated classes. This could happen, but is very unlikely on a systematic basis: teachers were rarely involved in the meetings and, if so, only one or two of them were. Also, this bias should not form only after the first term.

Nevertheless, the rest of this section provides further evidence suggesting that differences between treated and untreated volunteers represent actual effects of the program. Specifically, we show that the direct effects of the program vary across pupils and that these variations are associated with actual variation in pupils' exposure to the program.

A. Impact on first-borns

First-borns are the first children, in a family, to enter middle-school. The school and its curriculum are new to their families and, as a consequence, parents of 6th graders who are the eldest child in the family are likely to get more new information from the program than parents of other 6th graders. Furthermore, these parents do not have to share their involvement effort across several middle-school siblings, nor is parents' influence mitigated by that of older children in the same family. In fact a significant fraction of first-borns are lone children for whom parental influence is likely to be strong and unmitigated.

Overall, we expect that if the impact on volunteers' children reflects increased levels of parental involvement, it is stronger when the child is the only child in middle school. In contrast, even if the judgment of teachers or the administration was biased towards treated volunteers, there is no reason why it should be even more biased when the pupil is a first-born or lone child. Therefore, stronger impacts on first-borns from volunteer families add strength to the evidence that the measured impacts are real and not the result of reporting bias. In the data, we can identify first-borns using information provided by families in the registration form before the start of the school year.

Table 10 presents the impact of treatment for all cognitive and non-cognitive variables, where the treatment dummy is interacted with volunteer status and a first-born dummy. Very strikingly, for all variables but the teacher mark in Mathematics (for which there is no impact whatsoever), point estimates are larger when the child is a first-born. This does not hold for non-volunteers, but there is no reason why it should, as spillovers derive from peer interaction, not from parental influence. The difference between first-borns and other children among volunteers is most of the time non-significant, as we lack power for such a detailed analysis, but it is very systematic on 9 variables that are different in their scope and in their source. We take this result as compelling evidence that differences between treated and control children of volunteer parents are driven by actual parental influence and not by reporting bias.

B. Long-term impacts

Finally, we present impacts on treated children in June 2010, at the end of academic year 2009-2010, that is about 18 months after the end of the program (December 2008). This is relevant per se, as it indicates whether the program has long-lasting effects. Long-term impacts also provide meaningful evidence on whether reported effects on treated pupils are robust to changes in class composition and pedagogical team.

As it turns out, the institutional practice in France is to reshuffle classes each year. Principals use this practice to influence future peer interactions and in particular to prevent dynamics of misbehavior, by separating clusters of troublesome pupils across different classes. From 6th to 7th grade, choice of language options and other special programs can further imply reallocation. In our specific case, this practice implied that, with few exceptions, class-groups, as defined in the year of intervention, no longer exist as a unit of interactions in the second year. During 6th grade, non-volunteers in the treatment group were exposed to 23.9% of class-mates whose families had been invited to the meetings (the proportion in control classes is 0%) whereas, as a consequence of attrition and reshuffling, during 7th grade, the average non-volunteer from the treatment group had only 4.3% more invited classmates than the average control non-volunteer.

Within this context, any effect which relies on peer exposure can no longer be detected with sufficient power. On the other hand, reshuffling makes it unlikely that teachers would be able to label pupils within mixed classes as belonging to the treatment group. Therefore, any long-term impact that we can observe on former pupils from treated classes is not affected by systematic response behavior.

Table 11 shows long-term impacts for available variables. We do expect both smaller effect sizes and larger standard errors, due to smaller sample sizes as a result of attrition.[15] On the volunteers sample, measured differences between (former) treatment and control groups tend to be slightly smaller and are less often significant (except for French mark and good conduct), but they are quite in line with the hypothesis of persistent benefits from increased

[15] As discussed above, attrition among 7th graders is not related to experimental arms.

parental involvement. The persistence of effects confirms that teacher grades are not influenced by biased reporting and strengthens the relevance and efficiency of this program.

