The Making of Civic Virtues:
A School-Based Experiment in Three Countries

Simon Briole, Marc Gurgand, Éric Maurin, Sandra McNally,
Jenifer Ruiz-Valenzuela and Daniel Santín*

May 31, 2024

Abstract

This paper shows that schools can foster the transmission of civic virtues by helping students to develop concrete, democratically chosen, collective projects. We draw on a RCT implemented in 200 middle schools in three countries. The program leads students to conduct citizenship projects in their communities under the supervision of teachers trained in the intervention. The intervention caused a decline in absenteeism and disciplinary sanctions at school, alongside improved academic achievement. It also led students to diversify their friendship network. The program has stronger effects when implemented by teachers who are initially more involved in the life of the school.

---

*Simon Briole: Center for Environmental Economics Montpellier, Montpellier 34000, France, simon.briole@umontpellier.fr. Marc Gurgand: Paris School of Economics, Paris 75000, France, marc.gurgand@psemail.eu. Éric Maurin: Paris School of Economics, Paris 75000, France, eric.maurin@psemail.eu. Sandra McNally: London School of Economics, London WC2A 2AE, and University of Surrey, Surrey GU2 7XH, UK, S.Mcnally1@lse.ac.uk. Jenifer Ruiz-Valenzuela: Universitat de Barcelona and IEB, 585 08007 Barcelona, Spain, and Centre for Economic Performance, London School of Economics, London WC2A 2AE, UK, jruizv@ub.edu. Daniel Santín: Complutense University of Madrid, 28040 Madrid, Spain, dsantin@ccee.ucm.es.

1 Introduction

Even though there is no universally accepted definition, the civic sense in a society is commonly measured by the respect that citizens show to the rules of collective life, their involvement in the definition of these rules as well as the priority they give to general interest over private interests. In modern democracies, good citizenship is also commonly understood to include tolerance for the diversity of religions, sexual minorities and political opinions as well as support for the idea of equal rights for all citizens, regardless of gender, race and origin.

These civic virtues have long been identified as central to the stability of democratic societies and to their economic development.[1] The importance of cultivating them in younger generations is constantly reaffirmed, and most modern school curricula include a civic education program (Heater, 2004). However, the effectiveness of civic education in schools is still a matter of debate, and there is no real consensus on how best to teach civics in societies as diverse as modern ones (Campbell, 2019). At a deeper level, it is not even clear whether civic sense in modern societies can be considered a form of human capital that can accumulate or depreciate based on implemented policies, including educational ones. The fundamental institutions and values of modern democracies are being rejected by increasingly large sections of the population, despite an unprecedented rise in educational levels (Dalton, 2017; Carothers and O'Donohue, 2019).

In this article, we demonstrate that fostering altruism, tolerance, political participation, and respect for collective rules among young adolescents is achievable by assisting their teachers in implementing a pedagogy based on student empowerment, and integrating it into the design and implementation of concrete civic-oriented projects. These results are in line with the long-held hypothesis that schools are among the places where children can best develop their civic sense by learning to cooperate in practice around projects that they have chosen and that concern them directly (Dewey, 1915; Williams, 2003).

---

[1]See Putnam (1993), Tabellini (2008) or Guiso et al. (2011). For a discussion of the different liberal and republican conceptions of civic virtue, see e.g. Burtt (1993).

These findings are based on a large-scale randomized experiment conducted in a sample of more than 200 middle schools in three different countries (France, Greece and Spain). The intervention was designed in the aftermath of the Paris terrorist attacks in 2015 and is part of joint efforts in several European countries to promote civic spirit, religious tolerance, democratic values and equal rights on the old continent.

In these schools, during the 2018-2019 school year, about 320 teachers and 6,200 grade 8 and 9 students participated in the experiment. Half of participating schools within each country were randomly selected to implement the program (called *Active Citizenship Program*, hereafter ACT). Teachers from treatment schools first attended a two-day training program at the beginning of the academic year and then supervised the implementation of a concrete civic-oriented project with their students. Students first had to elicit a project democratically, and then run it over the school year. The idea behind this approach is that empowering students through the concrete practice of civic engagement and the exercise of democratic rules may help them develop the skills necessary to adopt civic behavior. A majority of the projects selected are designed to show solidarity and empathy towards people and students facing difficulties (for example, performing a small play in a hospital or retirement home; organizing a day to raise awareness of the problems faced by disabled people or migrants; organizing homework help for students in difficulty, etc.).

As stated in the pre-analysis plan, program effects are identified through pre- and post-intervention surveys that measure students' social and political engagement, their adherence to civic values as well as the size and diversity of their friendship networks. We also have information on student participation in the Global Climate Strike for Future, which took place on March 2019, at the end of the intervention. This provides us with a concrete measure of students' propensity for political participation, a dimension often particularly difficult to identify among young people. On the French site of the experiment, we also use administrative data on unjustified absences, late arrivals and disciplinary sanctions, of which the most severe (exclusions) arise from acts of incivility or violence towards teachers and other students. These data provide us with objective measures of students' ability to respect teachers, classmates, and the rules of school life, namely

some of the most basic dimensions of citizenship that a middle-school student may or may not demonstrate. We also have administrative data on grades given by teachers, which allows us to assess the impact of the treatment on students' academic engagement, a final dimension mentioned in the pre-analysis plan.

The comparison of treatment and control students before and after the intervention reveals that the intervention produced a significant improvement in student behavior. We observe a decline in unjustified absences and late arrivals as well as in disciplinary sanctions, especially the most serious ones. The summary index of student behavior increases by 27% of a SD. We also observe a significant improvement in grades assigned by teachers (13% of a SD), including in subjects where no teachers were involved in the experiment, i.e., in subjects where teaching methods and curriculum content remain unchanged. The same comparison of treated and control students also reveals that the intervention helps students to develop their network of friends and to open up to peers with different profiles. Specifically, we show that the program has a sizeable, positive impact on the number of 'socially different' friends (as measured by gender, social origin or geographic origin), while having no effect on the number of 'socially similar' friends. This greater openness may be due in part to project themes chosen by project participants, which are often about fighting prejudice and discrimination based on gender or origin. It is likely also due to teachers' decisions to structure small groups to be more diverse than initial friendship networks. In fact, we show that the treatment effect on network heterophily is stronger when participants are assigned to a small group with a higher share of socially different classmates.

Hence, the introduction of an active-learning method for civic education is accompanied by a significant improvement in a range of objective measures of students' compliance with school rules, their academic engagement and their ability to interact with people who do not share their cultural or social background. On a more subjective level, student surveys conducted at the beginning and end of the academic year reveal that students who benefit from the intervention feel more able to act on political and social issues. In line with this empowerment, they do show greater levels of social engagement in and out of school, as well as higher participation in the

Global Climate Strike for Future. Very similar results were found at the Greek, Spanish and French sites.

In the pre-analysis plan written before the intervention, we hypothesized that the program's effects might be particularly strong on the most successful and committed students. We test this hypothesis both in the way set out in the pre-analysis plan and using Machine Learning techniques designed to explore the heterogeneity of treatment effects. These analyses confirm that the students most strongly affected by the intervention tend to be those who had the best grades and were already the most altruistic and interested in politics at baseline. Above all, our heterogeneity analysis reveals the key role played by teachers. When we measure the involvement of teachers by their baseline participation in the various school councils, we see that treatment effects on students' civic behavior or political participation are much larger when the program is implemented by the most involved teachers, in line with the recent literature on the effect of teachers on the transmission of non-cognitive skills (Jackson, 2018; Kraft, 2019; Petek and Pope, 2022). Student-centered pedagogy is hard to use efficiently because teachers must balance student autonomy with the right level of guidance, a question that has long been debated in educational science (Lazonder and Harmsen, 2016). The complexity of such teaching places high demands on teachers, with strong heterogeneity in their ability to meet these demands (Crawford, 2007). In the context of civic education, it seems that the teachers who are most civic-minded and accustomed to collective deliberation are also the ones best equipped to tackle this pedagogical challenge.

Importantly, the treatment effect heterogeneity across teachers does not appear to be mediated by the nature of the project implemented. Similarly, despite differences in the types of projects chosen, treatment effects are not significantly different across countries. The key success factor of the program does not seem to be the precise theme of the project chosen, but the fact that the theme is debated and chosen democratically by local students.

To our knowledge, our paper is one of the few to provide clean evidence on the causal effects of civic-specific education on students' behavior and civic outcomes as well as on the distribution of these effects across students. It contributes to the long-standing economic literature that ex-

plores the relationship between education and citizenship (e.g., Milligan et al. (2004); Dee (2004); Larreguy and Marshall (2017)). It also contributes to ongoing debates on the impact that civic education in schools can have on students' attitudes and values (Campbell et al., 2012; Isac et al., 2014; Donbavand and Hoskins, 2021). Drawing on a large-scale randomized experiment conducted in the US, Green et al. (2011) find that an enhanced civics curriculum that tries to foster awareness, as well as understanding of constitutional rights and civil liberties, increases knowledge about civil liberties but does not change attitudes. Although the curriculum emphasizes classroom discussion and active student participation (similarly to this study), the pedagogy is not structured around student-designed projects as in the ACT intervention.

Our paper also adds to a body of evidence obtained from the evaluation of student-based programs in civic education, such as Democracy Prep (Gill et al., 2020) or Student Voice (Syvertsen et al., 2009). The former study is of a single charter school in New York City that describes its mission as 'to educate responsible citizen scholars for success in the college of their choice and a life of active citizenship'. The authors find positive effects on later voter participation and interpret this as evidence for the efficacy of a school culture built around a civics-orientated mission. In a related contribution Cohodes and Feigenbaum (2023) evaluate the effect of attending charter schools in Boston on academic and voting outcomes. They find a substantial effect on the probability of voting in the first presidential election after a student turns 18, though this is only for girls. The current study is more focused on the effect of a specific civics curriculum and pedagogy, as well as considering a wider set of outcomes (including school behavior, democratic participation and heterophily of friendship networks). The idea that learning by doing is an essential method of learning, including learning the skills and behaviors that make life in society possible, goes back to Aristotle and has a very long history in both economics and philosophy (Dewey, 1897; Arrow, 1962).

On a broader level, we contribute to the literature that explores the formation of pro-social values and skills during youth. In line with one central assumption of Cunha and Heckman (2007), several recent studies focus on early childhood and confirm that programs that enrich young children's in-

teractions with their environment and develop their ability to understand the perspectives of others are able to promote pro-social behavior (Cappelen et al., 2020; Alan et al., 2021). Our article demonstrates that it is still possible to continue to develop pro-social values and skills in older children during adolescence.

The remainder of our paper is structured as follows. In Section 2, we present the ACT intervention, the context in which it was formulated, its content and its experimental design. In Section 3, we describe the data as well as the measures we are using. Section 4 shows the main results of the intervention while Section 5 develops two heterogeneity analyses: as set out in the pre-analysis plan; and a data-driven version, using Machine Learning techniques. We conclude in Section 6.

2 Institutional Context and the ACT Intervention

The experiment took place in three countries where the school curriculum emphasizes respect for collective rules and civic participation, and where students have the opportunity to experience a commitment to public interest by being elected as student representatives on class or school councils.[2] However, in these three countries civics is still not taught as such at the university level and the teachers who teach it in middle schools or high schools are not specialists in the subject. As a result, civic education is often taught in a very teacher-centered way, following textbooks closely, with teachers not mastering the subject enough to venture into too free an interaction with students (Bozec, 2016). Furthermore, in all three countries, there are recurring debates about the legitimacy of the State to impose values in civic education, particularly when these values may conflict with traditional or religious norms of at least some sections of society.

The hypothesis motivating the ACT intervention is that these difficulties can be overcome by shifting the focus of civic education from the transmission of civic knowledge and values, conveyed through teacher-centered

---

[2]Additional information on civic education in France, Greece and Spain are provided in Appendix A.

practices, to the development of concrete civic skills and behavior through the use of student-centered teaching methods, where the topics studied are the result of debate and democratic choice of the students themselves. The aim is to give students more ownership over the learning process and reduce distrust arising from the imposition of particular viewpoints.

2.1 The ACT intervention

The ACT intervention took place during the 2018-2019 school year. It is based on two basic principles: empowering middle-school students, and having them define and implement concrete collective projects, designed to show solidarity and empathy towards people in their environment.

2.1.1 Teacher's training sessions

Teachers are trained at the beginning of the school year over two days. These training sessions are reserved for teachers in the treatment group. The aim of these sessions is to promote teaching techniques that give students the opportunity to debate, exercise their autonomy and make collective choices. Once the content of the program and the training material has been defined, the only marginal cost of the intervention is the remuneration paid to trainers. The two-day sessions had an average of 8 teachers per trainer. Trainers are typically former teachers with a special certificate. Assuming they are paid similarly to teachers, or slightly above, this implies a cost per trained teacher in the 50-100 euros range, depending on the country specific wage scheme, or 2-5 euros per student.

2.1.2 ACT projects

The ACT projects were designed and implemented by grade 8 and 9 students in treatment classes during the 2018-2019 school year, from October to April. During the first phase of implementation, students are assigned to groups of 4 to 5 students. Each group is tasked with identifying a project for the class, that must deal with one of three themes: discrimination, social inclusion or cultural diversity. The students must also specify the group of people for whom the project is primarily intended, such as students in other classes in the same school, students in another school, community

8

groups, etc. Students must also state the objectives of their project such as : to raise public awareness, inspire change, promote dialogue, bring people together, etc.

Once this preparatory work is complete, each group presents its project to the class. A vote is then organized to elect the project that the whole class will carry out. Following the vote, the teacher helps the students develop an action plan and allocate tasks amongst themselves. Teachers are advised to spend about 20 hours with their students on the project. These hours are taken from the time usually allocated to civic education.

Table 1 shows that the elected projects cover all three possible themes: fighting discrimination (64%), social inclusion (53%) and cultural diversity (29%).[3] They target groups that are most often victims of violence and discrimination, such as women, sexual minorities or people of immigrant background. Out-of-school projects (44% of the total) include collecting food from supermarkets for homeless people; visiting a retirement home so as to perform a short play; or visiting a nearby elementary school to hold a workshop related to gender equality. In-school projects include producing posters to speak out against xenophobia, racism or discrimination to their schoolmates; setting up an online quiz to detect isolation and following this up with activities to encourage interaction among students; helping non-native speakers to overcome difficulties with the local language or organizing private tutoring for those of their schoolmates with academic difficulties.

2.2 Experimental design

Recruitment of volunteer middle-schools took place between February and June 2018. It was restricted in each country to public schools in a subset of educational regions.[4] Only volunteer schools and volunteer teachers entered the experimental design. Then, between July and September 2018, public authorities collected the names of the teacher(s) and students that the volunteer schools planned to include in the program, should the school be allocated to the treatment group, and communicated them to the

---

[3]Note that each project can correspond to several themes and/or target populations.

