



Network

Rimini Centre for Economic Analysis Working Paper Series

wp 21-08

How effective is financial education? Evidence from the Emilia-Romagna region

**Stefania Mignani
Marcello Pagnini**

HOW EFFECTIVE IS FINANCIAL EDUCATION? EVIDENCE FROM THE EMILIA-ROMAGNA REGION

Stefania Mignani * and Marcello Pagnini **

Abstract

This paper describes the results from a questionnaire administered to students of schools in the Emilia-Romagna Region who participated in the Banca d'Italia-Miur project in the field of economic and financial education during the 2016-7 school year. The results show a significant increase in students' knowledge following a 10-hour course attended by schools teachers and held by personnel of the Bank of Italy. In particular, between the pre and post teaching activity phase, students increased the number of correct answers by about 5 units out of the 30 items in the questionnaire, compared to the improvement achieved by a group of students not involved in the teaching activity linked to the course held by the Bank of Italy. Our findings are robust to a wide set of controls suggested by literature on the topic.

1. Introduction¹

A rapidly expanding literature shows that individuals facing financial problems behave in ways differing significantly from courses of action that would be suggested by economic rationality². These cognitive and behavioural biases have been detected in decisions on savings, indebtedness, evaluation of financial asset risks and returns, in strategies for risk diversification and in many other contexts³. The pervasiveness of these problems appears even more worrying, since these do not only influence the choices made by ordinary citizens but end up having an impact also on the strategies followed by professionals in the financial markets (Bodnaruk and Simonov, 2015).

The debate on these issues is broad. Some argue that these phenomena do not necessarily have to be taken as examples of irrational behaviour but that, on the contrary, they can be considered as instances of a procedural rationality to which to resort under conditions of

* Department of Statistical Sciences of the University of Bologna.

** Bank of Italy, Bologna branch, economic research unit.

¹ The opinions expressed in the present work are those of the authors and do not reflect those of the institutions they belong to. The authors would like to thank Carlo Guiatti (Bank of Italy of Bologna) for the precious collaboration provided in all the management phases of the survey and Roberto Manfredi (Bank of Italy of Bologna) for the IT and data management aspects. We also thank the Emilia-Romagna Regional School Office and all the schools and teachers who took part in the courses and the test. Finally, we are extremely grateful to Magda Bianco, Riccardo Ferretti, Sauro Mocetti and participants in seminars held at the Bank of Italy and the University of Bologna (Rimini branch) for comments and suggestions.

² A collection of these cases is skilfully described and interpreted in the now famous contribution by Kahneman (2011). Ferretti et al. (2011) offer an excellent review of these issues within the context of the financial economics literature.

³ On the ways individual evaluate financial assets risk and returns, see Ganzach (2000) and the examples reported in the previously mentioned contribution by Kahneman (2011). On risk diversification, see Bernatzi and Thaler (2001).

limited and lacking information (Gigerenzer, 2015). Others argue that under certain circumstances orthodox financial economics would not even be able to provide clear indications on what would be the optimal strategies to be pursued by savers in their financial choices (see Willis, 2011).

In any case, there is a certain convergence between the opinions of scholars on the fact that financial knowledge and skills of individuals are weak and that in this sense it is necessary to provide remedy⁴.

The uncertainties regarding the diagnosis of the problems deriving from scarce knowledge and skills of individuals in the field of finance also end up influencing the debate about the prognosis. Some argue that there is a need for targeted training interventions in the field of finance (Lusardi and Mitchell, 2014), others do not believe that financial education is a valid tool in this sense (Willis, 2011) and claim that alternative strategies, e.g. courses in basic mathematics, can be more effective (Cole et al., 2014).

Financial education interventions differ considerably under various aspects, depending on, for example, the nature of the course promoter (public or private), the channels used for their implementation (schools, classroom lessons, video materials, mass media), recipients (students or adults), purposes (to act on the knowledge of the subjects concerned or on their behaviour), and finally the methodologies used.

To try to shed light on the effectiveness of these initiatives, a wide empirical literature has investigated the ability of financial education programs to increase the knowledge of individuals and/or to direct their financial behaviour towards more appropriate choices, ie toward outcomes more in line with what would be dictated by sound (rational) economic principles.

The increase of this literature has made it possible to develop contributions assessing the impact of financial education not directly but by comparing, with appropriate statistical methodologies, the results of other works published on the subject, thus allowing an evaluation in relative terms through the adoption of a common metric⁵. Fernandes et al. (2014), examining about 90 articles on the effectiveness of economics and finance courses, show that financial education programs contribute little both to increasing the knowledge of individuals and to improving the quality of their financial strategies. In addition, they find that the magnitude of the effects is negatively correlated with the rigor of the techniques used to measure them. For example, the intensity of the impact tends to decrease when we move from purely descriptive studies to those that use approaches that reproduce natural experiments, and finally, to studies, considered among the most rigorous, that resort to appropriate randomization (see below).