VI. Conclusion

Governments and schools are increasingly enthusiastic about improving child outcomes through parenting programs. Parents are sometimes seen as a reserve of underutilized inputs, waiting to be called upon to contribute, at low cost, to the process leading to better school outcomes. This drive has delivered an abundance of policy initiatives, but remarkably little rigorous evidence on whether, and how, interventions fostering parental involvement in education are successful.

This article provides experimental evidence that middle-school classes in poor neighborhoods are less exposed to truancy and to episodes of misbehavior when parents receive invitations and support to become effectively involved in their children's school education. Teaching and learning activities take place in a more cooperative environment and pupils' marks improve.

Generally speaking, our results show that low levels of parental involvement are not a fatality in poor neighborhoods. Schools have the critical ability to trigger higher levels of involvement among some parents, and this can be enough to improve the outcomes for all children. On a population of pre-adolescents, our findings show that these interventions first and mostly deliver improvements in their behavior at school, which might be instrumental for gains in achievement. Furthermore, the results of this study not only stress the influence that parental behavior has on pupils, but also the role of peer-pressure in shaping pupils' behavior. Taken together, these two influences redefine high level of parental involvement as a club good at the class level, rather than a private investment: all pupils in a same class benefit from higher monitoring and involvement efforts by some parents.

A central debate in education is whether remedial programs should be targeted at the individual level. Our results on spillovers demonstrate some benefits of universal provision of parenting programs over the alternative of targeting at-risk families only. Providing support to entire communities has the advantage of minimizing the stigma associated with individual targeting. In the context of parenting programs, this does not come necessarily at the cost of smaller benefits for individual pupils, given the large spillovers at play.

Despite universal provision, the evaluated program had low take-up rates among potential beneficiaries. The net benefits from this intervention could probably be increased if more parents took up the program. Even the most rational among parents do not internalize the large positive externalities of their efforts on classmates. The provision of small and targeted incentives to compensate some parents for their effort to attend school-based meetings could be a way to improve take-up and cost-effectiveness.

Aizer, Anna (2004), "Home alone: supervision after school and child behavior", *Journal of Public Economics*, vol. 88, pp. 1835-1848.

Angrist, Joshua and Kevin Lang (2004), "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program", *American Economic Review*, vol. 94(5), pp. 1613-1634.

Avvisati, Francesco, Bruno Besbas and Nina Guyon (2010), "Parental Involvement in Schools: A Literature Review", *Revue d'Economie Politique*, vol. 120, no. 5, pp. 761-778, September-October 2010.

Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster and Stuti Khemani (2008), "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India", *American Economic Journal: Economic Policy*, vol. 2(1), pp. 1-30.

Björkman, Martina and Jakob Svensson (2009), "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda", *Quarterly Journal of Economics*, vol. 124(2), pp. 735-769.

Carrell, Scott E. and Mark L. Hoekstra (2010), "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids", *American Economic Journal: Applied Economics*, vol. 2(1), pp. 211-228.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach and Danny Yagan (2011), "*How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star"*, NBER Working Paper No. 16381

Cunha, Flavio and James J. Heckman (2008), "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation", *Journal of Human Resources*, vol. XLIII(4), pp. 739-782.

Desforges, Charles and Alberto Abouchaar (2003), "*The Impact of Parental Involvement, Parental Support and Family Education on Pupil Achievements and Adjustment: a Literature Review*", Department for Education and Skills, Research Report No. 433.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006), "The Effects of Cognitive and Non cognitive Abilities on Labor Market Outcomes and Social Behaviors", *Journal of Labor Economics*, vol. 24(3), pp. 411-482.

Hill, Nancy and Dinan Tyson (2009), "Parental Involvement in Middle School: A Meta-Analytic Assessment of the Strategies that Promote Achievement", *Development Psychology*, vol. 45(3), pp. 740-763.

Hoover-Dempsey, Kathleen V. and Howard M. Sandler (1995), "Parental Involvement in Children's Education: Why Does it Make a Difference?", *Teachers College Record*, vol. 97, pp. 310-331.