[4]The list of regions and further details on the recruitment protocol are given in online Appendix B.

evaluation team. A total of 270 schools expressed their interest in participating in ACT and provided those lists. Randomization then took place between September and October 2018. We first formed school strata (from 2 to 6 schools each) on the basis of similar characteristics of the schools (e.g. location and size of schools, average student social and immigration background or metrics of student achievement in previous years). Schools were then randomly allocated to the treatment and control groups within strata. Subsequently, a small number of schools stopped responding to surveys and participating in the program (1 school in France, 15 in Greece and 4 in Spain). We drop the different strata to which these schools belonged. In the end, we kept a total of 85 strata comprising 108 treatment schools and 109 control schools. This paper focuses on this set of schools. They correspond to a total of 323 volunteer teachers and 6211 listed students, of which 3194 and 3017 are in the treatment and control schools respectively (see Appendix Table D1). Although this sample was not designed to be representative of middle school students in each country, the baseline characteristics of the students in our sample are not very different from population averages, as estimated by the PISA 2018 surveys (see online Appendix B). In the following, whenever we use this basic sample to assess the effect of being assigned to the treatment group on a particular endline outcome, we will first check that the response rate for this specific outcome is not affected by the treatment (no differential attrition) and that, among respondents, there is no correlation between the different baseline characteristics and the probability of being treated.

3   Data and measurement

Our analysis draws on administrative data on student behavior and academic performance as well as on online surveys designed to measure the impact of the intervention on students' civic-mindedness, students' friendship networks and teachers' practices. Data from online surveys were collected on all three sites of the experiment, at the beginning and the end of the school year. These surveys targeted all volunteer teachers and all students that appeared on the class lists sent by the schools before ran-

domization. Administrative data are only available for the French site of the experiment, but they have the advantage of being exhaustive.[5]

3.1 Administrative data

For each of the students participating in the French part of the experiment, we had access to administrative data on students' unjustified absences and late arrivals as well as on the number and the nature of disciplinary sanctions students were subject to over the school year. It should be stressed that the measurement of absenteeism and disciplinary problems by the school administration is very objective as it corresponds to strong legal obligations.[6]

For each of the students in the French site of the experiment, we also observe the grades received at the end of each term in each of the 11 subjects that the students take in middle school. We observe the grades received in the first quarter (before the implementation of the program) and at the end of the last quarter (after the implementation), so it is possible to test the impact of the program on end-of-year grades holding initial grades constant. As discussed in the online Appendix C, the curriculum of several of these subjects (such as Sport or French language) requires the development of social skills similar to those promoted by the ACT program, and we can expect to observe a treatment effect on the assessments received in these subjects.

3.2 Teacher Survey

The teacher baseline survey includes questions on teacher demographics, professional background and civic engagement at school and outside of school. In particular, we know whether the teacher participates in the various school councils (board of directors, disciplinary council, etc.). The teacher endline survey provides information on whether the teacher participated in the ACT training sessions, whether he or she has implemented

---

[5]We did not collect information on classes with non-volunteer teachers of participating schools, so it will not be possible to test for spillover effects on their students.

[6]See online Appendix C.

a citizenship project, distinguishing between those developed within the framework of the ACT program and others.

We also collected information about the teacher's teaching practices. We have information on the frequency with which teachers have students work in small groups, set up whole-class discussions or have students make oral presentations. We also have information on how often students suggest classroom activities themselves or express their opinions about lessons. Using these questions, we constructed a summary *Teaching Practices* index that captures the extent to which teachers use student-centered methods that fit the principles exposed during the ACT training sessions.[7]

3.3 Student Survey

The student baseline survey includes questions on students' own demographic characteristics and family background, and whether students had ever been elected as a representative on the class council or the school parliament/student council. In the remainder of the paper, we use student experience as a representative to measure the student's civic-mindedness at baseline.[8] In addition, specific questions were asked at the end of the school year to collect information on students' participation in citizenship projects.

Student surveys also collected information on friendship networks at the beginning and end of the school year. In each school, each student on the list was presented with the names of all the other students on the class list and asked whether or not each person was a friend. Student surveys also measured the effect of the intervention on social engagement, tolerance and support for equal rights, the main attitudes the ACT program aims to improve. We constructed an index of social engagement using student responses to questions about: (i) student civic engagement in their school

---

[7]Throughout the paper, whenever we construct an index from a set of outcomes, we follow Anderson's procedure (Anderson, 2008). This procedure involves: (i) switching the sign of outcomes where necessary so that the positive direction always indicates a "better" outcome; (ii) normalizing each outcome and (iii) computing a weighted average of normalized outcomes to build the corresponding index, where each weight corresponds to the inverse of the covariance matrix of the outcomes.

[8]Table D2 in the online appendix confirms that there are indeed strong differences in baseline levels of civic skills between representatives and non-representatives.

community (such as tutoring of younger students); (ii) civic engagement outside of school (such as volunteering in an association aimed at helping the community); and (iii) altruistic behavior (adapted from the Self-Report Altruism scale (Rushton et al., 1981)). To measure tolerance, students were asked the extent to which they favor social interactions with individuals who share their views on religion. Finally, support for equal rights is based on survey questions measuring the extent to which students agree with general statements on the equality of rights between citizens. Based on our indexes of social engagement, tolerance and support for equal rights, we constructed a summary *Civic Attitudes* index.[9]

To assess the impact of the intervention on students' ability to participate in democratic processes, we constructed a *Political Self-Efficacy* index, using students' responses to a standard set of questions about their political knowledge and their self-confidence in talking about and participating in politics, adapted from Niemi et al. (1991). We also constructed an *Interest in Politics* index using students' responses to questions such as how often they talk about political issues with their parents and friends, or the likelihood that they will take part in different forms of traditional political engagement in adulthood. Finally, to get around the difficulty of measuring the political participation of 14-years-old students, we take advantage of the Global Climate Strike, which took place on March 2019 in all three countries during the experiment: we use student participation in the strike as a concrete measure of political participation. This measure has the notable advantage of being well aligned with new forms of youth political participation (Dalton, 2008). Based on our indexes of *Political Self-Efficacy* and *Interest in Politics* as well as on a dummy indicating student participation in the Global Climate Strike, we constructed a summary *Democratic Participation* index. Table D3 in the online appendix provides descriptive statistics showing that female and high SES students show a higher level of civic-mindedness than male and low-SES students on most summary indices, in line with international surveys (Schulz et al., 2018).

---

[9]Our indexes of social engagement, tolerance and support for equal rights are based on scales that are widely used in the political science and social psychology literature, the psychometric validity of which has been largely documented. For each scale, we further check that Cronbach's alphas are above 0.7 in all national samples.

4 Results

In this section, we provide an evaluation of the effects of being assigned to the treatment group on the implementation of the program on the one hand and on students' outcomes on the other. We base our analyses on the following regression model:

$$Y_{isr} = \alpha + \beta T_s + \gamma X_{is} + \delta_r + \epsilon_{isr} \tag{1}$$

where Y_{isr} is the outcome of interest for student (or teacher) i in school s and strata r. T_s is the binary treatment indicator, which equals one if school s is in the treatment group and zero otherwise, and X_{is} is a vector of controls selected in each individual regression through a Lasso procedure (Belloni et al., 2014). Potential controls include student pre-determined characteristics (gender, age, geographical origin, family background, experience as representative, civic outcomes at baseline, grades in the first quarter) as well as teacher sociodemographic and professional characteristics (age, gender, experience, seniority, certification level, experience with citizenship teaching and training, implementation of citizenship projects over the last two years, personal engagement for the community at school and outside of school) and class size. These controls may be selected at the individual student-level or averaged at the school-level.[10] Finally, δ_r represents a full set of dummies indicating the strata used for randomization and ϵ_{isr} the residuals. The estimated $\hat{\beta}$ is the intention to treat effect.

In the case of missing baseline data and complete endline data, we impute missing covariate values. For this, we replace missing values with their mean values and include dummies indicating missing values for each covariate. Our results are not sensitive to imputing the missing covariate values. When estimating the effect of the program on sub-indexes, we report p-values of the coefficient of the treatment variable adjusted for the False Discovery Rate, using the Benjamini and Hochberg (1995) procedure, in order to account for multiple hypotheses testing. Finally, following Abadie et al. (2023), we cluster standard errors at the school level in all regressions.

---

[10]We checked that the estimated effects on most outcomes of interest remained similar (and statistically significant) when all controls were removed. However, the use of controls provides generally more precise estimates.

4.1 Implementation of the program

Before estimating the effects of the program on students' civic outcomes, it is important to identify the extent to which the program was actually implemented in the treatment group and the changes this implementation induced in teacher practices. To explore these questions, we measured the effect of being assigned to the treatment group on (1) the probability that teachers completed the ACT-specific training at the beginning of the school year, (2) the probability that teachers had their students implement an ACT project during the school year, and (3) the type of pedagogy implemented by teachers during the school year. In addition, we measured the effect of the treatment on the probability that students took part in a citizenship project during the school year (whether through ACT or not).

For each of these four outcomes, the analysis is conducted on the sample of individuals who are observed at baseline and for whom the outcome is measured at endline. Appendix Tables E1 and E2 show that these working samples represent between 70% and 80% of the initial sample, but that there is no significant difference in missing rates between the treated group and the control group. These tables further show that there are no differences in baseline characteristics between the treatment and control groups for each of the four working samples.

Table 2 reports the impact of being assigned to the treatment group for each outcome under consideration. The table first confirms that the vast majority of volunteer teachers in treated schools participated in the fall training sessions and supervised the implementation of an ACT citizenship project during the following academic year. In contrast, the proportions of volunteer teachers who participated in the fall training sessions or supervised an ACT citizenship project are negligible in the control group.[11]

The comparison of teaching practices in the control and treatment schools confirm that teachers in treatment schools were indeed influenced by the ACT training. Using our Teaching Practices index, Table 2 shows that, on average, teachers from treated schools declare practices that better

---

[11]To implement the ACT citizenship projects, teachers in treated schools also appear to have followed the protocol provided during the training sessions. In particular, they declare having spent about 20 hours on the projects with the students, which is in line with training guidelines (see Table D4 in the appendix).

fit the principles exposed during the training sessions, by about 43% of a standard deviation (SD).

At the end of the school year, students were also asked whether they had taken part in a citizenship project in their school. Reassuringly, the average proportion of students who report having participated in a citizenship project is much higher in treated schools than in control schools (by about 42 percentage points). In treated schools, about three quarters of students report having participated in a citizenship project, compared to less than one third in control schools. The fact that the proportion of students who participated in a citizenship project does not reach one hundred percent in treated schools reflects the fact that some projects were aborted very early in the year and that some students eventually refused to participate or did not get involved. All such occurrences were observed in qualitative work. Conversely, the fact that the proportion of students who participated in a citizenship project was not negligible in the control group confirms that project-based pedagogy is not unknown to teachers, and suggests that a significant fraction of teachers in the treated group would have conducted a citizenship project anyway, even if they had not been assigned to that group.

4.2 Treatment effect on student behavior and academic achievement

In the previous section, we showed that treated classes are, as expected, classes where civic education based on group projects and student initiative is used much more extensively than in the other classes. In this section, we assess whether these changes in teaching methods have been accompanied by changes in student behavior and achievement. Specifically, we report the effect of the treatment on the number of unjustified absences, late arrivals, disciplinary sanctions as well as on grades assessed by teachers, as reported in students' official academic records at the end of each academic term. We distinguish between two levels of sanctions according to their severity. The most severe sanctions (exclusions) are pronounced against acts of incivility and violence towards teachers and schoolmates. In France, as in many other countries, the problems posed by these acts

are a central issue for the national education system, in part because they contribute to the deterioration of the working conditions of teachers and to their disaffection.

For each outcome, the analysis is conducted on the working sample of individuals who are observed in the baseline and for whom the outcome is measured at the endline. Appendix Tables E3 and E4 show that this working sample represents about 98% of the initial sample and that there is no significant difference in missing rates between the treated and control group. These tables also confirm that there is no difference in baseline characteristics between treatment and control groups for each outcome. Appendix Table F1 further confirms that there is no effect of treatment assignment on grades received in the first quarter of the school year, before program implementation.

Table 3 reports the impact of being assigned to the treatment group for each outcome (noting that all outcomes are scaled such that a positive sign denotes a better outcome). The regression results reveal a significant and positive impact on both behaviors and academic performance. In particular, the intervention resulted in a significant decrease in unjustified absences and a significant reduction in acts of violence and incivility that warrant the most serious sanction, namely exclusion from the school. We also detect a decrease in late arrivals, even though the effect is only marginally significant at standard levels. In the end, the intervention leads to an improvement of the *School Behavior* index of about 27% of a SD.[12] This result is consistent with the idea that promoting active-learning methods in civics can lead students to be more respectful of other members of the school community and the basic rules of school life. The comparison of the distribution of the *School Behavior* index for the treated and control groups further reveals that the intervention leads to an increase in the most civic-minded behaviors to the detriment of those just below or close to average (Figure 1 (a)).

---

[12]As this summary index is standardized, the estimated effect does not necessarily correspond to an average of the estimated effects on the primary outcomes. For reference, the estimated effect on this summary index is of the same order of magnitude as the average effect reported by Valdebenito et al. (2018) from their meta-analysis of 37 randomized interventions aimed at reducing school exclusions and suspensions.

The improvement in student behavior in the treated group was accompanied by an improvement in academic performance. The regression results in Table 3 shows an improvement of 13% of a SD in the average grades received post-treatment, during the last term of the school year. The comparison of the distribution of average grades for the treated and control groups further shows that the intervention leads to an increase in good and very good performances to the detriment of performances close to average (see Figure 1(b)).

To take one step further, Table 3 compares the estimated effects when we analyze volunteer and non-volunteer teachers separately. Unsurprisingly, the estimated impact is significantly higher when focusing on volunteer teachers, the majority of whom teach civics. But the estimated impact on teacher-assessed grades remains very significant (about 10% of a SD) when restricted to other teachers, those who did not volunteer for the program. This result is in line with the idea that improvements in grades are not a mere consequence of volunteer teachers wanting to promote the program, nor can they be interpreted as the consequence of the implementation of more student-friendly teaching methods, since there is no reason for teaching methods to have changed in the classes of the non-voluntary teachers.[13]

A potential problem with interventions requiring students to make an extra effort in one subject is that this may result in a reduction in their efforts in other subjects and adversely affect their performance in these subjects (Ly et al., 2020). In the particular case of our intervention, the opposite is true: performance is positively affected in History-Geography (the subject that includes civic education), but also in the other subjects, so that the estimated impact of the treatment on average grades remains almost unchanged (about 10%) even when we drop the observations of both volunteer teachers and all History-Geography teachers (Table 3). This result is in line with the assumption that the program promotes attitudes and behaviors that can pay off in many different contexts. Appendix Table F2 further shows the details of the treatment effects, subject by subject. It

---

[13]It should also be noted that non-volunteer teachers were not specifically surveyed (the information obtained on their grades is from the usual administrative records) and that it is unlikely that they were influenced by our experimental observation scheme.

confirms that the improvement in grades is noticeable in History-Geography (20%), but also in French Language (17%), Arts (14%) or Sports (12%), subjects in which teachers are asked to value students' ability to interact constructively with others.

4.3 Treatment effect on social interactions

To further investigate the changes associated with the program, it is also possible to use the information collected on friendship networks from all sites of the experiment. For each student on the list of potential participants, these data allow us to calculate the number of friends they have among the other potential participants at their school.[14] Among the friends of each participant, it is also possible to distinguish those who share the same social and geographical origins and the same gender as him/her (socially similar friends) and those who do not share all these characteristics (socially different friends, i.e., with at least one difference). Among the latter, we can even distinguish those with one difference, two differences or three differences. Again, the analysis is conducted on the sample observed at baseline and for whom outcomes are measured at end-line. Appendix Table E2 shows that there are no significant differences in missing rates or in baseline characteristics between treated and control groups.