Recently Miller et al. (2015) and above all Kaiser et al. (2017) have partly corrected the scepticism of the previous works. In the second contribution, in particular, using a set of 126 papers, it is shown that courses in economics and finance not only have statistically significant impacts on knowledge and behaviour, but also have economically sizeable effects, if evaluated with an appropriate metric. In addition, the authors describe a trend whereby the most recent works that use randomisation tend to find more intense effects than in the past, partially correcting the impression that the results depend dramatically on the statistical methodologies used.

⁴ See among others Lusardi et al. (2010) and the report by Consob on Italian investors, Linciano et al. (2016).

⁵ For the assessment of the impact of financial education initiatives, the two most used indicators are the partial R-squared, which measures the contribution to variance explained by the variable of interest, and the share in terms of standard deviation of the dependent variable attributable to the impact of treatment.

Given the above mentioned uncertainty about the outcomes, this work aims to contribute to this debate, moving from the experience that the Bank of Italy gained in the context of financial education courses organized in recent years for the teachers of the Italian schools (the so called “training the trainers” approach). More specifically, the present work uses the results obtained from questionnaires administered to high school students in the Emilia-Romagna Region to assess the increase of economic and financial knowledge of the subjects involved as a result of the didactic activities carried out by their teachers participating in the Bank of Italy financial education program.

Our contributions to the literature are as follows. Although we were not able to carry out a rigorous randomization, we were however able to use a comparison group formed by a sample of students not involved in the didactic activity. The criteria used to define this aggregate (control sample) should help reduce the risk of overestimating the effect of the courses (see below). In addition, in the econometric analysis we introduce a wide set of controls on school teachers’ and students’ characteristics that help mitigate the problem of omitted variables and that are rarely found in other studies on the topic.

Our results confirm that, there is a significant potential to improve students’ knowledge in the field of economics and finance, even in a region such as Emilia-Romagna, which is considered to be among the ones ranking highest in Italy in terms of efficiency of the organization of school programs. In the phase that preceded the implementation of the course, the correct answers to a questionnaire listing 30 items did not exceed 50 percent on average. Our results show that, all other things being equal, courses for trainers in Emilia-Romagna have increased the knowledge of students involved. Between the pre and post course phases, the correct answers increased by 5 units (18 percent more if the increase is assessed in terms of the share of correct answers) compared to the corresponding improvement observed for the control sample. This impact, apart from being remarkable in terms of magnitude, is also robust to a wide set of controls. In addition, the positive effect of the courses seems to extend to students with different performances, although it is less intense for those concentrated in the lower tail of the distribution. This raises the question of whether it might be appropriate to diversify the courses in order to reach these subjects more effectively. Finally, based on our estimates, we show that the courses have had a positive effect also on subject matters not strictly related to those dealt with during the courses, thus contributing to activate further knowledge and skills in students (spillover effects).

The rest of the work is organized as follows. In Section 2 we describe the design of the test and illustrate the data used. Section 3 illustrates the econometric estimates and the main findings. Some extensions and further robustness checks are discussed in Section 4. Some final remarks and indications for future work on the topic are discussed in Section.

2. The nature of the proposed test and data

2.1 The design of the test

Since 2008 the Bank of Italy, jointly with the Italian Ministry of Education, Universities and Research (Miur), has organized training courses for teachers in schools of all typologies and levels in the areas of economics and finance. The project, coordinated at a national level, is implemented on a regional basis by the Bank of Italy's branches in the regional capitals, which organize and manage the financial education program. In the 2016-17 school year, in the Emilia-Romagna Region, four meetings, lasting about two and a half hours each, were held on the following topics: a) currencies and other payment instruments (basic module); b) price stability; c) the financial system and d) TPL Vehicle insurance. In addition to the contents presented by the teachers of the Bank of Italy, the course also supplied teaching materials that schools teachers could use during their own lectures with their students. The meetings were held between January and March 2017 every two weeks at the branches of the Bank of Italy in Bologna, Forlì and Piacenza, for a total of 12 meetings. Teachers were free to adhere to the financial education program and participants received a formal certificate for their participation. A total of 121 teachers from schools in the region joined the meetings. In most cases the teachers who took part in the courses carried out an educational activity in which they partially or fully reported to their students the didactic contents learned.

Following a similar initiative (Romagnoli and Trifilidis, 2013), we carried out a test to assess the effectiveness of the courses. The test administered to the students consisted of a questionnaire of 30 items divided into 6 sections that covered the topics addressed by the courses: Currency, Bank accounts and payment instruments, Inflation, Financial system and types of financial activities, risk diversification, Insurance (see the text of the questionnaire in Appendix 1)⁶.