Hoover-Dempsey, Kathleen V. and Howard M. Sandler (1997), "Why Do Parents Become Involved in Their Children's Education?", *Review of Educational Research*, vol. 67, pp. 3-42.

Hoxby, Caroline (2000), "*Peer Effects in the Classroom: Learning from Gender and Race Variation*", NBER Working Paper No. 7867.

Lindqvist, Erik and Roine Vestman (2011), "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment", *American Economic Journal: Applied Economics*, vol. 3, pp. 101-28.

Moffitt, Robert A. (2001), "Policy Interventions, Low-Level Equilibria, and Social Interactions", pp. 45-82 of Durlauf, Steven N. and H. Peyton Young, (Eds.), *Social Dynamics*. The MIT Press, for Brookings Institution.

OECD (2010), “*PISA 2009 Results: What Students Know and Can Do - Student Performance In Reading, Mathematics and Science (Vol. I)*”, Organisation for Economic Co-Operation and Development.

Scott, Stephen, Thomas O'Connor and Annabel Futh (2006), “*What makes parenting programmes work in disadvantaged areas?*”, Joseph Rowntree Foundation.

Todd, Petra E. and Kenneth E. Wolpin (2007), “The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps”, *Journal of Human Capital*, vol.1, no.1, pp. 91-136.

Welsch, David M. and David M. Zimmer (2008), “After-School Supervision and Children's Cognitive Achievement”, *The B.E. Journal of Economic Analysis & Policy*, vol. 8, art. 49.

Table 1. Take-up Rates

	Test		Control	
	Volunteers	Non-vol's	Volunteers	Non-vol's
Initial workshops				
At least 1 debate	49.8	1.0	1.7	0.2
At least 2 debates	28.0	0.2	0.6	0.0
All 3 debates	12.3	0.1	0.6	0.0
Additional workshops				
Parenting	11.7	0.6	0.0	0.0
Internet	7.4	0.5	0.4	0.0
French as foreign language	3.2	0.0	0.0	0.0
Any of the above	16.4	0.9	0.4	0.0

All rates are expressed in percentage terms and are computed separately for each of the four groups.

Table 2. Impact of the Program on Parental Attitudes and Behavior

Dependent Variable	mean C	T - C		(se)	n.obs.
Panel A: Volunteers					
Parent self-assessment					
Global Parenting Score	0.005	0.152	**	(0.027)	758
School-Based Involvement Score	0.121	0.209	**	(0.048)	757
Home-Based Involvement Score	0.007	0.058	*	(0.034)	758
Understanding & Perceptions Score	-0.172	0.266	**	(0.074)	728
Reference teacher assessment					
Parent-School Interactions	0.794	0.034		(0.029)	735
Parental Monitoring of School Work	0.206	0.067	**	(0.034)	747
Panel B: Non volunteers					
Parent self-assessment					
Global Parenting Score	-0.013	0.014		(0.015)	1974
School-Based Involvement Score	-0.063	0.016		(0.025)	1973
Home-Based Involvement Score	-0.007	0.009		(0.022)	1973
Understanding & Perceptions Score	0.031	0.052		(0.040)	1917
Reference teacher assessment					
Parent-School Interactions	0.806	-0.032		(0.021)	2449
Parental Monitoring of School Work	0.213	0.022		(0.022)	2464

Score variables are averages of normalized and centered answers to questions in the corresponding section of the parent questionnaire. Coefficients can be interpreted as standardized effect-sizes. The first column is the mean of the row variable in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy, as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 3. Parental Attitudes and Behavior: Difference by Socio-Economic Status in Control Classes

Dependent Variable	mean NWC	WC - NWC		(se)
Parent self-assessment				
Global Parenting Score	-0.039	0.141	**	(0.023)
School-Based Involvement Score	-0.041	0.130	**	(0.037)
Home-Based Involvement Score	-0.045	0.177	**	(0.038)
Understanding & Perceptions Score	-0.011	-0.027		(0.083)
Reference teacher assessment				
Parent-School Interactions	0.788	0.100	**	(0.024)
Parental Monitoring of School Work	0.182	0.165	**	(0.033)