In this framework, the comparison of friendship networks at the end and the beginning of the year shows that the implementation of the program has no significant effect on the number of socially similar friends, but has a positive impact on the number of socially different friends (Table 4). The number of friends with at least one difference increases on average by about 0.22 (i.e., an 11% increase), and those with at least two differences by about 0.10 (i.e., a 15% increase).

When we further examine each separate component of friends' identity separately, we find that the intervention simultaneously increases the number of those who have a different gender, but also the number of those with a different social origin or with a different geographical origin. Leveraging the fact that the effects on the three components of heterophily are consistently

---

[14]For each student i, we count as "friends" those students whom i has placed on his or her friend list and who, in turn, have placed i on their own friend list.

positive, the corresponding synthetic index of network heterophily is significantly higher in treated than in control classes (+11% of a SD).[15] Similar effects are observed at all three sites of the experiment (see Appendix Table F3). As there is no decline in homophilic friendships, the total number of friendship ties tends to increase in treated classes, although the effect is only marginally significant at standard levels. Taken together, our findings are in line with the idea that project-based pedagogy is associated with intensified and more heterophilic interactions within classrooms.

This greater openness of friendship networks may be due to the fact that fighting prejudices and discrimination based on gender or origin is one of the themes most often chosen by program participants. It may also be due to the fact that, at the start of the year, participants have to work in small groups formed by their teachers, and this can help create new friendships. In France and Spain, we know the composition of these small groups, and we have been able to verify that they do indeed tend to be more diverse than the initial friendship networks observed at baseline, especially in terms of student gender. In these small groups, program participants interact on average with 44% of students of a gender different from their own, whereas their initial group of friends includes on average only around 21% of students of a different gender (see Appendix Table D5).

When we regress our endline measures of network heterophily on the interaction between the treatment dummy and variables characterizing small group composition, we find that the effect of the treatment on network heterophily tends to be stronger when participants are assigned to a small group with a higher share of socially different classmates, namely when they are forced into more heterophilic interactions within their small groups. In particular, the more participants interact with classmates who are socially different from them in their small groups, the more friends they have who are socially different from them at the end of the year and the fewer friends they have who are socially similar to them (Table 5). These results suggest that the effect of treatment on the heterophily of friendship networks de-

---

[15]The distribution of the synthetic index, separately by treatment status, is shown in Figure G1 in the online appendix.

rives at least in part from small-group work and from the efforts made by teachers to impose some gender and social diversity within small groups.[16]

4.4 Treatment effect on civic attitudes and democratic participation

By promoting group work and a less vertical pedagogy, the program clearly contributes to improving students' relationship with the school, teachers and classmates. As these traits are predictive of broader civic attitudes (Kupchik and Catlaw, 2015; Bacher-Hicks et al., 2024), we may also expect the program to have an impact on broader aspects of social and political engagement. In this section, we explore these issues as directly as possible using data from online surveys conducted before and after the intervention in all three sites of the experiment. More specifically, we provide evidence on the effect of being assigned to the treatment group on the two synthetic indicators of *Civic Attitudes* and *Democratic Participation* and their underlying dimensions. As discussed above, the *Civic Attitudes* index is based on students' social engagement, tolerance and support for equal rights (i.e., three secondary outcomes) while the *Democratic Participation* index is based on their political self-efficacy, interest in political life and participation in the Global Climate Strike (i.e., again three secondary outcomes). Compared to administrative data, these data have the disadvantage of being more prone to desirability bias, but they have the advantage of being measured in all three sites of the experiment.

For each of the two primary outcomes and the six secondary outcomes, the analysis is again conducted on the working sample of individuals who are observed at baseline and for whom the outcome is measured at end-line. Table E6 in the appendix shows that these working samples represent between 67% and 70% of the initial sample, but, again, there is no signifi-

---

[16]The recommendation made to teachers was to assign participating students to the small groups randomly. As it happens, there is no correlation between the gender (or social origin) of the participants and that of the other members of their small groups, in line with the random assignment assumption. However, baseline friends tend to be over-represented in small groups, and there is a correlation between participants' geographical origins and those of the other members of their small groups (see Appendix Table E5). The small groups are more diverse than initial friendship networks, but not as much as if assignment to small groups had been truly random.

cant difference in missing rates between the treated and the control group. These tables further show that there are no differences in baseline characteristics between the treatment and control groups for each of the thirteen working samples.

Table 6 reports the impact of being assigned to the treatment group for each of the eight outcomes. The results show a positive and statistically significant point estimate for both the *Civic Attitudes* and *Democratic participation* indexes. The intervention caused an increase in the *Civic Attitude* index of about 10.2% of a SD and an increase in the *Democratic Participation* index of about 8.4% of a SD.[17] Again, similar positive effects were observed at all three sites of the experiment (see online appendix Table F3).

The positive impact on the *Civic Attitudes* index is mostly driven by how the intervention affected their social engagement (and to a lesser extent, support for equal rights), something that is consistent with the fact that many of the projects chosen by students were related to helping others (elderly, minorities, other students). The positive impact on the *Democratic Participation* index is mostly driven by change in political self-efficacy and, to a lesser extent, to an increased participation in the Global Climate Strike: this may be connected with student empowerment during the whole project, and specifically the initial vote on potential projects, which can improve their ability and willingness to participle in democratic processes and debates.

The *Civic Attitudes* results could be an artifact of how projects were actually implemented in the ACT program. This index includes, among others, measures of engagement *at school* over the last school year, such as tutoring of younger students and participation in the school newspaper; and engagement *outside of school* over the last school year, such as volunteering in a humanitarian association or in an association aimed at helping the community or mentoring younger children (with homework, in sports, etc.). Given that some of the projects implemented during the ACT program can consist of tutoring other students or helping the community, the positive effect we find could be a direct measure of the implementa-

---

[17]The distributions of these synthetic indices, separately by treatment status, are shown in Figures G2 and G3 in the online appendix.

tion of the program, rather than of its effects. To test this hypothesis, we have been through the description of class projects provided in our teacher survey, and excluded all observations for which the project corresponds to one of the questions used for our social engagement measure. We also excluded the projects that could not be classified (non-response or ambiguous description). To the extent that such observations might overstate the impact measure, we can form a conservative estimate of the effect by running our regressions with these observations excluded. The main results are reported at the bottom of Table 6 (and detailed results in Table F4 in the online appendix). They show that the findings are robust: the effect on the *Civic Attitudes* index is now +7.1% of a SD, compared to +10.2% in the full sample, and the coefficient on the *Democratic Participation* index is now +7.9% of a SD instead of +8.4%. There is no evidence that our results are an artifact of how the program was implemented.

5 Heterogeneity analysis

The implementation of the ACT program coincides with an improvement in student behavior and academic performance. An important question, however, is whether this improvement has affected all students. In this section, we draw on the pre-analysis plan written before the start of the experiment and on Machine Learning techniques to shed light on this issue.

5.1 Pre-registered heterogeneity analysis

In the pre-analysis plan, we hypothesized that treatment effects might differ by gender and family background. Female students and high-SES students tend to have better behavior and better grades, (as confirmed by Table D3 in the online appendix) and we speculated that they might be more receptive to a school-based intervention such as ACT. It turns out that there is no significant variation in the impact of the treatment on school behavior, democratic participation or social interactions between female and male students or between students with different family backgrounds (see Tables F5 and F6 in the appendix). We do, however, detect

23

some heterogeneous effects on teachers' grades: the impact of the treatment on average grades is large and statistically significant for female students and for high SES students only, that is, for the strongest groups of students at baseline. To the extent that teacher evaluations are a measure of student engagement in school, this result is consistent with the idea that the intervention stimulates students' engagement all the more if they are committed to begin with. We also hypothesized that treatment effects might be stronger for students who had previous experience as student representatives, in line with a model in which civic skills acquired in the early school years (as a representative) and educational investments made later in school (this intervention) represent two types of inputs that are complements in the skill production function. Appendix Table F7 confirms that the impact of the treatment on school behavior, academic achievement, civic attitudes or democratic participation tends to be stronger for representatives than for the other students, although none of the impact differentials are statistically significant at standard levels.

5.2   Data-driven heterogeneity analysis

In the previous section, the comparison of treatment effects across pre-defined subgroups suggests that the program is most effective for the most civic-minded students at baseline. To take one step further, several Machine Learning techniques are now available to explore the heterogeneity of treatment effects in a data-driven manner, without pre-defined subgroup restrictions. We use one of these techniques, namely the generalized random forest (GRF) procedure introduced by Athey et al. (2019). This makes it possible to predict treatment effects for each student individually using all available information on his/her baseline characteristics (i.e., not simply information on the characteristics mentioned in the pre-analysis plan) and to test the existence of heterogeneity in these treatment effects.[18] Denoting Y the outcome under consideration, T the binary treatment and Z the set of baseline covariates, this procedure grows a causal forest to construct an

---

[18]To train our procedures, we use our baseline measures of students' civic skills (all scales and subscales), baseline measures of teachers' characteristics, friendship ties, the country of the school as well as dummies indicating experience as a student representative, gender, geographic origin and family socio-economic background.

estimate of the conditional average treatment effect $s_0(Z) = E(Y_1 - Y_0|Z)$, where Y_1 and Y_0 represent students' potential outcomes in treated and non-treated states. Following Athey and Wager (2019) and Chernozhukov et al. (2018), it is then possible to test for the existence of heterogeneity in $s_0(Z)$.

We conducted this test by considering our main outcome variables, namely, the dummy variable indicating participation in a citizenship project, the index of *School Behavior* and academic achievement constructed from administrative data, the index of *Civic Attitudes* and *Democratic Participation* constructed from online surveys and the variable capturing the degree to which friendship networks are heterophilic. The detailed results of the different tests are given in Table 7. They show that the null hypothesis of no heterogeneity in treatment effects is unambiguously rejected for five of the six outcome variables studied.

To further explore the sources of treatment effect heterogeneity, it is possible to identify the baseline variables that are most often used by the causal forest procedure to predict individual treatment effects. This analysis shows that the most important source of treatment heterogeneity comes from the variable indicating the extent to which the teacher in charge of the program is involved in the life of the school, as measured at baseline by the number of school councils in which s/he participates.[19] This involvement variable appears to be one of the main sources of treatment effect heterogeneity for each of the outcome variables for which the null hypothesis of no heterogeneity is rejected.

This finding is consistent with the long-standing literature that demonstrates teacher effect heterogeneity, including on student behavioral outcomes (see e.g. Jackson et al. (2014), Jackson (2018)).

To illustrate the importance of teacher involvement, Table 8 compares the effects of the treatment according to whether the teacher in charge of the intervention belongs to the most or least involved half of the teachers. The table confirms that the effect of the program on student behavior is

---

[19]This variable takes on values between 0 and 4. About 24% of the teachers do not participate in any school councils, 28% participate in 1 council, 21% participate in 2 councils, 17% in 3 councils and 9% in 4 councils. As shown in Appendix Table D6, the most involved teachers tend to be more experienced than those least involved. They are also (at baseline) closer to the pedagogical principles promoted by the program.

strong and significant for students taught by the most involved teachers, while it is much weaker and not statistically significant for those taught by the least involved teachers. When implemented by the most involved teachers, the intervention coincides with a marked reduction in absenteeism and disciplinary sanctions as well as with a significant diversification of friendship networks and a marked increase in students' participation in the Global Climate Strike. However, none of these effects are observed when the intervention is implemented by the least involved teachers.

The effectiveness of the teachers who are most involved in the life of their school is not simply a matter of their ability to get students to participate in citizenship projects, or even their ability to adapt their teaching practices. Indeed, the effect of the treatment on students' participation in citizenship projects is as strong for the least involved teachers as for the most involved ones and the effect of the treatment on teaching practices is even stronger for the least involved teachers than for the most involved ones.[20]

To take one step further, we also compared the projects chosen by the students of least involved and most involved teachers (Appendix Table D7). Chi-square tests do not reveal any significant differences in their distribution across topics (p-value = 0.50) or across target populations (p-value = 0.81). In the end, to have an effect on student outcomes, it does not appear to be sufficient to adopt a student-centered pedagogy or to help students democratically choose a specific type of citizenship project. For this type of intervention to be effective, it must also be implemented by teachers who are themselves civic-minded and accustomed to taking part in collective deliberations.

In addition to teacher involvement, baseline measures of civic outcomes (such as baseline social engagement) appear to be important sources of heterogeneity in treatment effects on civic attitudes and participation (see Table F8 in the appendix). The fact that the program had a significantly stronger impact on the civic outcomes of the students who were initially more socially engaged or more interested in politics is consistent with the findings obtained from the pre-analysis plan, which highlighted the impact of the program on students who have had experience as student repre-

---

[20]At baseline, the teaching practices implemented by the least involved teachers are less student-centered, but the intervention leads to a catch up.

sentatives. This set of results is in line with the idea that the ability to accumulate civic skills in adolescence depends on the civic skills one may have accumulated earlier, during childhood.

Finally, we can emphasize that the country in which the experiment takes place never appears to be a major source of heterogeneity in treatment effects. Appendix Table F3 confirms that there is no significant variation in treatment effects across the three sites of the experiment. This result is all the more striking given that the themes of the projects chosen by the students are significantly different from one site to another (Table 1).[21] Our findings are once again consistent with the assumption that there is no direct link between the type of project chosen and the success of the intervention - any theme can be suitable, provided it has been chosen democratically by local students.

6   Conclusion

Through a large-scale randomized controlled trial, this paper reveals the highly significant effects that implementing an active-learning method for civics education can have on students' attitudes and behaviors. The program involves training teachers to facilitate effective group work in their classrooms, so that small groups of students are able to independently discuss, elect and implement projects dealing with discrimination, social inclusion or cultural diversity. We first show that this training changed teaching practices across all three countries. We then show that, in all three settings, the program was followed by improvements in the quality of classroom social interactions and civic behavior. The program helps students make new friends (and friends less like themselves), enhance their social and political engagement, and leads them to better respect the internal rules of the school community (with fewer absences and serious sanctions). These results are in line with a long tradition dating back to Aristotle, Tocqueville or Dewey, which defends the idea that citizenship is learned primarily through practical investment in local, social and political life.

---

[21]When we compare the distribution of projects across topics for the three countries, chi-square tests reject the equality of the distribution.

The comparison of treatment and control groups also reveals that the program has stronger effects when it is implemented by teachers who are initially more involved in the life of the school. This finding complements the emerging literature on teacher effects on student attitudes, and suggests that active-learning methods require strong interpersonal skills and engagement, which are heterogeneously distributed among teachers. We also find that the effects of the intervention tend to be more important on students who are initially endowed with civic skills. This finding is in line with the idea that skills acquired in the early school years and school investments made later in adolescence are complementary inputs in the education production function. More research is needed to assess the effect of implementing new civics programs based on debates, group projects, and learning by doing much more systematically in the early grades.

Acknowledgments

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge, "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, 2023, *138* (1), 1–35.

Alan, Sule, Ceren Baysan, Mert Gumren, and Elif Kubilay, "Building Social Cohesion in Ethnically Mixed Schools: An Intervention on Perspective Taking," *The Quarterly Journal of Economics*, 2021, *136* (4), 2147–2194.