The questions were submitted in two different phases: before the teaching activity carried out in class by the schoolteachers (Phase 0 or baseline of the test) and after this activity (Phase 1 or follow up of the test). The items of the questionnaires were the same in the two waves of the test. Between Phase 0 and 1, we only proceeded to modify the numerical data for quantitative responses and randomly changed the order of the multiple-choice responses. This arrangement is done to attenuate the effects of the learning process during the compilation of the test in the time span between the two phases and to reduce the impact of guessing.

During the meetings it was explained to the teachers of the schools that the participation in the test was considered as an essential part of the project, nevertheless a part of teachers did not want or could not participate in the test; furthermore, an even smaller share of teachers participated in both phases.

We define all the students who received training in financial education and who participated in the two phases of the test as the main sample or the "treated" (henceforth denoted by the letter T). In order to be able to compare the performance of the students involved in the didactic activity between the two phases, data was collected also for a

⁶ To each question, they were associated three possible answers, one of these always being the correct one. The student could not therefore indicate more than one answer for the same question. In the case of no reply or multiple answer, the question was not positively evaluated. Furthermore, there was a fourth possibility, that of answering with "I do not know", which, according to some recent researches, improves the ability of questions to discern students' abilities.

group of students who were not involved in the activity and who nevertheless participated in the two phases of the test. We define this set as a control or untreated sample (denoted by the letter C).

Due to the characteristics of the project, this set was not created on the standard statistical criteria used to define a real control sample or appropriate randomization of treatment. Instead, we asked all the teachers attending the courses and adhering to the tests and who thought in more than one class not to carry out the teaching activity related to the Bank of Italy financial education program for a subset of classes. According to our rule, teachers were supposed to administer the test to the latter group of students featuring in the control or C sample. When it was not possible to proceed in this way, the same teacher was asked to find a colleague of the same school who was not involved in the financial education program but who was available for administering the test to their own students⁷. Also in this case, only a subset of teachers participated with their own classes to the control sample.

Despite not following the classical criteria for an appropriate randomization, we argue later on that the way in which we define the C sample should not generate relevant distortions in the estimation of the impact of the financial education program.

Finally, the Phase 0 questionnaire also included a section dedicated to self-assessment of students about their knowledge and skills in the field of economics and finance and their perceptions and attitudes towards risk (both allow improving the evaluation of the results). Finally, we collected data on students regarding age, gender and marks in Italian and in mathematics in the 2015-16 school year. This last set of information is only available for a subset of observations.

2.2 Data

A total of 31 teachers and 951 students participated either in Phase 0 or in Phase 1 of the test, from 50 classes of 20 upper secondary schools.⁸ Of these, 26 teachers and 525 students from 33 classes in 20 upper secondary schools in the region took part in both the two phases. Since we are interested in measuring the effect that the participation in the financial education course had on the difference between scores obtained by the students in the baseline (Phase 0) and in the follow-up (Phase1) from now on we will refer to this set for all the following analysis.

More than 70 percent of students are in technical and vocational schools, the rest is enrolled in lyceums (Table 1)⁹. The classes involved are mainly those from the second year on, the modal class is the fourth year (Table 2). The geographical location of the schools that participated in the test seems to be quite heterogeneous, although there is a concentration in the province of Piacenza.

⁷ The operational instructions given to the teachers included the following. “As an example, consider a teacher of Economics and Law who carries out his/her teaching activity in sections A and B of a fourth year in the same School. The teacher could carry out the didactic activity related to the Miur-Banca d’Italia financial education course in section B of the fourth year and not carry it out in the section A of the same year. The latter would then act as control sample. In the absence of these conditions, the test could be submitted to a class of the same year and the same school although it is taught by a different teacher (who nevertheless has to respects the condition of not having participated in the Miur-BdI financial education course)”.

⁸ The students who actually responded to Phase 1 are 715; however, for some of these subjects it was not possible to associate them with those of Phase 0 due to errors in identification codes.

⁹ Among the high schools involved there are also those specialized in the socio-economic area.

As for participation in the control sample, this involved 10 teachers for 159 students. For over 2/3 of the students in the control sample and 6 teachers, the test was carried out by the same teacher who took part in the T sample, in the same school and same year but obviously in a different section¹⁰. As for the rest, the other students took the test with a different teacher from the one teaching the main sample, who did not take part in the courses of the Bank of Italy but who carries out didactic activities in the same school and grade as the teacher of the main sample, though in a different section.

Compared to C, the treated sample T shows a higher incidence of students in the fourth and fifth classes and, for what concerns the type of schools, lyceums and in part technical institutes. Also for this reason, in the econometric section we will introduce explanatory variables controlling for these differences between the two samples.

Before examining students' performances, a reliability analysis of the test was carried out. This was motivated from the fact that the questionnaire included both questions typically used to verify the skills in the financial and economic field and other questions that were developed for the specific purposes of this project. The Cronbach alpha indicator showed for the Phase 0 data a value of 0.65 slightly below the threshold but still acceptable. For Phase 1 there is an improvement with an alpha that reaches a value of 0.76, with all items perfectly consistent. Also the biserial correlation coefficients show a similar pattern. Their values are within a reasonable range from 0.2 to 0.5 for all items except the first two items in Phase 0 but again those values improve in the phase 1.