Score variables are averages of normalized and centered answers to questions in the corresponding section of the parent questionnaire. Coefficients can be interpreted as standardized effect-sizes. The first column is the mean of the row variable in the non-white collar group. Column WC-NWC displays the coefficient from the regression of the row variable on a white collar dummy, as well as controls for gender, birth rank, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 4. Impact of the Program on Pupil's Non-cognitive Outcomes

Dependent Variable	mean C	T - C	(se)	n.obs.
Panel A: Volunteers				
Objective measure				
Absenteeism	4.252	-1.098 *	(0.589)	713
Pedagogical team assessment				
Behavioral score	-0.028	0.154 **	(0.039)	962
out of which:				
Discipl. Sanctions	0.110	-0.046 **	(0.021)	917
Good conduct	0.290	0.065 *	(0.033)	649
Honors	0.750	0.120 **	(0.055)	923
Reference teacher assessment				
Behavior in class	0.530	0.083 **	(0.035)	750
School work	0.542	0.023	(0.030)	759
Panel B: Non-Volunteers				
Objective measure				
Absenteeism	4.298	-0.554 *	(0.296)	2402
Pedagogical team assessment				
Behavioral score	0.011	0.073 **	(0.029)	3155
out of which:				
Discipl. Sanctions	0.115	-0.024 **	(0.012)	3014
Good conduct	0.348	0.046 *	(0.025)	2190
Honors	0.794	0.061	(0.049)	2964
Reference teacher assessment				
Behavior in class	0.591	0.022	(0.025)	2486
School work	0.574	0.011	(0.023)	2485

Absenteeism is counted in half-days; behavioral score is an average of normalized and centered dummies for sanctions, good conduct and honors; all other variables are dummies. The first column is the mean of the row variable in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy, as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 5. Effects on Non-Treated pupils: Dose-Response Relationship

	Absenteeism	Behavioral score	Discipl. Sanctions	Good conduct	Honors	Behavior in class	School work
T - C (NV): Many vol's	-.764**	.119**	-.035**	.058*	.176**	.057*	.071**
	(.350)	(.034)	(.016)	(.030)	(.054)	(.034)	(.027)
T - C (NV): Few vol's	-.183	-.003	-.005	.028	-.161	-.040	-.096**
	(.663)	(.054)	(.016)	(.049)	(.100)	(.043)	(.046)
Many vol's	.512	-.139**	.029	-.020	-.316**	-.063	-.104**
	(.503)	(.048)	(.020)	(.041)	(.083)	(.042)	(.034)
n.obs.	2402	3155	3014	2190	2964	2486	2485

This table presents an augmented version of regressions in Table 4 on the sample of Non-Volonteers only. Each column is a different regression. The treatment variable is fully interacted with dummies for low (first tercile) and high (second and third tercile) proportion of volunteers in the class. Other variables (not reported) are controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 6. Impact of the Program on Pupil's Cognitive Achievement

Dependent Variable	mean C	std C	T - C		(se)	n.obs.
Panel A: Volunteers						
Teacher marks						
French	10.736	3.717	0.456	**	(0.192)	897
Maths	10.974	4.254	0.124		(0.221)	899
Average Mark (all subjects)	11.586	2.878	0.240	**	(0.118)	902
Reference teacher questionnaire						
Progress over the school year	0.548	0.488	0.057	*	(0.033)	760
Uniform tests						
French	-0.067	1.000	0.020		(0.060)	801
Maths	-0.077	1.004	-0.035		(0.064)	792
Panel B: Volunteers						
Teacher marks						
French	11.099	3.717	0.053		(0.144)	2938
Maths	11.087	4.254	0.047		(0.149)	2964
Average Mark (all subjects)	11.717	2.878	0.032		(0.094)	2966
Reference teacher questionnaire						
Progress over the school year	0.612	0.488	0.002		(0.023)	2499
Uniform tests						
French	-0.009	1.000	0.040		(0.043)	2614
Maths	0.006	1.004	-0.002		(0.046)	2607