Anderson, Michael L, "Multiple Inference and Gender Differences in the Effects of Early Intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495.

Arrow, K, "The Economic Consequences of Learning by Doing," *Review of Economic Studies*, 1962, *29* (3), 155–173.

Athey, Susan and Stefan Wager, "Estimating Treatment Effects with Causal Forests: An Application," *Observational Studies*, 2019, *5* (2), 37–51.

_ , Julie Tibshirani, and Stefan Wager, "Generalized Random Forests," *The Annals of Statistics*, 2019, *47* (2), 1148–1178.

Bacher-Hicks, Andrew, Stephen B. Billings, and David J. Deming, "The School to Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime," *American Economic Journal: Economic Policy*, 2024.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal statistical society: series B (Methodological)*, 1995, *57* (1), 289–300.

Bozec, Géraldine, *Éducation à la citoyenneté: politiques, pratiques scolaires et effets sur les élèves*, CNESCO (Conseil national d'évaluation du système scolaire), 2016.

Burtt, Shelley, "The Politics of Virtue Today: A Critique and a Proposal," *American Political Science Review*, 1993, *87* (2), 360–368.

Campbell, David E, "What social scientists have learned about civic education: A review of the literature," *Peabody Journal of Education*, 2019, *94* (1), 32–47.

_ , **Meira Levinson, and Frederick M Hess**, *Making civics count: Citizenship education for a new generation*, Harvard Education Press, 2012.

Cappelen, Alexander, John List, Anya Samek, and Bertil Tungodden, "The Effect of Early-Childhood Education on Social Preferences," *Journal of Political Economy*, 2020, *128* (7), 2739–2758.

Carothers, Thomas and Andrew O'Donohue, *Democracies divided: The global challenge of political polarization*, Brookings Institution Press, 2019.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val, "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India," Technical Report, National Bureau of Economic Research 2018.

Cohodes, Sarah and James J Feigenbaum, "Why Does Education Increase Voting? Evidence from Boston's Charter Schools," Technical Report, National Bureau of Economic Research 2023.

Crawford, Barbara A, "Learning to teach science as inquiry in the rough and tumble of practice," *Journal of research in science teaching*, 2007, *44* (4), 613–642.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

Dalton, Russell J, "Citizenship Norms and the Expansion of Political Participation," *Political studies*, 2008, *56* (1), 76–98.

_ , *The Participation Gap: Social Status and Political Inequality*, Oxford University Press, 2017.

Dee, Thomas S, "Are there civic returns to education?," *Journal of Public Economics*, 2004, *88* (9-10), 1697–1720.

Dewey, John, *My Pedagogic Creed*, EL Kellogg & Company, 1897.

_ , *The School and Society*, University of Chicago Press, 1915.

Donbavand, Steven and Bryony Hoskins, "Citizenship education for political engagement: A systematic review of controlled trials," *Social Sciences*, 2021, *10* (5), 151.

Gill, Brian, Emilyn Ruble Whitesell, Sean P Corcoran, Charles Tilley, Mariel Finucane, and Liz Potamites, "Can charter schools boost civic participation? The impact of democracy prep public schools on voting behavior," *American Political Science Review*, 2020, *114* (4), 1386–1392.

Green, Donald P, Peter M Aronow, Daniel E Bergan, Pamela Greene, Celia Paris, and Beth I Weinberger, "Does knowledge of constitutional principles increase support for civil liberties? Results from a randomized field experiment," *The Journal of Politics*, 2011, *73* (2), 463–476.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Civic capital as the missing link," *Handbook of social economics*, 2011, *1*, 417–480.

Heater, Derek, *Citizenship: The civic ideal in world history, politics and education*, Manchester University Press, 2004.

Isac, Maria Magdalena, Ralf Maslowski, Bert Creemers, and Greetje Van der Werf, "The contribution of schooling to secondary-school students' citizenship outcomes across countries," *School Effectiveness and School Improvement*, 2014, *25* (1), 29–63.

Jackson, C. Kirabo, "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes," *Journal of Political Economy*, Oct 2018, *126* (5), 2072–2107.

Jackson, C Kirabo, Jonah E Rockoff, and Douglas O Staiger, "Teacher effects and teacher-related policies," *Annu. Rev. Econ.*, 2014, *6* (1), 801–825.

Kraft, Matthew A, "Teacher effects on complex cognitive skills and social-emotional competencies," *Journal of Human Resources*, 2019, *54* (1), 1–36.

Kupchik, Aaron and Thomas J. Catlaw, "Discipline and Participation: The Long-Term Effects of Suspension and School Security on the Political and Civic Engagement of Youth," *Youth & Society*, Jan 2015, *47* (1), 95–124.

Larreguy, Horacio and John Marshall, "The Effect of Education on Civic and Political Engagement in Nonconsolidated Democracies: Evidence from Nigeria," *Review of Economics and Statistics*, 2017, *99* (3), 387–401.

Lazonder, Ard W and Ruth Harmsen, "Meta-analysis of inquiry-based learning: Effects of guidance," *Review of educational research*, 2016, *86* (3), 681–718.

Ly, Son Thierry, Eric Maurin, and Arnaud Riegert, "A Pleasure that Hurts: The Ambiguous Effects of Elite Tutoring on Underprivileged High School Students," *Journal of Labor Economics*, 2020, *38* (2), 501–533.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos, "Does education improve citizenship? Evidence from the United States and the United Kingdom," *Journal of Public Economics*, 2004, *88* (9-10), 1667–1695.

Niemi, Richard G., Stephen C. Craig, and Franco Mattei, "Measuring Internal Political Efficacy in the 1988 National Election Study," *American Political Science Review*, 1991, *85* (4), 1407–1413.

Petek, Nathan and Nolan G. Pope, "The Multidimensional Impact of Teachers on Students," *Journal of Political Economy*, Aug 2022.

Putnam, Robert, "The prosperous community: Social capital and public life," *The american prospect*, 1993, *13* (Spring, Vol. 4.).

Rushton, J Philippe, Roland D Chrisjohn, and G Cynthia Fekken, "The altruistic personality and the self-report altruism scale," *Personality and individual differences*, 1981, *2* (4), 293–302.

Schulz, Wolfram, John Ainley, Julian Fraillon, Bruno Losito, Gabriella Agrusti, and Tim Friedman, *Becoming citizens in a changing world: IEA International Civic and Citizenship Education Study 2016 international report*, Springer Nature, 2018.

Syvertsen, Amy K, Michael D Stout, Constance A Flanagan, Dana L Mitra, Mary Beth Oliver, and S Shyam Sundar, "Using elections as teachable moments: A randomized evaluation of the student voices civic education program," *American Journal of Education*, 2009, *116* (1), 33–67.

Tabellini, Guido, "Institutions and culture," *Journal of the European Economic Association*, 2008, *6* (2-3), 255–294.

Valdebenito, Sara, Manuel Eisner, David P Farrington, Maria M Ttofi, and Alex Sutherland, "School-based interventions for reducing disciplinary school exclusion: A systematic review," *Campbell Systematic Reviews*, 2018, *14* (1), i–216.

Williams, Melissa S., "Citizenship as Identity, Citizenship as Shared Fate, and the Functions of Multicultural Education," in Kevin McDonough and Walter Feinberg, eds., *Citizenship and Education in Liberal-Democratic States: Teaching for Cosmopolitan Values and Collective Identities*, New York: Oxford University Press, 2003, p. 208–47.

Main Figures

Figure 1: School Behaviour and Average Grade by Treatment Status
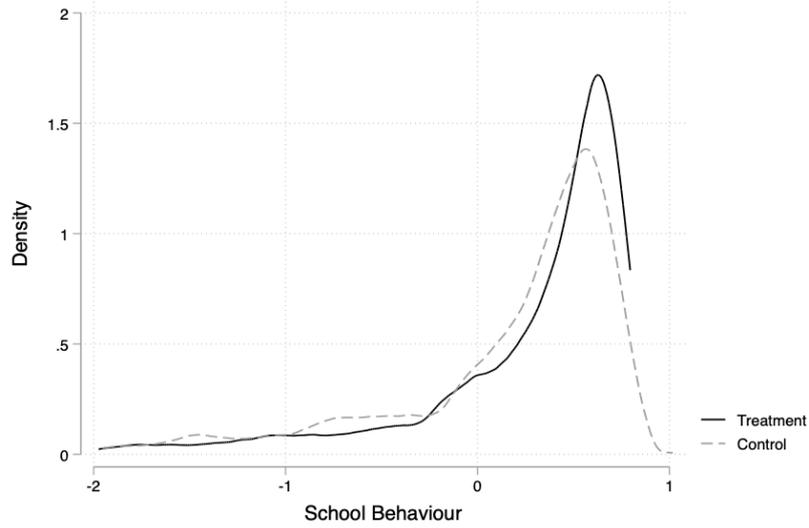
(a)
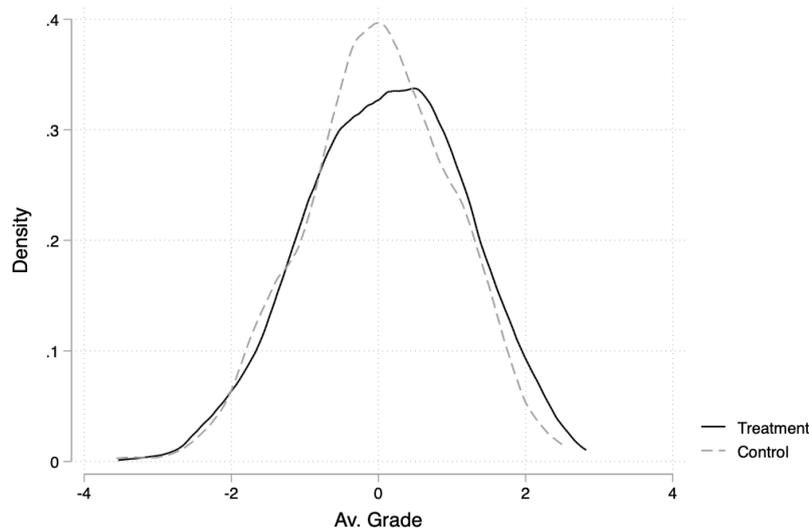
(b)

Note: Figures 1 (a) and 1 (b) respectively show the kernel distribution of the standardized student school behaviour index and end-of-the-school-year grades, by treatment status.

Main Tables

Table 1: Citizenship Projects

	(1) All	(2) France	(3) Greece	(4) Spain
Project topics				
Discrimination	0.64	0.56	0.71	0.69
	(0.48)	(0.50)	(0.46)	(0.47)
Social inclusion	0.53	0.50	0.71	0.47
	(0.50)	(0.51)	(0.46)	(0.50)
Cultural diversity	0.29	0.18	0.29	0.41
	(0.46)	(0.39)	(0.46)	(0.50)
Targeted population				
Elderly	0.15	0.12	0.12	0.18
	(0.35)	(0.33)	(0.34)	(0.39)
Homeless	0.12	0.14	0.17	0.08
	(0.33)	(0.35)	(0.38)	(0.28)
Migrants	0.26	0.14	0.42	0.31
	(0.44)	(0.35)	(0.50)	(0.47)
Women	0.19	0.14	0.08	0.29
	(0.39)	(0.35)	(0.28)	(0.46)
LGBT	0.11	0.10	0.04	0.16
	(0.32)	(0.30)	(0.20)	(0.37)
Disabled	0.26	0.34	0.29	0.16
	(0.44)	(0.48)	(0.46)	(0.37)
Other	0.25	0.30	0.17	0.24
	(0.44)	(0.46)	(0.38)	(0.43)
No specific group	0.20	0.28	0.12	0.16
	(0.40)	(0.45)	(0.34)	(0.37)
General orientation of the project				
School oriented project	0.56	0.46	0.52	0.67
	(0.50)	(0.50)	(0.51)	(0.48)
Out-of-school oriented project	0.44	0.54	0.33	0.38
	(0.50)	(0.50)	(0.48)	(0.49)
Observations	125	51	25	49

Note: This table shows the percentage of citizenship projects implemented in the treatment group that relate to each of the three topics covered by the ACT intervention, the population targeted by these projects and the share of in-school and out-of-school oriented projects. One project may correspond to multiple topics and/or targeted population. Standard deviations are in parentheses.

Table 2: Treatment Effects on Program Implementation

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
Teacher-level outcomes					
Participation in ACT training	0.024	0.949	0.026	0.000	247
Actual implementation of ACT project	0.040	0.902	0.029	0.000	245
Teacher Pedagogy	0.000	0.428	0.125	0.001	254
Student-level outcomes					
Student participation in a citizenship project	0.301	0.421	0.024	0.000	4133

Note: For each of the four row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school level whereas the fourth column shows the corresponding p-value. The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression. Attrition analysis and balance checks for the 3 first analysis samples are provided in Appendix Table E1. For the last analysis sample, they are provided in the first column of appendix Table E2.

Table 3: Treatment Effects on School Behavior and Academic Achievement

	(1) C	(2) T-C	(3) S.E.	(4) Unadj. p-val	(5) Adj. p-val	(6) N
School Behaviour	0.000	0.266	0.089	0.003	-	2251
Absence	0.000	0.249	0.104	0.017	0.034	2227
Punctuality	0.000	0.185	0.099	0.063	0.084	2227
Exclusion	0.000	0.190	0.062	0.002	0.009	2186
Smaller disciplinary sanctions	0.000	0.042	0.092	0.648	0.648	2241
Average Grade	0.000	0.126	0.040	0.002	-	2251
Av. Grade in subjects taught by volunteer teachers	0.000	0.267	0.063	0.000	-	2155
Av. Grade in subjects not taught by volunteer teachers	0.000	0.101	0.041	0.014	-	2251
Av. Grade in subjects not taught by volunteer teachers or HG teachers	0.000	0.098	0.042	0.019	-	2251

Note: Columns C, T-C, S.E. and N have the same meaning as in Table 2. Column "Adj. p-val" shows shows the p-value adjusted for false discovery rate (Benjamini and Hochberg, 1995) while column "Unadj. p-Val" shows the unadjusted p-value. Each line corresponds to a separate regression. Attrition analysis and balance checks for each of the eight analysis samples are provided in appendix Tables E3 and E4.

Table 4: Treatment Effects on Social Interactions

	(1) C	(2) T-C	(3) S.E.	(4) Unadj. p-val	(5) Adj. p-val	(6) N
Number of friends	3.650	0.195	0.130	0.133	-	4299
Nb of friends with 0 difference	1.552	-0.011	0.066	0.866	-	4298
Nb of friends with 1 difference or more	2.098	0.223	0.089	0.012	-	4299
Nb of friends with 2 differences or more	0.620	0.097	0.050	0.053	-	4299
Nb of friends with 3 differences	0.069	0.009	0.012	0.467	-	4299
Friendship Heterophily	0.000	0.107	0.041	0.010	-	4299
Nb of friends of different gender	0.892	0.124	0.056	0.028	0.075	4299
Nb of friends of different geo. origin	0.543	0.071	0.042	0.088	0.088	4299
Nb of friends of different social origin	1.353	0.119	0.061	0.050	0.075	4299

Note: Columns C, T-C, S.E., Adj. p-val, Unadj. p-val and N have the same meaning as in Table 3. Each line corresponds to a separate regression. Attrition analysis and balance checks for the sample are provided in the second column of appendix Table E2.