To start examining the results, we analyse the share of students who correctly answered the questions included in the survey (Table 3)¹¹.

The percentages of correct answers differ a lot both in Phase 0 and Phase 1: for some questions, the number of students who answered correctly does not reach 10 percent, while other questions have correct response rates between 80 and 90 percent.

The differences in the response rates considering all the 30 questions are similar for the controls and the treated. In general, and at a first glance, some weaknesses emerge in the area of knowledge of the institutional characteristics of currencies and of debit cards as means of payment, in the use of the notions of compound interest rate under various aspects, the relationship between the value of a security and its return, and finally the ability to read the movements in bank account balances over time. On the other hand, it seems that the students have some knowledge about the characteristics of the basic banking products (loans and deposits), on the effects of inflation and the characteristics of TPL Vehicle insurance. With regard to the definition of a financial risk, the ranking of financial assets in terms of risk and the notion of risk diversification, the performance appears to be mixed.

The similarity in the response patterns of samples C and T in Phase 0 is also confirmed by the high correlation between the two (the simple correlation coefficient is equal to .97, the Spearman correlation coefficient is approximately at the same level). In other words, the two groups of students seem to behave in a very similar way in Phase 0, facing similar difficulties in providing correct answers. It is interesting to note that the correlation

¹⁰ In only one case, a teacher carried out the test for both the main and the control samples, but with two different classes.

¹¹ Set $X_{sqg}=1$ if the student s ($s=1, \dots, N_g$), belonging to the group g , where $g \in \{P0, P1, C0, C1\}$ and P0 and P1 denote respectively the main sample in Phase 0 and 1 and C0 and C1 the students of the control sample in Phase 0 and 1, correctly answered the question q ($q=1, \dots, 30$), set $X_{sqg}=0$ otherwise. The means reported in the table 3

$$y_{qg} = \sum_s \left(\frac{X_{sqg}}{N_g} \right)$$

are defined by the following expression:

coefficient between the results achieved in Phase 1 by the two groups of students remains high at 0.89 and 0.90 – simple correlation coefficient and Spearman correlation coefficient respectively – but it is lower than the one just illustrated. In other words, the two samples seem to differ slightly more in their behaviour after the course.

The quantity of interest in this study is represented by the following difference:

$$\Delta y_T = E[y_{sT1} - y_{sT0}]$$

where the operator E denotes the expected value, y_{sT1} and y_{sT0} represent respectively the overall scores obtained in the questionnaire by the student s belonging to the sample T obtained in Phase 0 (pretreatment) and in the one after treatment (Phase 1)¹². Furthermore, we are interested in determining whether the difference in scores obtained between the two phases can be attributed to the effect of the treatment.

Although the scores show an improvement between the two phases, this evidence is not sufficient to conclude that the latter can be attributed to the effect of the course followed by the teachers. Firstly, there could be a learning effect of the students in compiling the questionnaire (Becchetti and Pisani, 2012). Secondly, in the interval between the two tests there could be multiple factors that lead to an increase in cognitive abilities of the students (e.g. advance of the ordinary school programs, improvement of cognitive abilities due to the passing of time, particularly relevant in age groups involved in our analysis).

For this reason, it is useful to look at sample C for which a similar quantity can be estimated:

$$\Delta y_C = E[y_{sC1} - y_{sC0}]$$

which, appropriately compared with the previous one, shows how much of the improvement observed in the sample T can be attributed to the effect of the course rather than to factors such as learning while filling in the questionnaire or improvement of students' skills between the two phases, which T and C share and hence are not related to the effects generated by the course.

In other words, we are interested in estimating the following quantity, defined from now on as Extra improvement:

$$\Delta y_E = \Delta y_T - \Delta y_C$$

To consider this method valid, however, it is necessary to ensure that the composition of the two groups of students in T and in C does not differ for systematic factors in the pretreatment phase which, if not properly controlled for, could determine the difference Δy_E instead of the treatment. In the next section, we will provide an estimate of these

¹² For most of the work the student's overall score will be defined by the correct answer share, which is: $y_{sg} = \sum_q \left(\frac{X_{sqg}}{30} \right)$ with $g \in \{T0, T1, C0, C1\}$, however we will also use the indicator based on the estimated scores described later in the text.

magnitudes and discuss the problem of the comparison between the two groups of students in T and C.

In order to obtain a measurement that is the most rigorous and comparable in different situations, the overall raw score was accompanied by an estimate of the unobserved student abilities, obtained using the Item Response Theory approach, typically used in educational research.