Teacher marks are /20; Average mark is an average of all subjects weighted by class hours; progress assessment is a dummy variable; uniform tests are normalized. The first column is the mean of the row variable in the control group and the second column is the empirical standard-error in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy, as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 7. Relationship between Cognitive and Non-cognitive Outcomes

	French Mark	Maths Mark	French Test Score	Maths Test Score
Discipl. Sanctions	-.426** (.059)	-.348** (.045)	-.050 (.062)	.039 (.053)
Absenteeism	-.042** (.003)	-.033** (.004)	.000 (.003)	.003 (.004)
French Test Score	.564** (.017)			
Maths Test Score		.681** (.019)		
French Mark			.625** (.022)	
Maths Mark				.764** (.021)
School F.E.	Yes	Yes	Yes	Yes
n.obs.	2609	2662	2609	2662
R sq.	.50	.63	.46	.58

French and Maths marks are teacher marks /20; French and Maths test scores are normalized uniform tests; Disciplinary sanctions is a dummy and Absenteeism is measured in half-days. Each column is a separate regression. Only control classes are used. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 8. Class-level Analysis

	T - C		se	n.obs.
Parenting Score	.293	**	(.117)	171
Non-Cognitive Outcomes				
Absenteeism	-.718	**	(.332)	142
Behavioral score	.096	**	(.029)	183
Discipl. Sanctions	-.031	**	(.013)	175
Good conduct	.040		(.027)	130
Honors	.089	*	(.046)	175
Behavior in class	.039		(.027)	146
School work	.015		(.024)	147
Cognitive Achievement				
French Mark	.201		(.160)	174
Maths Mark	.055		(.168)	175
Average Mark (all subjects)	.092		(.096)	175
French Test Score	.029		(.045)	165
Maths Test Score	-.002		(.049)	168

Column T-C displays the coefficient from the regression of the row variable's class-level mean on a treatment class dummy. All regressions include school fixed effects, as well as controls for average first-term marks in French and Math and for the class proportion of female, firstborn, white-collar, scholarship recipient and grade repeating pupils. Mean parenting scores are normalised so that the distance between the first (P25) and third (P75) quartile of the distribution of class scores is set to 1. The remaining row variables are simple class-level means of variables used in Tables 4 and 5. Each line corresponds to a separate regression. *: significant at 10% level; **: significant at 5% level.

Table 9. Effects of Being Selected in Treatment Group on First-Term Outcomes

Dependent Variable	mean C	T - C	(se)	n.obs.
Panel A: Volunteers				
Absenteeism	1.090	-0.132	(0.151)	759
Discipl. Sanctions	0.099	-0.030	(0.021)	874
Good conduct	0.370	0.016	(0.041)	690
Honors	0.956	-0.039	(0.074)	916
French Mark	11.927	-0.363	(0.243)	907
Maths Mark	12.252	-0.092	(0.253)	908
Average Mark (all subjects)	12.472	-0.071	(0.165)	914
Panel B: Non-Volunteers				
Absenteeism	1.089	0.090	(0.113)	2717
Discipl. Sanctions	0.080	-0.011	(0.011)	2920
Good conduct	0.412	0.024	(0.027)	2142
Honors	0.964	0.038	(0.038)	3034
French Mark	11.963	-0.146	(0.156)	2962
Maths Mark	12.541	0.089	(0.178)	2995
Average Mark (all subjects)	12.584	-0.035	(0.094)	3008

This table presents variables identical to Tables 4 and 5, but *measured as of the end of the First Term*. The first column is the mean of the row variable in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy, as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 10. Subgroup Analysis: Effects on First-borns and Other Children