Table 5: Student Working Groups' Composition: Effects on Social Interactions

	(1) C	(2) $\hat{\beta}_1$ (T)	(3) $\hat{\beta}_1$ s.e.	(4) $\hat{\beta}_1$ p-val	(5) $\hat{\beta}_2$ (TxShare)	(6) $\hat{\beta}_2$ s.e.	(7) $\hat{\beta}_2$ p-val	(8) N
Friendship Heterophily	0.059	0.127	0.075	0.091	0.231	0.115	0.045	2693
Number of friends	3.747	-0.002	0.197	0.991	0.249	0.235	0.289	2693
Nb of friends with 0 difference	1.525	-0.311	0.107	0.004	0.021	0.144	0.886	2693
Nb of friends with 1 difference or more	2.222	0.318	0.141	0.024	0.360	0.227	0.112	2693
Nb of friends with 2 differences or more	0.667	0.138	0.086	0.107	0.352	0.169	0.037	2693
Nb of friends with 3 differences or more	0.073	0.027	0.029	0.364	0.010	0.055	0.860	2693

Note: This table shows the results of regressing student friendship outcomes on a treatment dummy and on this treatment dummy interacted with the share of students with at least two differences in the working group. Column (1) displays the mean of the row variable in the control group; columns (2) to (4) respectively show the point estimate, standard error and p-value associated to the treatment dummy; columns (5) to (7) respectively show the point estimate, standard error and p-value associated to the treatment dummy interacted with the share of students with at least two differences in the working group. The last column displays the size of the analysis sample, namely students who are observed at endline and for whom working group composition is known. Each line corresponds to a separate regression. All regressions control for strata fixed-effects. Standard errors are clustered at the school level.

Table 6: Treatment Effects on Civic Attitudes and Democratic Participation

	(1) C	(2) T-C	(3) S.E.	(4) Unadj. p-val	(5) Adj. p-val	(6) N
Civic Attitudes	0.000	0.102	0.033	0.002	-	4244
Social engagement	0.000	0.090	0.041	0.028	0.084	4244
Tolerance	0.000	0.027	0.031	0.375	0.375	4119
Equal rights	0.000	0.050	0.033	0.126	0.189	4110
Democratic Participation	0.000	0.084	0.033	0.011	-	4294
Political self efficacy	0.000	0.092	0.029	0.002	0.005	4241
Interest in political life	0.000	0.003	0.032	0.928	0.928	4294
Participation in Climate strike	0.000	0.068	0.039	0.081	0.121	4244
Civic Attitudes - projects unrelated to Social engagement	0.000	0.071	0.038	0.062	-	3469
Democratic Participation - projects unrelated to Social engagement	0.000	0.079	0.036	0.028	-	3509

Note: Columns C, T-C, S.E., Adj. p-val, Unadj. p-val and N have the same meaning as in Table 3. Each line corresponds to a separate regression. Attrition analysis and balance checks for the sample are provided in the second column of appendix Table E6. In the last two rows of the table, projects directly related to our social engagement measure are excluded from the analysis. More detailed results on this restricted sample are presented in appendix Table F4.

Table 7: Generalized Random Forests: Tests for Heterogeneity

	(1) Citizen. project	(2) School Behaviour	(3) Av. Grade	(4) Civic Attitudes	(5) Democratic Part.	(6) Friend. Heterophily
Heterogeneity	3.04 (0.32)	2.03 (0.43)	2.49 (0.35)	0.56 (0.45)	1.04 (0.35)	1.72 (0.39)

Most important variables:

1	Teacher resp	Teacher resp	Int. in pol	Soc. Engag.	Teacher resp	Teacher resp
2	Soc. Engag.	Sport Grade	Sport Grade	Teacher resp	Pol. Self-eff	Friends diff geo
3	Pol. Self-eff	Pol. Self-eff	Female teacher	Equal rights	Int. in pol	Friends one diff
4	Int. in pol	Soc. Engag.	Soc. Engag.	Pol. Self-eff	Soc. Engag.	Nb friends

Note: For each outcome variable, the first row of the table shows the estimated coefficient of the regression of the conditional average treatment effect (CATE) on its causal forest estimate, following Chernozhukov et al. (2018). Rejecting the assumption that this coefficient is zero is tantamount to rejecting that the variance of the CATE is zero, i.e. rejecting that there is no heterogeneity in the effects of the treatment. The next rows show the four most important variables determining the heterogeneity of treatment effects (i.e., those that are most often used by the causal forest procedure to predict individual treatment effects), by order of importance.

Table 8: Treatment Effects by Teacher Baseline School Involvement

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
High involvement					
Participation in a citizenship project	0.280	0.441	0.041	0.000	1999
School Behaviour	0.000	0.395	0.125	0.002	1045
Av. Grade	0.000	0.128	0.048	0.007	1045
Civic Attitudes	0.000	0.136	0.048	0.004	2065
Democratic Participation	0.000	0.142	0.048	0.003	2087
Participation in Climate strike	0.000	0.166	0.061	0.007	2063
Friendship Heterophily	0.000	0.215	0.053	0.000	2088
Teacher Pedagogy	0.000	0.392	0.241	0.104	136
Low involvement					
Participation in a citizenship project	0.319	0.443	0.050	0.000	2025
School Behaviour	0.000	0.030	0.132	0.817	1206
Av. Grade	0.000	0.181	0.060	0.002	1206
Civic Attitudes	0.000	0.046	0.069	0.505	2070
Democratic Participation	0.000	0.002	0.074	0.975	2098
Participation in Climate strike	0.000	-0.055	0.098	0.577	2072
Friendship Heterophily	0.000	-0.003	0.105	0.978	2102
Teacher Pedagogy	0.000	0.579	0.274	0.035	118

Note: Columns C, T-C, S.E., p-val, and N have the same meaning as in Table 2. Each line corresponds to a separate regression.

For online publication

Appendix A   Civic Education in Participating Countries

Instructional time

In France, citizenship education is taught as a separate subject for the whole 12 years of primary and secondary education, from age 6 to 18. In primary education, the average instructional time devoted to this subject is 30 hours per year. It is 28 hours in lower secondary education and 16 hours in upper secondary education. In primary and secondary education, citizenship education is also integrated in the curriculum of other subjects (history, geography, philosophy).

In Greece, citizenship education is taught as a separate subject for 4 years in primary and secondary education (at age 10-11, 13-14, 15-17). The instructional time devoted to this subject is on average 8 hours per year in primary education, 15 hours per year in lower secondary education and 15 hours in upper secondary education. In primary education, citizenship education is integrated in the curriculum of the other subjects.

In Spain, citizenship education is taught as a separate subject for 4 years in primary and secondary education (at age 10-12, 14-15, 16-17). The instructional time devoted to this subject is on average 8 hours per year in primary education, 17 hours per year in lower secondary education and 35 hours in upper secondary education. In primary and secondary education, citizenship education is also integrated in the curriculum of various other subjects.

Skills to be acquired and assessment

In France and Spain (as in many other countries), four skills are defined as essential by the national curriculum for students to become active and responsible citizens:

- Civic-related skills (participating in society through, for example, volunteering, and influencing public policy through voting and petitioning);

- Social skills (living and working with others, resolving conflicts);

- Communication skills (listening, understanding and engaging in discussion);

- Intercultural skills (establishing intercultural dialogue and appreciating cultural differences).

In Greece, however, civic-related skills as defined in this way are not included in the national curriculum.

In all three countries, educational authorities provide tools to help teachers assess the civic knowledge, skills and attitudes acquired by students through a range of subjects or through other school experiences. In France, students' social and civic competences are specifically evaluated by teachers at various key points in compulsory education (2nd and 5th years of primary education and last year of lower secondary education), using a standardized personal booklet. In all three countries, students' marks in citizenship education (taught as a separate compulsory subject) are generally taken into account to decide transition to the next level of education. For example, in France, the final written exam for lower secondary education addresses French language, mathematics, history-geography and civic education.

Class councils

All three countries in our experiment have established official regulations for the creation of councils at the class level. Class councils are formal bodies set up to deal with class-level matters. They usually meet several times a year, for example at the end of each term of instruction. Their composition varies depending on official regulations and/or school decisions, but they generally include representatives of teachers, students and parents. Students' representatives are elected by the students in the class. Their most common role is consultative. They help circulate information

between teachers and students and bring student problems to the attention of teachers.

Student councils and school governance

The student council's mandate relates mainly to formulating rules governing every-day school activities. The acquisition of educational materials, such as textbooks and software, and the supervision of budgetary matters are also activities which fall within the remit of student councils. However, student councils do not enjoy real decision-making power in any of the activities in which they are involved. Their role is advisory and is to ensure that students' views are heard. In France and Greece, members of the student councils are directly elected by all students in the school. In Spain, student councils are composed of both class representatives and members of school governing bodies who are directly elected.

In all three countries, students also participate in school governing bodies. In France and Spain, student representatives appointed to school governing bodies are directly elected by all the students of the school. In Greece, they are nominated by the student council. As representatives on school governing bodies, students are involved in decisions concerning the development of the school educational plan, the establishment of the rules governing school life, the choice and organization of extra-curricular activities and the supervision of budgetary matters. Student representatives play a mostly consultative role.

Appendix B   Experimental Sample

The experiment took place in France, Greece and Spain during the 2018-2019 school year. The program was defined jointly by the educational authorities of the three countries as well as the English ones. England was initially part of the experiment, but we had to exclude it from the evaluation because of problems recruiting enough schools and resulting implementation issues.[22] In this appendix, we provide additional information on how the schools in our experimental sample were selected and how representative they are.

As mentioned in the main text, school recruitment was limited geographically for practical reasons and to keep costs down. France and Spain targeted a subset of educational regions, scattered over the national territory, whereas in Greece, recruitment was limited to the Attica region. Specifically, we have schools from 6 different French educational regions (Aix-Marseille, Amiens, Nancy-Metz, Nantes, Orleans-Tours and Versailles) and 13 Spanish regions (Andalucıa, Aragon, Asturias, Cantabria, Castilla-La Mancha, Castilla y Leon, Comunidad Valenciana, Extremadura, Galicia, Islas Baleares, Madrid, Murcia, Ceuta and Melilla). There was no other predefined eligibility criterion, except that all schools were public schools.

To enroll in the program, schools had to provide the name of (at least) one volunteer teacher, as well as a list of the students who would participate in the program, should the school be assigned to the treatment group. In the vast majority of volunteer schools, only one teacher volunteered to participate and enrolled a class he or she taught. However, no constraints were imposed on the number of participating teachers or classes, so in a few cases, several teachers and/or classes enrolled in the experiment. Tables B1 and B2 provide descriptive statistics about students and teachers enrolled in the experiment.

---

[22]By September 2018 only 8 schools had been recruited and a new time table had to be agreed. This ultimately led to 42 schools recruited on a revised protocol, with class projects starting very late in the year. Thus, the statistical power and scope of the intervention are much lower in England and not easily comparable to other countries. There were also problems with attrition of schools after recruitment. See the European Commission report (Briole et al., 2020) for a full account of the evaluation process in English schools.

Although this sample was not designed to be representative of middle school students in each country, the student characteristics in our sample are not very different from population averages. In particular, the proportions of students with few (or many) books at home are quite similar in our samples and in the PISA survey. For instance, we observe about 21% of students with no more than 10 books at home in the French sample, 8% in the Greek sample and 11% in the Spanish sample, while the PISA survey reports 21%, 10% and 10% respectively. There is a long standing literature that has consistently found that the book-at-home indicator provides one of the best proxy for student socioeconomic status and subsequent academic achievement (see e.g. Eriksson et al. (2021)). This variable is measured in exactly the same way (and with very few missing values) in our three national samples and in the PISA 2018 student survey.[23]

We also observe about 13.2% of students who have already repeated a grade in the French, 3.5% in the Greek and 23.5% in the Spanish samples, while the OECD PISA 2018 survey reports repetition rates of 16.5%, 4% and 28.5% respectively. This result is in line with the idea that the baseline academic level of our student samples is representative of that of the general population.

In the French sample, we also observe eligibility for the financial aid received by students with low-income parents: the proportions are about 24% in our sample, compared to about 25.5% for all French students in middle school.[24]

Another feature to judge the external validity of this experiment is that, although this is a small sample of countries, it is striking that our results are very homogeneous across the three countries, in spite of differences in their education systems and civic education traditions. In Greece, the historical pre-eminence of Orthodox Church is enshrined in the constitution. The law organizing the education system states that one of the aims of education is to help pupils have belief in the authentic elements of the Christian Orthodox tradition. In Spain, Catholicism was also for a long time the state religion, but it has not been since the end of the dictatorship

---

[23]See https://www.oecd.org/education/pisa-2018-assessment-and-analytical-framework-b25efab8-en.htm

[24]See: https://www.education.gouv.fr/reperes-et-references-statistiques-sur-les-enseignements-la-formation-et-la-recherche-2019-3806

and the approval of the Spanish constitution in 1978. Religion classes are still taught in schools by teachers appointed by the bishops and paid by the state. In France, Catholicism is no longer a state religion since 1905, but secular civic education must deal with a much larger Muslim minority than in Spain or Greece.

Table B1: Student characteristics

	(1) All	(2) France	(3) Greece	(4) Spain
Female	0.51	0.52	0.50	0.51
	(0.50)	(0.50)	(0.50)	(0.50)
Age (September 2018)	14.06	13.63	14.04	14.51
	(0.72)	(0.68)	(0.35)	(0.67)
European origin only	0.81	0.76	0.90	0.79
	(0.40)	(0.43)	(0.30)	(0.41)
Experience as student representative	0.35	0.30	0.53	0.30
	(0.48)	(0.46)	(0.50)	(0.46)
Nb of books at home	1.97	1.77	2.16	2.06
	(1.27)	(1.32)	(1.20)	(1.23)
N	4,299	1,649	932	1,718

Note: This table shows the average characteristics of students enrolled in the experiment, namely their gender, age, a dummy indicating that all parents and grand-parents were born in Europe, a dummy indicating experience as a student representative and the number of books at home. These statistics are displayed for our main sample (Column 1) and separately by country (Columns 2 to 4). Standard deviations are in parentheses.

Table B2: Teacher characteristics

	(1) All	(2) France	(3) Greece	(4) Spain
Teaching experience (years)	18.02 (7.63)	15.34 (6.65)	21.13 (5.94)	19.48 (8.24)
Seniority in the school (years)	8.10 (6.63)	8.25 (6.06)	9.64 (6.87)	7.42 (7.02)
Citizenship teaching experience (years)	9.50 (9.28)	9.84 (8.94)	4.33 (4.95)	11.02 (10.15)
Female teacher	0.67 (0.47)	0.72 (0.45)	0.85 (0.37)	0.56 (0.50)
Age	45.06 (7.91)	40.46 (6.80)	49.92 (6.19)	47.74 (7.18)
Subjects taught				
National language	0.10 (0.30)	0.08 (0.27)	0.33 (0.48)	0.04 (0.19)
History-Geography	0.35 (0.48)	0.61 (0.49)	0.15 (0.37)	0.17 (0.38)
Foreign or ancient language	0.12 (0.33)	0.12 (0.33)	0.23 (0.43)	0.08 (0.28)
Social sciences	0.14 (0.35)	0.00 (0.00)	0.85 (0.37)	0.03 (0.16)
Philosophy, citizenship, religion	0.39 (0.49)	0.07 (0.25)	0.13 (0.34)	0.79 (0.41)
Other (science, math, art, sport, technology)	0.18 (0.39)	0.24 (0.43)	0.00 (0.00)	0.19 (0.39)
N	254	105	39	110

Note: This table shows the average characteristics of teachers in our sample, for the pooled sample of countries participating in the experiment (column (1)) and separately by country (columns (2) to (4)). Standard deviations are in parentheses.