In the present work, a two-parameter model for binary data was considered (Hambleton and van der Linden, 1997) to estimate both the characteristics of the items, like the difficulty of each question and its discriminatory power, and the ability of each student. The latter estimate is obtained with the joint contribution of all the responses, assuming a normal distribution with 0 mean and unit variance.

Before illustrating the results of the econometric estimates, we will report some descriptive statistics on the raw and estimated scores (Table 4)¹³.

The table shows that the raw score obtained by students in the T and C groups in Phase 0 does not exceed 50 percent, confirming the relatively low level of knowledge in economics and finance that have already emerged in other contributions. The mean raw and estimated scores are slightly higher for the treated students in T than for those in C in Phase 0 (see also Figure 1). However, tests on the difference between mean and medians clearly indicate that these differences are not significantly different from 0¹⁴. Therefore, at least from the point of view of the performance of the students in the two groups, no significant differences seem to emerge that can be referred to the pre-treatment phase.

As expected, the comparison of the evolution of the distribution of the scores for T and C between the two phases clearly shows an improvement for both groups of students, even if it is clear that the rightward shift in the distribution is more accentuated for T with respect to C (Figure 2).

In the case of raw scores and without conditioning on any explanatory variable we obtain $\Delta y_T = 0.125$ (s.e. 0.0153***) which is significantly different from 0 at 0.01 level of significance¹⁵. For C we obtain: $\Delta y_C = 0.024$ (s.e. 0.0134*) significantly different from 0 at 0,1 level, and finally the extra improvement is equal to $\Delta y_E = 0.100$ (s.e. 0.0200***), significant at 0.01

3. Main econometric evidences

As already discussed in the previous section, it is important to make sure that the differences between the scores of the T and C samples are not due to unobserved systematic factors that could lead us to erroneously attribute the extra improvement

¹³ In this case, the various moments of the distribution are defined starting from: $y_{sg} = \sum_q \left(\frac{X_{sqg}}{30} \right)$ and from the quantities defined by the estimated scores.

¹⁴ The results described were obtained for the mean test by regressing with the OLS estimation method the score in Phase 0 on a dummy equal to 1 if the student is part of the T sample (student treated) and 0 otherwise. For the test on the difference between medians, the qreg2 stata procedure was used that allows the clustering of errors as explained in the note of Table 7. The permanence of a positive difference in favour of the students in the T sample, although not significant, can still be easily explained by the fact that this sample shows a greater share of students in lyceums and more advanced classes (fourth and fifth year), who usually tend to perform better than others.

¹⁵ From here on, the estimates are reported providing the estimated value of the parameter and the standard error in brackets. As always: ***, ** and * indicate statistical significance levels of 1, 5 and 10 percent respectively. Finally, standard errors are always clustered at the level of the single class and section.

obtained by the students in T to the course in financial education rather than to alternative uncontrolled and unobserved factors.

To this end, we adopt a specification that is standard in the literature on the subject:

$$y_{s1} = \alpha + \beta y_{s0} + \gamma TR_{s \in h} + \sum_d \delta_d I_{ds} + \tau' SC_s + \theta' PR_s + \vartheta' ST_s + \varepsilon_{s1} \quad (1)$$

where the overall score for the whole questionnaire obtained by the student s in phase 1 has been regressed on a constant, the same score obtained in the base line (Phase 0), a treatment dummy $TR_s = 1$ if s belongs to T , with $b \in \{T, C\}$ and 0 otherwise, plus a set of variables that are related to three types of controls: a) teacher characteristics, with dummy $I_{ds}=1$ if the student s belongs to a class of teacher d and 0 otherwise; b) types of school/class, represented by the vector SC ; c) methods of conducting the test, represented by the PR vector; d) individual characteristics of the student, included in the ST vector¹⁶.

a) On the first aspect it must be noted that the teaching activity carried out by the personnel of the Bank of Italy was held by the same teacher for each module, and the material used was the same for the three locations where the courses were held, while school teachers were free to attend from one to four modules. Furthermore, the latter were given complete autonomy regarding the ways of teaching the contents learned during the courses (including for example the hours of teaching to be dedicated to the discussion of the various topics). Also, the school teachers differ with regard to the subjects taught, school in which they work, age, experience, individual attitudes, etc. To take into account all these potential factors that may have an impact on the variable of interest, a teacher fixed effect was introduced in the specification¹⁷.

b) Another important source of heterogeneity that, if not appropriately controlled for, could affect the results is represented by the characteristics of the school and the class in which the student is enrolled. In order to consider this problem, two sets of dummy variables were introduced indicating the type of school (technical, vocational and lyceum) and the class attended (from the first to the fifth year) respectively.

c) Another set of controls refer to context variables that describe the modalities with which the tests were carried out. Although a test length of 75 minutes was recommended, school teachers could decide for themselves how long the test would last, and although we indicated the periods in which to perform the test, they were also free to choose the dates of the test, provided it was held in the course of the 2016-17 school year and respecting the criteria for defining Phases 0 and 1. Finally, the test could be administered on paper or on computer. To take into account all these aspects, the following variables were introduced in the specification: the length of the test expressed in minutes, the time interval between Phases 0 and 1, expressed in number of days, and finally a dummy = 1 if the modality of compilation of the test was on computer and 0 otherwise. The role of all these factors in influencing the scores of the students is treated by a wide literature on the subject¹⁸.

d) Finally, the decisive factors that can explain both the cross sectional and time variability in the scores include the abilities and behavioural attitudes of individual students. To this end, the specification includes several groups of dummy variables: the first one is derived from the students' self-evaluation about their own financial literacy, the second reports the

¹⁶ For a detailed description of the definition of the variables used in the regression, see Appendix 2.