Panel A: Non-Cognitive Outcomes

	Absenteeism	Behavioral score	Discipl. Sanctions	Good conduct	Honors	Behaviour in Class	School Work
T - C (V): first-borns	-1.653**	.204**	-.069**	.072	.163**	.149**	.102**
	(.746)	(.051)	(.025)	(.046)	(.082)	(.045)	(.041)
T - C (V): others	-.495	.096	-.019	.058	.071	.010	-.065
	(.861)	(.059)	(.030)	(.047)	(.086)	(.050)	(.046)
first-borns	.097	-.009	.016	-.001	.038	-.009	.052
	(.889)	(.064)	(.029)	(.053)	(.096)	(.046)	(.047)
n.obs.	713	962	917	649	923	750	759
p-value: differential effects on first-borns	.296	.167	.172	.831	.471	.034	.008

Panel B: Cognitive Achievement

	French Mark	Maths Mark	Avg. Mark (all subjects)	Progress (ref.tch.ass.)	French Test Score	Maths Test Score
T - C (V): first-borns	.328	.201	.262*	.101**	-.053	.006
	(.233)	(.260)	(.144)	(.043)	(.080)	(.077)
T - C (V): others	.611**	.031	.213	.008	.103	-.080
	(.253)	(.295)	(.166)	(.045)	(.081)	(.087)
first-borns	.525**	.057	.141	.066	.189**	-.090
	(.229)	(.274)	(.155)	(.046)	(.092)	(.081)
n.obs.	897	899	902	760	801	792
p-value: differential effects on first-borns	.339	.611	.806	.117	.149	.401

This table presents an augmented version of regressions in Tables 4 and 5 on the sample of Volonteers only. Each column is a different regression. The treatment variable is fully interacted with dummies for first-born within the family. Other variables (not reported) are controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table 11. Long Term Effects (18 Months After the Intervention)

Dependent Variable	mean C	std C	T - C	(se)	n.obs.
Panel A: Volunteers					
Non-Cognitive Outcomes					
Absenteeism	4.456	8.085	-0.539	(0.493)	714
Good conduct	0.287	0.475	0.095 **	(0.038)	612
Teacher marks					
French	10.775	3.859	0.340 *	(0.176)	772
Maths	10.485	4.382	0.022	(0.226)	768
Average Mark (all subjects)	11.343	2.958	0.193	(0.131)	774
Panel B: Volunteers					
Non-Cognitive Outcomes					
Absenteeism	0.021	0.156	0.001	(0.004)	3090
Good conduct	0.333	0.475	0.020	(0.022)	2083
Teacher marks					
French	10.706	3.859	-0.004	(0.151)	2552
Maths	10.418	4.382	0.158	(0.173)	2538
Average Mark (all subjects)	11.379	2.958	0.015	(0.108)	2561

This table presents some of the variables in Tables 4 and 5 but *measured at the end of the year following the intervention*. The first column is the mean of the row variable in the control group and the second column is the empirical standard-error in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy (*as of the year of the intervention*), as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table A1. Differences in Pre-Treatment Characteristics