Appendix C Administrative Data

In this appendix we provide additional information on how administrative data on absenteeism, disciplinary problems and teacher grades are collected in France. We also provide information on the official curriculum of subjects taught in French middle schools, particularly those that foster the development of skills close to those promoted by the ACT program.

Data on absences and disciplinary problems

The measurement of absenteeism and disciplinary problems by the school administration corresponds to a legal obligation. In particular, French law is very specific about the legal responsibilities of schools and how they should record and handle truancy. At the beginning of each class, teachers must inform the school principal immediately of any unauthorized absence and the principal must contact the parents as soon as possible to identify the cause of the absence. In case an accident happens to an absent child, the school remains responsible until parents are informed of the absence. In such a context, it is not likely that recorded truancy could be affected by teachers' subjective perceptions or by the empathy that they may have for some parents or children. Similarly, the exclusion of students (temporary or permanent) can only be decided after a fairly formal procedure. It involves the meeting of a disciplinary council led by the school principal and composed of elected representatives of teachers, parents and students, during which the student threatened with exclusion has the opportunity to explain his or her behavior.

Table C1 provides descriptive statistics about the measures of absenteeism, late arrivals and sanctions, as observed in our French sample. It shows that about 25% of students in our control group experienced at least one exclusion from school during the school year, in line with the fact that incivility and violence represent a significant problem in many public schools (Fréchou, 2023) . Table C1 also shows that there is an average of about 8 unjustified absences per year per student, in line with national trends. In our sample, about 3.9% of students are absent at least 4 half-days per month, a proportion of high-absenteeism students very similar

to the 3.5% that the Ministry of Education reports on average for French middle schools (Cristofoli, 2020).

Table C1: School behavior: Descriptive statistics

Variable	Mean	Std. Dev.	N
Unjustified absences (half-days)	7.893	19.061	944
At least 1 unj. absence	0.693	0.462	944
At least 4 unj. absences per month	0.039	0.194	944
Nb of late arrivals	4.088	7.332	999
Nb of exclusions	0.677	2.202	958
At least 1 exclusion	0.252	0.434	958
At least 3 exclusions	0.078	0.269	958
Nb of other sanctions	1.417	3.205	1013
At least one other sanction	0.479	0.5	1013
At least 3 other sanctions	0.174	0.379	1013

Note: This table shows descriptive statistics on the total number of unjustified absences, late arrivals, exclusions and other disciplinary sanctions over the school year for students in French administrative sample. These statistics are computed using the control group only.

Data on teacher grades

For each of the schools in the French site of the experiment, we had access to the administrative register that includes the grades received by students at the end of each of the three terms of the school year. For each of the 11 subjects that students are required to take in middle school, we observe the grades received in the first quarter (before the implementation of the program) and at the end of the last quarter (after the implementation), so that it is possible to test the impact of the program on end-of-year grades holding initial grades constant.

To the extent that the program may affect civic skills and change students' attitudes, it can be expected to have an effect on their academic effort, the quality of the relationships with teachers and, ultimately, on teachers' evaluations. This is particularly the case for History-Geography, since the curriculum of this discipline includes civic education and since the majority of the teachers who volunteer for the program are History-

Geography teachers. Specifically, two-thirds of the teachers involved in the program are History-Geography teachers.[25] In this subject, the improvement in student performance could be the consequence of a pedagogy that is considered more enjoyable by the students, but also of a teaching content that is considered more interesting. To the extent that the intervention is able to elicit deeper changes in students' attitudes and behaviors, the effect should also be detected in other subjects, even if their teaching methods are unchanged and their teachers hardly ever participate in the program. This is, for example, the case in Sports.

In the French system, Sports is as much focused on the quality of social interactions and respect, as it is on athletic performance. The curriculum for Sports states: "Sports education develops access to a rich field of practices, with strong cultural and social implications, important in the development of the personal and collective life of the individual. Throughout schooling, Sports education aims to form a lucid, autonomous, physically and socially educated citizen, with a view to living together. It leads children and adolescents to seek well-being and to care about their health. It ensures the inclusion in the class of students with special educational needs or with disabilities. Sports education initiates to the pleasure of sports practice."

In the end, the social skills required to succeed in Sports are not far removed from some of those that the ACT program seeks to promote and, as a result, Sports scores may provide an indicator of program effectiveness.

The program can also be expected to have an impact in subjects where students are assessed on their ability to express personal views and listen to those of others, such as in French language or Arts. In French language, for instance, the curriculum states explicitly that students are assessed in part on their ability to "participate in a debate constructively and with respect for the other's speech". Also, one of the major themes that students must address with their French language teachers through novels, poetry and plays is called : "Living in society and participating in society".

In each school, we also know the subjects taught by teachers participating in the experiment, so that it is possible to test whether the impact of

---

[25]In contrast, only about 8% are French language teachers and less than 2% for Sports or Arts.

the program on grades is similar in the subjects of teachers participating in the experiment compared to the subjects of teachers outside the experiment. If the impact on grades is only perceptible for volunteer teachers, it cannot be completely ruled out that it reflects an effort on their part to make the intervention appear successful (a "social desirability bias"). If, on the other hand, the impact on grades is perceptible even for teachers who are not involved in the experiment, it can be interpreted as reflecting a deeper change in the students themselves, namely the acquisition of behavioral skills whose effects are felt beyond the context in which they are taught.

Appendix D   Sample characteristics and citizenship projects implemented

Table D1: Number of Schools, Teachers and Students, by Treatment Status and Country

	(1) All	(2) France	(3) Greece	(4) Spain
Number of schools				
Total	217	75	47	95
Treated schools	108	37	23	48
Control schools	109	38	24	47
Number of volunteer teachers on initial lists				
Total	323	123	67	133
Teachers in Treated school	161	60	35	66
Teachers in Control schools	162	63	32	67
Number of students on initial lists				
Total	6211	2269	1808	2134
Students in Treated school	3194	1202	884	1108
Students in Control schools	3017	1067	924	1026

Note: This table shows the number of schools, students and teachers in the sample of the experiment, by country and treatment status.

Table D2: Outcomes at baseline (standardized), by student experience as representatives

	(1)	(2)	(3)
	Represent.	*Non represent.*	Diff.
Civic Attitudes	0.118	-0.078	0.196***
Social Engagement	0.245	-0.188	0.433***
Tolerance	0.031	-0.015	0.046
Equal rights	0.022	-0.000	0.022
Democratic Participation	0.158	-0.122	0.280***
Political self efficacy	0.153	-0.118	0.271***
Interest in political life	0.167	-0.123	0.290***
N	1872	3291	
School Behavior	0.070	0.034	0.036
Absence	0.073	0.079	-0.006
Punctuality	0.025	0.029	-0.004
Exclusion	0.100	-0.002	0.102
Smaller sanctions	-0.001	0.027	-0.029
Av. Grade	0.249	-0.077	0.326***
History-Geography	0.293	-0.112	0.405***
Sport	0.240	-0.112	0.353***
N	282	657	

Note: This table shows the average baseline civic skills and academic performance of students in our sample, by student experience as representative, for the pooled sample of countries participating in the experiment. Statistics for the School Behavior index and sub-indexes and for grades are computed on the control group only. * p<0.10, ** p<0.05, *** p<0.01.

Table D3: Outcomes at baseline (standardized), by student type

	(1)	(2)	(3)	(4)	(5)	(6)
	Female	*Male*	Diff.	*High SES*	*Low SES*	Diff.
Civic Attitudes	0.099	-0.110	0.209***	0.069	-0.079	0.148***
Social Engagement	0.065	-0.121	0.186***	0.019	-0.077	0.096***
Tolerance	0.085	-0.085	0.170***	0.075	-0.067	0.142***
Equal rights	0.160	-0.157	0.317***	0.115	-0.102	0.217***
Democratic Participation	0.004	-0.041	0.044	0.136	-0.172	0.308***
Political self efficacy	-0.052	0.022	-0.074**	0.112	-0.151	0.263***
Interest in political life	0.062	-0.100	0.161***	0.158	-0.189	0.348***
N	2588	2543		2655	2500	
School Behavior	0.118	-0.091	0.208***	0.122	-0.056	0.177***
Absence	0.033	0.004	0.030	0.124	-0.041	0.166**
Punctuality	0.122	-0.087	0.209***	0.091	-0.084	0.175**
Exclusion	0.171	-0.163	0.335***	0.125	-0.060	0.184**
Smaller sanctions	0.174	-0.145	0.319***	0.139	-0.067	0.206***
Av. Grade	0.097	-0.099	0.197***	0.136	-0.110	0.247***
History-Geography	0.097	-0.118	0.216***	0.130	-0.122	0.252***
Sport	-0.042	0.022	-0.064	0.042	-0.021	0.063
N	505	530		521	479	

Note: This table shows the average baseline civic skills and academic performance of students in our sample, by student gender and social origin, for the pooled sample of countries participating in the experiment. Statistics for the School Behavior index and sub-indexes and for grades are computed on the control group only. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table D4: Citizenship Projects: Additional Features

	(1) All	(2) France	(3) Greece	(4) Spain
Hours spent in class on the project				
Total hours spent	21.26	20.86	21.17	21.74
	(10.44)	(6.04)	(8.08)	(14.61)
Preparation phase	9.47	9.51	8.92	9.70
	(5.46)	(3.92)	(3.81)	(7.32)
Implementation phase	11.40	10.94	13.04	11.02
	(8.54)	(5.66)	(6.56)	(11.46)
Implementation of ACT protocol key features				
Students voted to chose project	0.98	0.94	1.00	1.00
	(0.16)	(0.24)	(0.00)	(0.00)
Students worked in small groups	0.94	0.98	1.00	0.87
	(0.24)	(0.15)	(0.00)	(0.34)
Student groups formed randomly	0.86	0.95	0.89	0.76
	(0.35)	(0.22)	(0.32)	(0.43)
N	122	50	24	48

Note: This table describes the average characteristics of citizenship projects implemented over the 2018-2019 year by students in the treatment group, based on the endline teacher survey. Standard deviations are in parentheses.

Table D5: Peers' Characteristics: Descriptive Statistics

Variable	Mean	Std. Dev.	N
Class peers			
Prop. of stud with diff. gender	0.485	0.143	605
Prop. of stud with diff. geo origin	0.382	0.23	578
Prop. of stud with diff. social origin	0.453	0.184	615
Prop. of stud with diff. one difference	0.694	0.178	615
Prop. of stud with diff. two differences	0.274	0.161	615
Small working group peers			
Prop. of stud with diff. gender	0.436	0.301	609
Prop. of stud with diff. geo origin	0.189	0.274	609
Prop. of stud with diff. social origin	0.400	0.296	609
Prop. of stud with diff. one difference	0.709	0.276	609
Prop. of stud with diff. two differences	0.281	0.273	609
Friendship network at baseline			
Prop. of stud with diff. gender	0.212	0.271	552
Prop. of stud with diff. geo origin	0.202	0.299	552
Prop. of stud with diff. social origin	0.408	0.32	552
Prop. of stud with diff. one difference	0.613	0.333	552
Prop. of stud with diff. two differences	0.186	0.252	552

Note: This table shows the average characteristics of peers in the class, in the small working groups formed for the project and in the group of friends at baseline, for students in the treatment group for which the composition of the small working group is known.

Table D6: Teacher Characteristics by Teacher Involvement

	(1) High-involvement	(2) Low-involvement
Teaching experience (years)	19.11 (9.23)	16.71 (8.12)
Seniority in the school (years)	7.68 (5.99)	6.96 (5.16)
Citizenship teaching experience (years)	7.05 (8.46)	6.99 (7.78)
Female teacher	0.68 (0.47)	0.80 (0.40)
Teacher Pedagogy (baseline)	0.05 (0.96)	-0.05 (1.04)
Observations	181	193

Note: This table shows the average number of years of teaching experience (in total, in their current school, and specific to citizenship), gender and baseline Teacher Pedagogy index of teachers in our sample, separately on the half of the most involved teachers (Column 1) and the half of the least involved teachers (Column 2).

Table D7: Citizenship Projects: By Teacher Involvement

	(1) High-involvement	(2) Low-involvement
Project topics		
Discrimination	0.62 (0.49)	0.67 (0.47)
Social inclusion	0.53 (0.50)	0.53 (0.50)
Cultural diversity	0.31 (0.47)	0.27 (0.45)
Targeted population		
Elderly	0.15 (0.36)	0.15 (0.36)
Homeless	0.13 (0.34)	0.11 (0.31)
Migrants	0.28 (0.45)	0.24 (0.43)
Women	0.21 (0.41)	0.16 (0.37)
LGBT	0.13 (0.34)	0.09 (0.29)
Disabled	0.29 (0.46)	0.22 (0.42)
Other	0.24 (0.43)	0.27 (0.45)
No specific group	0.22 (0.42)	0.18 (0.39)
Observations	68	55

Note: This table shows the percentage of citizenship projects implemented in the treatment group that relate to each of the three topics covered by the ACT intervention, the population targeted by these projects and the share of in-school and out-of-school oriented projects, computed on the subsample of the half of the most involved teachers (Column 1) and the half of the least involved teachers (Column 2). One project may correspond to multiple topics and/or targeted population. Standard deviations are in parentheses.

Appendix E Attrition and balancing checks

Table E1: Attrition Analysis and Baseline Balance Checks for the Teacher Samples used in Table 2

	(1) ACT training	(2) ACT implementation	(3) Pedagogy
Attrition			
Observation not missing	-0.024	-0.011	0.008
	(0.042)	(0.041)	(0.040)
	[0.806]	[0.799]	[0.822]
N	323	323	323
Balancing			
Female	-0.040	-0.029	-0.035
	(0.068)	(0.067)	(0.066)
	[0.690]	[0.687]	[0.691]
Experience	-0.911	-1.112	-0.982
	(0.926)	(0.937)	(0.915)
	[18.28]	[18.22]	[18.36]
Seniority	0.216	0.220	0.184
	(0.886)	(0.875)	(0.838)
	[7.82]	[7.88]	[7.91]
School responsibilities	-0.126	-0.139	-0.132
	(0.117)	(0.116)	(0.116)
	[0.066]	[0.057]	[0.046]
Engagement out of school	0.013	-0.001	0.014
	(0.150)	(0.147)	(0.145)
	[0.042]	[0.043]	[0.038]
Years teaching citizenship	0.348	0.540	0.376
	(1.116)	(1.130)	(1.089)
	[8.937]	[9.010]	[9.048]
Studied citizenship init. training	0.059	0.066	0.046
	(0.060)	(0.059)	(0.058)
	[0.358]	[0.361]	[0.356]
Studied citizenship professional development	-0.022	-0.035	-0.032
	(0.057)	(0.059)	(0.055)
	[0.492]	[0.496]	[0.489]
Citizen project over last 2 years	0.014	0.009	0.023
	(0.062)	(0.062)	(0.060)
	[0.623]	[0.620]	[0.610]
Teacher Pedagogy index (Baseline)	-0.243	-0.219	-0.237
	(0.124)	(0.126)	(0.118)
	[0.000]	[0.000]	[0.000]
N	247	245	254

Note: The top panel of the table refers to the sample of teachers who participate in the experiment (N=323). For each of the three outcomes that measure the implementation of the program (i.e., training participation, project implementation, pedagogy), this top panel shows the result of regressing a variable indicating that the observation for this outcome is not missing on a treatment dummy, controlling for strata fixed effects. For each of the three outcomes, the bottom panel of the table refers to the sample of teachers who participate in the experiment for which the observation is not missing. For each outcome and each baseline variable, the bottom panel shows the result of regressing the baseline variable on a treatment dummy, controlling for strata fixed effects. Standard errors (in parentheses) are clustered at the school level. Variable means in the control are within brackets.