¹⁷ For a similar approach see Walstad et al. (2010) and Bosshardt and Watts (1990).

¹⁸ On the importance of checking for the variables of the context in which the test takes place, see among others Koundinya et al. (2016).

evaluation of the degree of risk that the student attributes to certain actions and finally the propensity to undertake such actions. It must be noted that the variable on self-assessment can be positively correlated, as well as to the cognitive abilities of the student, also to a degree of self-esteem that in turn could reflect the degree of optimism of the subject involved¹⁹. In other specifications, we will introduce other individual characteristics of the students that will be explained later.

As an alternative to equation (1), we could have adopted the traditional diff-in-diff specification that reads as follows:

$$y_{st} = \alpha + \beta POST + \gamma TR_{s \in h} + \delta TR_{s \in h} * POST + \vartheta' X_{st} + \varepsilon_{st} \quad (2)$$

where y_{st} is the student s score in phase t ($t=0,1$) and $POST$ is a dummy variable equal to 1 for $t=1$ and = 0 otherwise and X_{st} is vector of student time varying characteristics. In the latter case the parameter of interest for gauging the effect of the financial education programme would be δ .

We prefer specification in equation (1) for the reasons explained in Mc Kenzie (2012). Basically, the author shows that the latter improves the power of the test as compared to (2) whenever the correlation between y_0 and y_1 is not too high. In our case, the correlation between scores in the base line and the follow up is positive and significant but sensitively below .5 (a threshold that Mc Kenzie indicates as a relatively low value). In any case, we will resort to specification in (2) in the robustness check section.

Table 5 shows some descriptive statistics for a subset of variables used for regression and also a test to verify whether the C and T samples differ significantly under these profiles. The reported p values do not seem to indicate systematic differences between the two samples regarding the methods of carrying out the tests, the skills and attitudes of the students and finally the composition of the two groups by gender and nationality. Therefore, not only the two samples are sufficiently similar to each other with regard to performance of the students in the pre-treatment phase but also they are homogeneous from the point of view of observable characteristics of the students who compose them.

Going back to the econometric analysis, the standard errors of the parameters are always clustered at the level of class and section. Moreover, having conditioned the score in Phase 1 on the base line outcomes in 0 implies that the effect of the explanatory variables can be interpreted as impact on the difference between the scores obtained in the two phases. Finally, the parameter of interest is γ which measures the quantity Δy_E conditional on the controls introduced in the specification. The results for the estimated γ are shown in Table 6.

The impact of the course on the financial knowledge of the students involved in the teaching activity in the classroom is remarkable. The estimated value of γ is always positive and significantly different from 0 in all columns. In column I, the proposed specification includes as explanatory variables, in addition to the treatment dummy, only

¹⁹ As regards the importance of checking, as well as for the cognitive abilities of the students, also for some traits of their behavioural attitudes, in order to limit the problem of the variables omitted in the estimates, see Fernandes et al. (2014). Meier and Sprenger (2007) show how the lack of control for some of these characteristics, for example the temporal preferences of the subjects participating in a test, can create an upwards bias in the evaluation of the effectiveness of a course due to a problem with the selection of the sample of the subjects treated. Finally, Puri and Robinson (2007) show how a moderate degree of optimism can positively influence various economic choices.

the students' score in Phase 0. In this case the extra improvement in the response rate of the sample T is equal to about 12 percent and corresponds to about 3 extra correct answers compared to those indicating the improvement of the control sample. Once the other controls described above (see columns II through V) were introduced, the impact of the course further increased. In column II (baseline specification), it is equal to an extra improvement of scores for the treated students of 18 percent, which corresponds to about 5 extra correct answers per questionnaire. The latter result is substantially confirmed if additional controls picking up student characteristics are added to the baseline specification (grades in math, gender, nationality, see column IV)²⁰. Finally, in column III the estimated score was used as a dependent variable, obtaining results similar to those described above.

As motivated above our preference goes to specification in (1). Despite that and as a further robustness check, we also use the diff in diff methodology as represented in equation (2). Through that, it is possible to introduce a student fixed effects thereby getting rid of any residual problem related to the time invariant student unobserved characteristics that might bias our evaluation of the financial education program. Results (see column VI in Table 6) are extremely encouraging: estimated parameter for the interaction term between the POST and the treatment dummies is positive and significant at 1 per cent level. The magnitude of the effect is also remarkable as it amounts to an extra improvement of 3 correct answers for the treated student²¹.