	mean C	T - C	(se)	n.obs.
Panel A: Volunteers vs. Non-Volunteers				
Parents				
Employment status	0.86	-0.004	(0.013)	4276
Two-parents Household	0.73	0.046 **	(0.016)	4308
White-collar	0.18	0.028 *	(0.015)	4308
Scholarship recipient	0.32	0.020	(0.018)	4308
Children				
Girl	0.49	-0,036 *	(0.019)	4308
First-born	0.54	0.002	(0.018)	4308
Past Grade Repetition	0.26	-0.021	(0.015)	4308
French Test Score (Sept. 08)	0.04	-0,073 *	(0.041)	3820
Maths Test Score (Sept. 08)	0.02	-0.017	(0.044)	3831
Panel B: Test vs. Control Volunteers				
Parents				
Employment status	0.85	-0.020	(0.021)	970
Two-parents Household	0.76	-0.024	(0.024)	970
White-collar	0.21	-0.015	(0.024)	970
Scholarship recipient	0.35	-0.038	(0.025)	970
Children				
Girl	0.46	-0.006	(0.030)	970
First-born	0.54	0.016	(0.028)	970
Past Grade Repetition	0.25	0.026	(0.028)	970
French Test Score (Sept. 08)	-0.07	-0.012	(0.079)	903
Maths Test Score (Sept. 08)	-0.04	-0.057	(0.074)	914
Panel C: Test vs. Control Non-Volunteers				
Parents				
Employment status	0.86	0.005	(0.012)	3306
Two-parents Household	0.73	0.010	(0.015)	3338
White-collar	0.18	0.006	(0.013)	3338
Scholarship recipient	0.32	0.021	(0.015)	3338
Children				
Girl	0.49	0.007	(0.012)	3338
First-born	0.54	-0.025	(0.018)	3338
Past Grade Repetition	0.26	-0.014	(0.017)	3338
French Test Score (Sept. 08)	0.04	-0.041	(0.049)	2917
Maths Test Score (Sept. 08)	0.02	0.032	(0.042)	2917

All variables except test scores are dummies measured before the beginning of the year; French and Maths test scores are uniform standardized test submitted at the beginning of the year. The first column is the mean of the row variable in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy, as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.

Table A2. Response and Attrition Rate (7th grade measures)

	response rates			observations		within-school attrition	
	mean C	T - C	(se)	individuals	schools	class	ind.
Non-Cognitive Outcomes							
Absenteeism	0.71	-0.003	(0.012)	3075	31	0.08	0.22
Good conduct	0.62	0.009	(0.021)	2695	31	0.10	0.29
Teacher marks							
French	0.77	-0.000	(0.012)	3324	34	0.08	0.23
Maths	0.76	0.005	(0.013)	3306	34	0.08	0.23
Average Mark (all subjects)	0.77	0.000	(0.012)	3335	34	0.08	0.23

This table presents response rates for the variables used in Table 11 (measured at the end of the year following the intervention). The explained variable is a dummy for the availability of information for each of the row variables. The first column is the mean of this variable (thus the response rate) in the control group. Column T-C displays the coefficient from the regression of this variable on a treatment class dummy (as of the year of the intervention). Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. Individuals and schools observations indicate the number of schools for which information could be obtained and the number of resulting individuals. Within-school attrition gives the attrition rate of entire classes within available schools and of individuals within available classes. *: significant at 10% level; **: significant at 5% level.

Table A3. Response Rates to Parent Questionnaire

	response rates			respondents		within-school attrition	
population	mean C	T - C	(se)	individuals	schools	class	ind.
non volunteers	0.61	-0.029	(0.028)	1974	31	0.07	0.32
volunteers	0.66	-0.022	(0.037)	627	31	0.11	0.25
volunteers (incl. call-back)	0.80	-0.023	(0.028)	758	31	0.02	0.16

This table presents response rates to the parent questionnaire. The explained variable is a dummy for the availability of quetionnaire. The first column is the mean of this variable (thus the response rate) in the control group for each of the populations defined in the rows. Column T-C displays the coefficient from the regression of this variable on a treatment class dummy. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classese are in parenthesis. Individuals and schools observations indicate the number of schools for which information could be obtained and the number of resulting individuals. Within-school attrition gives the attrition rate of entire classes within available schools and of individuals within available classes. *: significant at 10% level; **: significant at 5% level.