Table E2: Attrition Analysis and Baseline Balance Checks for the Student Samples used in Tables 2 and 4

	(1) Citizen. project	(2) Friendship
Attrition		
Observation not missing	-0.005	-0.000
	(0.022)	(0.022)
	[0.721]	[0.750]
N	6,211	6,211
Balancing		
Civic Attitudes index	0.006	0.001
	(0.036)	(0.035)
	[0.000]	[0.000]
Democratic Participation index	-0.045	-0.047
	(0.046)	(0.044)
	[0.000]	[0.000]
Age	-0.008	0.000
	(0.045)	(0.046)
	[14.07]	[14.07]
Female	-0.006	-0.012
	(0.017)	(0.017)
	[0.520]	[0.517]
European origin	-0.031	-0.030
	(0.016)	(0.017)
	[0.831]	[0.828]
High SES	-0.008	-0.009
	(0.021)	(0.020)
	[0.530]	[0.525]
Nb siblings	-0.061	-0.068
	(0.054)	(0.054)
	[1.802]	[1.817]
Representative	0.007	0.008
	(0.013)	(0.013)
	[0.352]	[0.352]
N	4,133	4,299

Note: The top panel of the table refers to the sample of students who participate in the experiment (N=6,211). For each of the two outcomes that measure the implementation of the program (i.e., participation in a citizenship project, friendship), this top panel shows the result of regressing a variable indicating that the observation for this outcome is not missing on a treatment dummy, controlling for strata fixed effects. For each of the two outcomes, the bottom panel of the table refers to the sample of students who participate in the experiment for which the observation is not missing. For each outcome and each baseline variable, the bottom panel shows the result of regressing the baseline variable on a treatment dummy, controlling for strata fixed effects. Standard errors (in parentheses) are clustered at the school level. Variable means in the control group are within brackets.

Table E3: Attrition Analysis and Baseline Balance Checks for the Samples used in Table 3

	(1) CB Idx	(2) Absence	(3) Punctuality	(4) Exclusion	(5) Smaller sanc.
Attrition					
Observation not missing	0.003	0.031	0.031	0.033	0.010
	(0.006)	(0.025)	(0.025)	(0.022)	(0.010)
	[0.982]	[0.959]	[0.959]	[0.921]	[0.973]
N	2, 290	2, 290	2, 290	2, 290	2, 290
Balancing					
Age	-0.073	-0.051	-0.051	-0.098	-0.071
	(0.095)	(0.095)	(0.095)	(0.096)	(0.096)
	[13.69]	[13.68]	[13.68]	[13.71]	[13.68]
Female	0.008	0.010	0.010	0.009	0.007
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
	[0.490]	[0.487]	[0.487]	[0.489]	[0.491]
High SES	-0.029	-0.035	-0.035	-0.027	-0.030
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	[0.507]	[0.510]	[0.510]	[0.501]	[0.507]
Financial aid	0.044	0.044	0.044	0.039	0.043
	(0.038)	(0.038)	(0.038)	(0.040)	(0.039)
	[0.216]	[0.216]	[0.216]	[0.223]	[0.218]
Nb siblings	-0.168	-0.160	-0.160	-0.176	-0.171
	(0.081)	(0.081)	(0.081)	(0.083)	(0.081)
	[2.338]	[2.333]	[2.333]	[2.357]	[2.346]
Grade 8	0.069	0.069	0.069	0.099	0.069
	(0.096)	(0.096)	(0.096)	(0.096)	(0.097)
	[0.614]	[0.614]	[0.614]	[0.584]	[0.610]
Delayed student	0.032	0.032	0.032	0.033	0.034
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	[0.127]	[0.127]	[0.127]	[0.129]	[0.124]
N	2,251	2,227	2,227	2,186	2,241

Note: The top panel of the table refers to the French sample of students for which administrative data was collected. For each of the 5 civic outcomes measured in this data (i.e., School Behaviour index, Absence, Punctuality, Exclusion and Smaller sanctions), this top panel shows the result of regressing a variable indicating that the observation for this outcome is not missing on a treatment dummy, controlling for strata fixed effects. For each of the 5 outcomes, the bottom panel of the table refers to the sample of students who participate in the experiment for which the observation is not missing. For each outcome and each baseline variable, the bottom panel shows the result of regressing the baseline variable on a treatment dummy, controlling for strata fixed effects. Standard errors (in parentheses) are clustered at the school level. Variable means in the control group are within brackets.

Table E4: Attrition Analysis and Baseline Balance Checks for the Samples used in Tables 3 and F2

	(1) Av. grade	(2) Sport	(3) Biology	(4) For. Lang. 1	(5) Art	(6) Music	(7) For. Lang. 2	(8) French	(9) Hist-Geo	(10) Math	(11) Phys-Chem	(12) Techno
Attrition												
Observation not missing	0.003	0.002	0.002	0.003	0.002	0.002	-0.004	0.003	0.003	0.003	0.003	0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]	[0.982]
N	2,290	2,290	2,290	2,290	2,290	2,290	2,290	2,290	2,290	2,290	2,290	2,290
Balancing												
Age	-0.073	-0.074	-0.074	-0.073	-0.074	-0.074	-0.079	-0.073	-0.073	-0.073	-0.073	-0.073
	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)
	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]	[13.68]
Female	0.008	0.008	0.008	0.008	0.008	0.008	0.007	0.008	0.008	0.008	0.008	0.008
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]	[0.489]
High SES	-0.029	-0.029	-0.029	-0.029	-0.029	-0.032	-0.029	-0.029	-0.029	-0.029	-0.029	-0.029
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]	[0.507]
Financial aid	0.044	0.043	0.043	0.044	0.043	0.043	0.046	0.044	0.044	0.044	0.044	0.044
	(0.038)	(0.039)	(0.039)	(0.038)	(0.039)	(0.039)	(0.039)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]	[0.216]
Nb siblings	-0.168	-0.171	-0.168	-0.171	-0.171	-0.165	-0.165	-0.168	-0.168	-0.168	-0.168	-0.168
	(0.081)	(0.080)	(0.080)	(0.081)	(0.080)	(0.080)	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)
	[2.314]	[2.314]	[2.314]	[2.314]	[2.314]	[2.314]	[2.338]	[2.314]	[2.314]	[2.314]	[2.314]	[2.314]
Grade 8	0.069	0.070	0.070	0.069	0.069	0.070	0.070	0.077	0.069	0.069	0.069	0.069
	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)	(0.096)
	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]	[0.614]
Delayed student	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032	0.032
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]	[0.127]
N	2,251	2,250	2,250	2,251	2,250	2,250	2,242	2,251	2,251	2,251	2,251	2,251

Note: The top panel of the table refers to the French sample of students for which administrative data was collected. For the average grade and for each of the 11 subject-specific grades collected in this data, the top panel shows the result of regressing a variable indicating that the observation for this outcome is not missing on a treatment dummy, controlling for strata fixed effects. For each of the 12 outcomes, the bottom panel of the table refers to the sample of students who participate in the experiment for which the observation is not missing. For each outcome and each baseline variable, the bottom panel shows the result of regressing the baseline variable on a treatment dummy, controlling for strata fixed effects. Standard errors (in parentheses) are clustered at the school level. Variable means in the control group are within brackets.

61

Table E5: Student Working Groups' Composition: Balancing Tests

	(1) female	(2) high SES	(3) French origin
Proportion of female	0.029		
	(0.032)		
Proportion of high SES		0.026	
		(0.041)	
Proportion of students with French origin			0.110*
			(0.051)

Note: This table shows the results of regressing student characteristics on the characteristics of other students in the same working group, on the sample of students for which group composition is known. These regressions control for the characteristics of other students in the class and for class fixed-effects to get rid of the bias due to correlations between individual characteristics and characteristics of the other members of the working group. Standard errors (in parentheses) are clustered at the school level. * p<0.10, ** p<0.05, *** p<0.01.

Table E6: Attrition Analysis and Baseline Balance Checks for the Samples used in Table 6

	(1) CA Idx	(2) Soc. engag.	(3) Tolerance	(4) Equal rights	(5) DP Idx	(6) Pol. Self-eff	(7) Pol. int.	(8) Clim Strike
Attrition								
Observation not missing	-0.002	-0.002	-0.003	-0.008	-0.001	0.001	-0.001	0.002
	(0.022)	(0.022)	(0.023)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)
	[0.740]	[0.740]	[0.718]	[0.718]	[0.749]	[0.738]	[0.749]	[0.738]
N	6,211	6,211	6,211	6,211	6,211	6,211	6,211	6,211
Balancing								
Civic Attitudes index	-0.001	-0.001	0.006	0.008	0.000	-0.006	0.000	-0.004
	(0.035)	(0.035)	(0.037)	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Democratic Participation index	-0.042	-0.042	-0.043	-0.041	-0.047	-0.046	-0.047	-0.046
	(0.044)	(0.044)	(0.045)	(0.046)	(0.044)	(0.044)	(0.044)	(0.044)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Age	-0.005	-0.005	-0.005	-0.012	-0.004	-0.003	-0.004	-0.002
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
	[14.07]	[14.07]	[14.07]	[14.08]	[14.07]	[14.07]	[14.07]	[14.07]
Female	-0.007	-0.007	-0.005	-0.006	-0.011	-0.009	-0.011	-0.005
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
	[0.516]	[0.516]	[0.520]	[0.522]	[0.517]	[0.517]	[0.517]	[0.517]
European origin	-0.032	-0.032	-0.031	-0.031	-0.031	-0.032	-0.031	-0.034
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
	[0.802]	[0.802]	[0.804]	[0.806]	[0.800]	[0.802]	[0.800]	[0.802]
High SES	-0.008	-0.004	-0.004	-0.006	-0.008	-0.004	-0.008	-0.006
	(0.020)	(0.020)	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)	(0.020)
	[0.525]	[0.525]	[0.525]	[0.529]	[0.529]	[0.525]	[0.525]	[0.525]
NB siblings	-0.059	-0.059	-0.063	-0.077	-0.061	-0.048	-0.061	-0.061
	(0.054)	(0.054)	(0.054)	(0.055)	(0.053)	(0.054)	(0.053)	(0.054)
	[1.809]	[1.809]	[1.804]	[1.807]	[1.813]	[1.804]	[1.813]	[1.804]
Representative	0.003	0.003	0.007	0.011	0.008	0.006	0.008	0.004
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
	[0.354]	[0.354]	[0.352]	[0.350]	[0.351]	[0.353]	[0.351]	[0.353]
N	4,244	4,244	4,119	4,110	4,294	4,241	4,294	4,244

Note: The top panel of the table refers to the sample of students who participate in the experiment (N=6,211). For each of the 8 civic outcomes measured in the student endline survey (i.e., Civic Attitudes index, Social engagement, Tolerance, Civic values, Democratic Participation index, Political self-efficacy, Interest in political life and Participation in Climate strike), this top panel shows the result of regressing a variable indicating that the observation for this outcome is not missing on a treatment dummy, controlling for strata fixed effects. For each of the 8 outcomes, the bottom panel of the table refers to the sample of students who participate in the experiment for which the observation is not missing. For each outcome and each baseline variable, the bottom panel shows the result of regressing the baseline variable on a treatment dummy, controlling for strata fixed effects. Standard errors (in parentheses) are clustered at the school level. Variable means in the control group are within brackets.

Appendix F Robustness checks and heterogeneity of Treatment effect

Table F1: Treatment Effects on Grades in the First Quarter

	(1) C	(2) T-C	(3) S.E.	(4) Unadj. p-val	(5) Adj. p-val	(6) N
Av. Grade - First quarter	0.000	-0.039	0.041	0.332	-	2251
History-Geography - First quarter	0.000	-0.002	0.055	0.971	0.971	2250
Sport - First quarter	0.000	-0.037	0.038	0.322	0.708	2248
Biology - First quarter	0.000	-0.132	0.074	0.073	0.399	2251
Foreign Lang. 1 - First quarter	0.000	0.033	0.060	0.580	0.797	2251
Art - First quarter	0.000	-0.091	0.077	0.241	0.662	2248
Music - First quarter	0.000	-0.012	0.072	0.868	0.955	2250
Foreign Lang. 2 - First quarter	0.000	-0.039	0.059	0.508	0.797	2242
French - First quarter	0.000	0.082	0.053	0.122	0.446	2250
Math - First quarter	0.000	0.019	0.059	0.748	0.914	2251
Physics-Chemistry - First quarter	0.000	-0.159	0.059	0.007	0.078	2250
Technology - First quarter	0.000	0.052	0.064	0.411	0.753	2251

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects. The third column shows the standard errors clustered at the school level. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Table F2: Treatment Effects on Grades in all Subjects

	(1) C	(2) T-C	(3) S.E.	(4) Unadj. p-val	(5) Adj. p-val	(6) N
Av. Grade	0.000	0.126	0.040	0.002	-	2251
History-Geography	0.000	0.196	0.056	0.000	0.005	2251
Sport	0.000	0.118	0.055	0.032	0.109	2250
Biology	0.000	0.032	0.068	0.636	0.778	2250
Foreign Lang. 1	0.000	0.040	0.053	0.452	0.624	2251
Art	0.000	0.138	0.067	0.040	0.109	2250
Music	0.000	0.057	0.076	0.450	0.624	2250
Foreign Lang. 2	0.000	0.010	0.052	0.854	0.879	2242
French	0.000	0.170	0.068	0.012	0.067	2251
Math	0.000	0.048	0.064	0.454	0.624	2251
Physics-Chemistry	0.000	0.077	0.049	0.114	0.252	2251
Technology	0.000	-0.009	0.057	0.879	0.879	2251

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school level. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Table F3: Treatment Effects by Country

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
France					
Participation in a citizenship project	0.254	0.541	0.048	0.000	1560
Civic Attitudes	0.000	0.199	0.070	0.004	1619
Democratic Participation	0.000	0.077	0.060	0.197	1647
Friendship Heterophily	0.000	0.155	0.059	0.009	1649
Greece					
Participation in a citizenship project	0.380	0.430	0.031	0.000	906
Civic Attitudes	0.000	0.101	0.071	0.158	922
Democratic Participation	0.000	0.128	0.086	0.138	930
Friendship Heterophily	0.000	0.094	0.115	0.413	932
Spain					
Participation in a citizenship project	0.298	0.392	0.041	0.000	1667
Civic Attitudes	0.000	0.115	0.054	0.033	1703
Democratic Participation	0.000	0.080	0.048	0.098	1717
Friendship Heterophily	0.000	0.118	0.061	0.055	1718

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school level. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Table F4: Replication of Table 6 after dropping Projects Related to our
Measure of Social Engagement

	(1) C	(2) T-C	(3) S.E.	(4) Unadj. p-val	(5) Adj. p-val	(6) N
Civic Attitudes	0.000	0.071	0.038	0.062	-	3469
Social Engagement	0.000	0.081	0.051	0.114	0.234	3469
Tolerance	0.000	0.002	0.036	0.957	0.957	3370
Equal rights	0.000	0.058	0.041	0.156	0.234	3360
Democratic Participation	0.000	0.079	0.036	0.028	-	3509
Political self efficacy	0.000	0.083	0.032	0.010	0.030	3466
Interest in political life	0.000	0.001	0.039	0.973	0.973	3509
Participation in Climate strike	0.000	0.050	0.041	0.225	0.338	3469

Note: This table replicates Table 6 when we drop the 40 schools that implemented a project directly related to our endline measure of social engagement or which project could not be classified.