The improvement in students' financial knowledge is quite large compared with the findings in other contributions in the same stream of literature, but remains in any case a comparable order of magnitude with at least a part of the works on the subject. Romagnoli and Trifilidis (2013) have assessed the effects of the courses held by the Bank of Italy in past years using a questionnaire similar to ours but including the section on currencies and payment instruments only. In short, they estimate for upper second grade schools Δy_T equal to about 15 percent without however conditioning for any explanatory variable and without the option of a control sample to estimate the extra improvement²². Becchetti e Pisani (2012) use a randomized treatment procedure and for a sample of Italian high school students, conditioned to a series of explanatory variables, and obtain an estimate of the extra improvement attributable to a course in finance, defined by Δy_E , very similar to ours, especially for what concerns the estimates referring to schools located in northern Italy, which are more easily comparable with our sample of schools in Emilia-Romagna. Using treatment randomization, Brugiavini et al (2018) estimate an improvement in knowledge, for a sample of Italian university students following the screening of a short video on finance, of the response rate between 10 and 12 percent. Waldstad et al (2010) estimate the effects of a financial education course on the knowledge of a sample of students in the US. By constructing a control sample with similar criteria to those followed in this work, they obtain results comparable in size to those described above.

Resorting to a large sample of Brazilian schools and distributing the treatment randomly, Bruhn et al. (2016) find that a financial education course has a significant effect on the

²⁰ In the latter case, the number of observations is drastically reduced because such information is available for a subset of students. Furthermore, as proof of robustness we have also introduced the student's assessment in Italian among the controls without this modifying the results commented in the text.

²¹ In unreported evidence (available on request) we also experiment the diff in diff methodology by using alternative specifications like that in column III in Table 6 or another one where we use student fixed effects plus a set of controls interacted with the POST dummy. In all cases results widely confirmed those reported in the text.

²² The difference between the pre and post teaching activity averages is always significant at one percent.

extra improvement for the treated sample.²³ However, the magnitude of the effect amounts approximately to 25 percent of the standard deviation (SD) of the dependent variable. For comparison, in our work the order of magnitude, using the same metric, would be 150 percent and 71 per cent SD if we consider, respectively, our preferred specification in (1) and the diff in diff specification with student fixed effects in (2). Lührmann et al. (2015) for a sample of German teenagers, Hospido et al. (2015) for an aggregate of Spanish students obtain, with methodologies that are in part similar, estimates of Δy_E , which are more similar to the ones found in the previously mentioned work than to ours. However, to partially mitigate the differences with these works, we can mention the already cited work by Kaiser et al. (2017), which shows that the average increase in knowledge due to financial education courses held in schools, identified in 35 researches, amounts to about 37.3 percent of the standard deviation, a share that further increases if one considers that the effect is usually higher for countries with a high per capita income like Italy (see Table 1 on page 619 of the aforementioned paper).

An interesting aspect investigated by the literature on the subject concerns the possibility that the financial education courses provided to students are more or less effective depending on the abilities of the students themselves, in other words we ask ourselves if there is an appreciable heterogeneity in the effects of the treatment. An improvement in our knowledge along this direction could help us to decide whether differentiating didactic activity in order to increase its general usefulness instead of offering the same contents to all the students.

To investigate this aspect, we replicated the basic specification in column II of Table 6 using quantile regression (Table 7). From these additional findings, it emerges that the impact of the courses is positive for all students, regardless of their performance. In fact, the estimated parameter of interest is always positive and significantly different from 0 and never falls below 10 percent for all percentiles considered. However, a sort of heterogeneity in the effect of the treatment does emerge. The positive impact of the courses seems to be more intense for the students with a score between the tenth percentile and the median performance, while it seems to be lower for the students with scores above the 75th percentile and, above all, for those in the lower tail of the distribution (5 percentile scores). In other words, financial education courses seem to have a positive effect on the financial literacy of all students, even if some problems of effectiveness in comparative terms seem to emerge for students with weak performances.

4. Discussion

Thanks to a pre and post treatment phase, a control sample and a large group of regressors, the proposed empirical analysis should generate an estimate that reflects a genuine causal relationship between the attendance of the financial education courses, on the one hand, and the improvement of the financial knowledge of the students interested in the course, on the other. However, the way in which we define the control sample, as already mentioned, did not ensure that the treatment was randomly distributed among the students. More specifically, the latter was defined by all the teachers that were in the sample T and who accepted to carry out the two phases of the test either directly or indirectly, through a colleague, for a class similar to the one included in T but not involved in the teaching activity related to the courses.