Table A4. Response Rate (Administrative data)

	response rates			observations		within-school attrition	
	mean C	T - C	(se)	individuals	schools	class	ind.
Non-Cognitive Outcomes							
Absenteeism	0.73	0.001	(0.006)	3115	26	0.07	0.06
Behavioral score	0.96	-0.001	(0.006)	4117	34	0.08	0.04
Discipl. Sanctions	0.91	-0.003	(0.006)	3931	32	0.07	0.04
Good conduct	0.64	0.014	(0.023)	2839	27	0.17	0.08
Honors	0.91	-0.022	(0.014)	3887	33	0.08	0.06
Cognitive Achievement							
French Mark	0.88	0.015	(0.017)	3835	33	0.09	0.07
Maths Mark	0.89	0.001	(0.013)	3863	33	0.09	0.07
Average Mark (all subjects)	0.89	0.002	(0.015)	3868	33	0.09	0.07
French Test Score	0.81	-0.020	(0.024)	3415	32	0.13	0.13
Maths Test Score	0.81	-0.011	(0.017)	3399	32	0.11	0.15

This table presents response rates for the administrative variables used in Tables 4 and 5. The explained variable is a dummy for the availability of information for each of the row variable. The first column is the mean of this variable (thus the response rate) in the control group. Column T-C displays the coefficient from the regression of this variable on a treatment class dummy. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. Individuals and schools observations indicate the number of schools for which information could be obtained and the number of resulting individuals. Within-school attrition gives the attrition rate of entire classes within available schools and of individuals within available classes. *: significant at 10% level; **: significant at 5% level.

Table A5. Response Rate (Reference Teacher Questionnaire)

	response rates			observations		within-school attrition	
	mean C	T - C	(se)	individuals	schools	class	ind.
Behavior in class	0.78	-0.039	(0.034)	3236	30	0.19	0.07
School work	0.78	-0.027	(0.033)	3244	30	0.18	0.07
Progress	0.78	-0.027	(0.033)	3259	30	0.18	0.07
Parent-School Interactions	0.75	-0.008	(0.035)	3184	30	0.18	0.09
Parental Monitoring of School Work	0.76	-0.018	(0.032)	3211	30	0.19	0.08

This table presents response rates to the teacher questionaire. The explained variable is a dummy for the availability of information for each of the row variable. The first column is the mean of this variable (thus the response rate) in the control group. Column T-C displays the coefficient from the regression of this variable on a treatment class dummy. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. Individuals and schools observations indicate the number of schools for which information could be obtained and the number of resulting individuals. Within-school attrition gives the attrition rate of entire classes within available schools and of individuals within available classes. *: significant at 10% level; **: significant at 5% level.

Table A6. Differences in Parents' Behavior Between Test and Control Volunteers (Raw Indicators)

Question	mean C	T - C	(se)
Global Parenting Score	0.005	0.152**	(0.027)
School-Based Involvement Score	0.121	0.209**	(0.048)
Several individual appointments with teachers	0.23	0.058*	(0.031)
Has attended parents/teachers meetings	0.80	0.076**	(0.026)
Has participated in parents' organizations	0.25	0.116**	(0.033)
Home-Based Involvement Score	0.007	0.058*	(0.034)
Precise knowledge of child's grades	0.44	0.040	(0.034)
Sometimes helps with homeworks	0.89	-0.006	(0.022)
Child does not watch TV daily after 9pm	0.81	0.036	(0.025)
Child spends less than 1 h/d on other screens	0.88	0.028	(0.020)
Understanding & Perceptions Score	-0.172	0.266**	(0.074)
Knowledge of optional courses offered	0.76	0.101**	(0.028)
Has never been anxious about violence	0.26	0.015	(0.031)
Clear ideas about high-school plans	0.26	0.044	(0.033)
Satisfied with school	0.81	0.060**	(0.021)
Never been summoned to the school	0.72	0.097**	(0.029)

Score variables are averages of normalized and centered answers to questions in the corresponding section of the parent questionnaire; other variables are dummies from the raw questions asked in the parent questionnaire. Coefficients can be interpreted as standardized effect-sizes. The first column is the mean of the row variable in the control group. Column T-C displays the coefficient from the regression of the row variable on a treatment class dummy, as well as controls for gender, birth rank, white collar, scholarship recipient, grade repetition, first-term marks in French and Math, and school fixed effects. Each line corresponds to a separate regression. Robust standard-errors allowing for correlated residuals within classes are in parenthesis. *: significant at 10% level; **: significant at 5% level.