67

Table F5: Treatment Effects by Student Gender

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
Female					
Participation in a citizenship project	0.317	0.462	0.029	0.000	2081
School Behaviour	0.000	0.284	0.102	0.005	1093
Av. Grade	0.000	0.163	0.037	0.000	1093
Civic Attitudes	0.000	0.045	0.045	0.314	2119
Democratic Participation	0.000	0.066	0.042	0.118	2139
Friendship Heterophily	0.000	0.112	0.051	0.027	2140
Male					
Participation in a citizenship project	0.286	0.422	0.028	0.000	1977
School Behaviour	0.000	0.219	0.105	0.037	1135
Av. Grade	0.000	0.076	0.053	0.151	1135
Civic Attitudes	0.000	0.170	0.045	0.000	2043
Democratic Participation	0.000	0.089	0.040	0.025	2072
Friendship Heterophily	0.000	0.101	0.046	0.028	2075

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school level. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Table F6: Treatment Effects by Student Family Background

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
High SES					
Participation in a citizenship project	0.307	0.476	0.029	0.000	2183
School Behaviour	0.000	0.256	0.106	0.016	1057
Av. Grade	0.000	0.205	0.043	0.000	1057
Civic Attitudes	0.000	0.095	0.048	0.047	2225
Democratic Participation	0.000	0.102	0.045	0.022	2247
Friendship Heterophily	0.000	0.071	0.047	0.130	2251
Low SES					
Participation in a citizenship project	0.294	0.400	0.029	0.000	1949
School Behaviour	0.000	0.228	0.091	0.013	1119
Av. Grade	0.000	0.071	0.062	0.251	1119
Civic Attitudes	0.000	0.110	0.042	0.008	2018
Democratic Participation	0.000	0.082	0.043	0.055	2045
Friendship Heterophily	0.000	0.097	0.045	0.033	2046

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Table F7: Treatment Effects by Experience as Representative

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
Representatives					
Participation in a citizenship project	0.382	0.379	0.029	0.000	1423
School Behaviour	0.000	0.244	0.110	0.027	586
Av. Grade	0.000	0.235	0.069	0.001	586
Civic Attitudes	0.000	0.190	0.050	0.000	1464
Democratic Participation	0.000	0.082	0.048	0.089	1480
Friendship Heterophily	0.000	0.092	0.052	0.075	1483
Non Representatives					
Participation in a citizenship project	0.258	0.454	0.030	0.000	2649
School Behaviour	0.000	0.215	0.080	0.007	1374
Av. Grade	0.000	0.107	0.044	0.016	1374
Civic Attitudes	0.000	0.037	0.041	0.376	2716
Democratic Participation	0.000	0.061	0.039	0.117	2749
Friendship Heterophily	0.000	0.097	0.044	0.026	2750

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school level. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Table F8: Treatment Effects by Student Baseline Social Engagement

	(1) C	(2) T-C	(3) S.E.	(4) p-val	(5) N
High baseline Social Engagement					
Participation in a citizenship project	0.356	0.434	0.030	0.000	2058
School Behaviour	0.000	0.212	0.094	0.024	985
Av. Grade	0.000	0.142	0.056	0.012	985
Civic Attitudes	0.000	0.167	0.041	0.000	2114
Democratic Participation	0.000	0.121	0.043	0.005	2135
Friendship Heterophily	0.000	0.097	0.051	0.055	2136
Low baseline Social Engagement					
Participation in a citizenship project	0.243	0.458	0.028	0.000	2049
School Behaviour	0.000	0.166	0.074	0.025	996
Av. Grade	0.000	0.128	0.045	0.004	996
Civic Attitudes	0.000	0.041	0.050	0.411	2104
Democratic Participation	0.000	0.044	0.038	0.255	2132
Friendship Heterophily	0.000	0.109	0.045	0.014	2136

Note: For each of the 12 row variables, the first column (column C) displays the mean of the row variable in the control group; the second column (column T-C) displays the coefficient from the regression of the row variable on a treatment dummy, controlling for strata fixed effects as well as for a set of controls selected from the full set of baseline variables through a Lasso procedure (Belloni et al. (2014)). The third column shows the standard errors clustered at the school level. The fourth column shows the corresponding unadjusted p-value while the fifth column shows the p-value adjusted for false discovery rate (Benjamini and Hochberg (1995)). The last column displays the size of the analysis sample, namely the sample of individuals who are observed at baseline and for whom the row variable is measured at endline. Each line corresponds to a separate regression.

Appendix G Student Outcomes by Treatment Status: Kernel Distributions

Figure G1: Distribution of Friendship Heterophily by Treatment Status
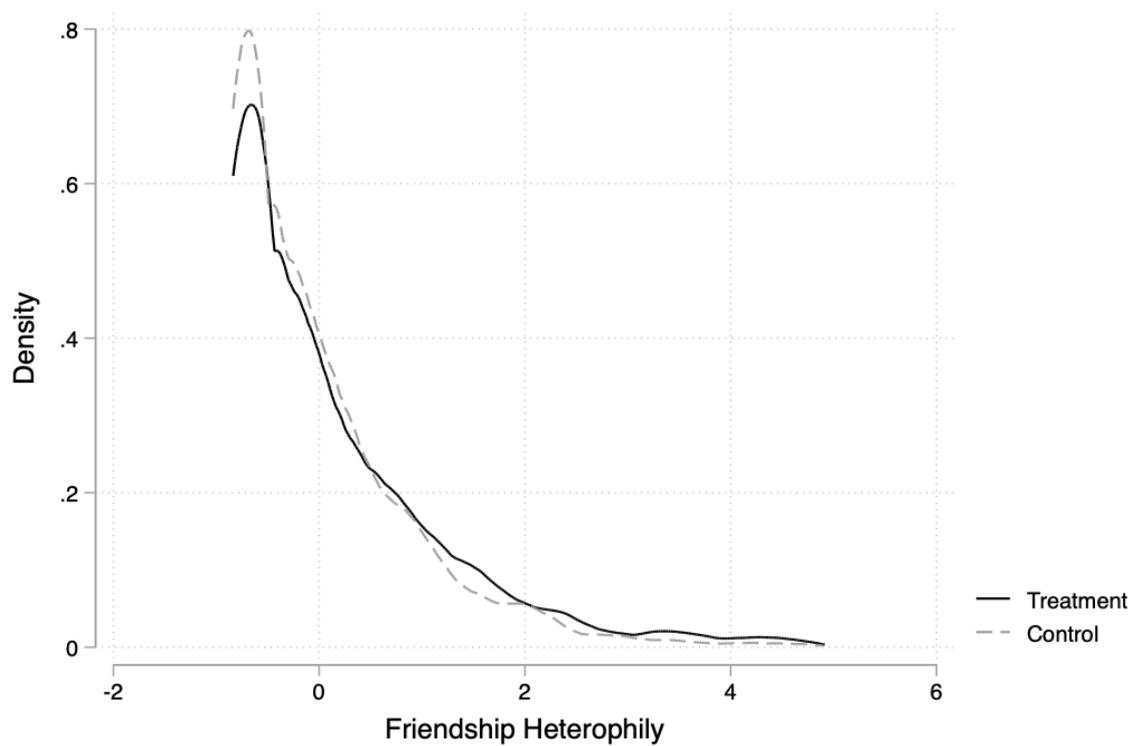

Note: Figure G1 show the kernel distribution of the standardized student friendship heterophily index, by treatment status.

Figure G2: Distribution of Civic Attitudes by Treatment Status
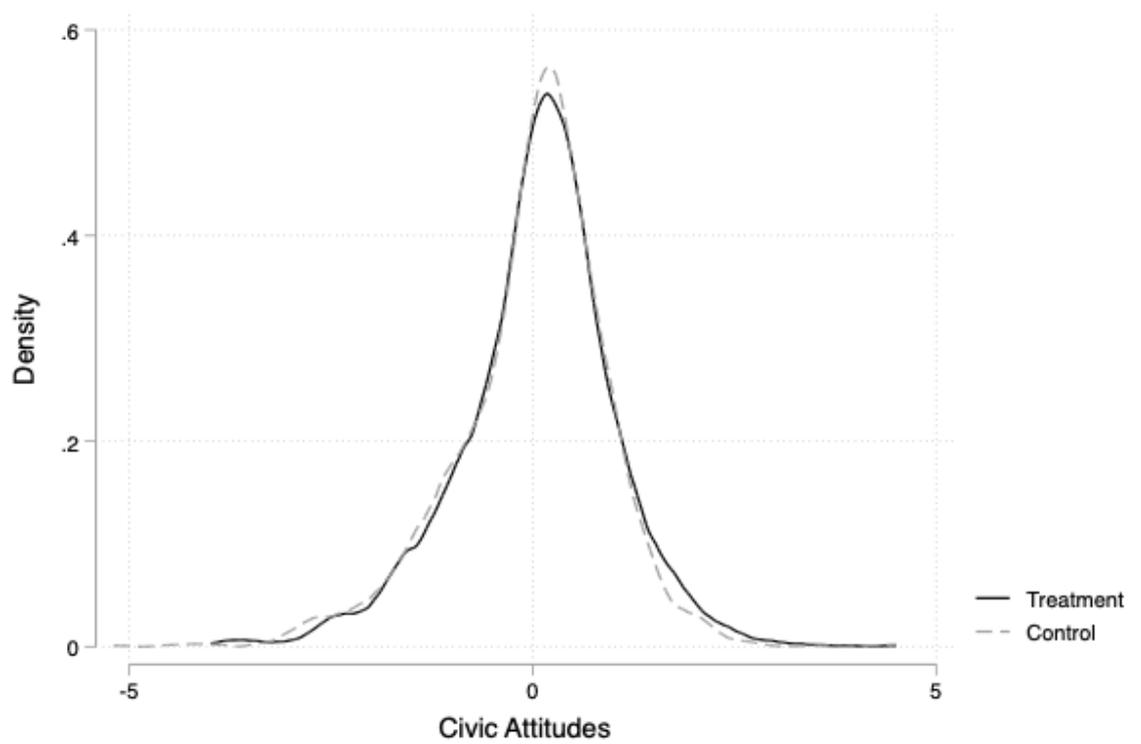
Note: Figure G2 show the kernel distribution of the standardized student civic attitudes index, by treatment status.

Figure G3: Distribution of Democratic Participation by Treatment Status
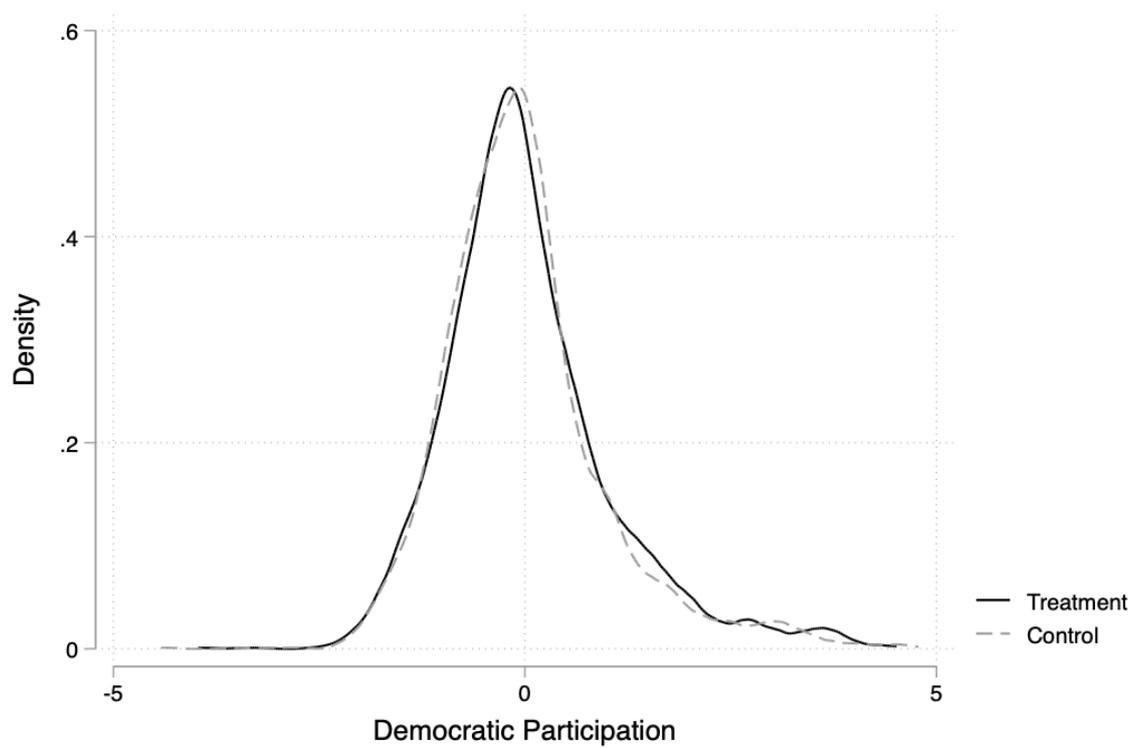
Note: Figure G3 show the kernel distribution of the standardized student democratic participation index, by treatment status.

Additionnal References

Briole, Simon , Marc Gurgand, Laura Lopez-Torres, Éric Maurin, Sandra McNally, Jenifer Ruiz-Valenzuela, Daniel Santin, Luis Schmidt, Gabriela Sicilia, Rosa Simancas and Yannis Tsirbas, "Active Citizenship Project (ACT) Quantitative Final Report", European Commission, 2020.

Cristofoli, Sophie, "En 2018-2019, l'absentéisme touche en moyenne 6% des elèves du second degré public", Note d'Information, n° 20.18, DEPP 2020.

Eriksson, Kimmo, Jannika Lindvall, Ola Helenius, and Andreas Ryve, "Socioeconomic status as a multidimensional predictor of student achievement in 77 societies," in "Frontiers in Education," Vol. 6 Frontiers 2021, p. 464.

Fréchou, Hélène, "Résultats de l'enquête Sivis 2021-2022 auprès des ecoles publiques et des collèges et lycées publics et privés sous contrat", Note d'Information, n° 23.02, DEPP 2023.