²³ Other papers using similar randomized experiment to assess the impact of financial education programs include Bover et al (2018) for Spain and Frisancho (2018) for Peru.

These rules of composition for C should limit possible distortions in the assessment of the impact of the financial education course, which might be caused by the fact that for the control sample teachers were selected with particular characteristics that, if not properly controlled for, could lead to an overestimation of the effect of the courses. The classic example in this situation is represented by the case in which the subset of the teachers who accept to be part of the aggregate C are those who show, for example, a strong propensity to improve and invest in new knowledge, able to produce in turn a significant increase in students' knowledge and skills, regardless of the effect of the financial education course. If this should be the case, the simultaneous presence in T and C of these teachers should generate an improvement in the performance of the students of both groups, thus making the parameter measuring the effect of the financial education course insignificant²⁴. Therefore, the characteristics of the control sample could generate a downward distortion of the estimated parameter of interest, making the estimates of the impact of the financial education course contained in Table 5 conservative.

In any case, in T there are teachers who did not participate in the control sample and in C there is a marginal presence of teachers who are not present in T. To ensure that our results are not influenced by the different composition of the two samples we narrowed the sample to those teachers present in both T and C (see column V in Table 6). Proceeding in this direction, we see a sharp fall in the number of observations but the impact of the financial education course remains positive and significant, also for this small sample, with an extra improvement in the score of students of 19 percent.

In terms of balancing the T and C samples, we already observed the latter is more concentrated on fourth and fifth year classes. To rebalance the two sample a class dummy was introduced in the baseline specification. Despite that, it is still possible that treated students benefit more than others for the fact of being exposed for a longer period to the teaching activity of the more motivated teachers overrepresented in sample T. This would generate an positive bias in our results. To control for that, we restricted the sample to student in IV and V classes. Even for this case, the estimated parameter for the treatment effect is positive, large and extremely significant (the estimated parameter for the treatment dummy is .197, s.e. 0.008***).

Another important aspect in the evaluation of the effectiveness of the course concerns the heterogeneity of its effects across types of questions. For example, the students could answer some questions by drawing directly on the information provided in the courses for their teachers, others required the activation on the part of the student of additional logical-deductive skills, or covered areas beyond the scope of the subjects treated in the course. In order to verify whether the course was effective in improving the answers of the students also in these areas, we selected eight questions for which the course did not directly provide the possibility of answering and computed the total amount of correct answers²⁵.

²⁴ Another situation that would tend to produce a downwards distortion of the effect of the course could be that in which the students of the treated and non-treated classes communicate with each other about the course and the test results, thus reducing the possible differences between the performance of the two groups (treatment contamination). On the contrary, what might contribute to produce an upward distortion is a situation in which the teacher who participated with two different classes to T and C decides knowingly, in his/her teaching activity, to penalize the second in order to bring out a greater effectiveness of the course. However, given the design of the experiment, there is no reason a teacher should adopt such a strategy; therefore, we consider this event to be not very plausible.

²⁵ In particular, the group of questions identified referred to the calculation of simple and compound interest (questions 5 and 7), the ability to compare financial balances and the effects of inflation over time (questions 9 and 15), the relationship between price and return of a title (question 20), the choice of the type of loan (question 21), diversification of risk (questions 25 and 26).

Using the total score in these 8 questions as dependent variable, we proceeded to replicate the econometric analysis carried out previously, the results are shown in Table 8.

As expected, for questions that dealt with issues less directly related to the course the students achieved an extra improvement of scores lower than the one obtained with other questions (see Table 6). However, the estimated parameter for the treatment remains positive and significant and approximately amounts to 10 percent in the baseline specification (see column II). In other words, the course seems to have been effective in improving the students' knowledge and skills in economics and finance beyond the areas directly addressed by the courses for teachers.

5. Final comments and indications for future research

This work belongs to a stream of literature dealing with the evaluation the effectiveness of teaching initiatives in the field of the financial education. With regard to the courses held by the Bank of Italy in Emilia-Romagna, the results are encouraging. The students who took part in a test before and after the teaching activities carried out in the classroom registered a significant increase in their financial competence, which also proves robust to a series of controls typically used in recent statistical and econometric literature. In addition, positive effects also seem to extend to economic and financial matters not directly dealt with in the courses (spillover effects). Moreover, the improvements seem to be transversal to the different performances of the students, although some concern emerges regarding effectiveness in relation to the students with results in the low tail of distribution.

In the future, we would like to extend the evaluation of the impact of the courses to financial choices actually carried out by the students. Finally, in addition to quantify the effects, there is also the question about the more effective channels through which to promote an improvement in young people awareness of these matters²⁶.

²⁶ On these aspects, see Drexler et al. (2014) that show how under certain conditions it is better to convey simple and less complex rules and concepts. Berg and Zia (2017) argue that a significant improvement in knowledge can be achieved by mobilizing the emotional components of individuals, spreading information and knowledge to be transmitted through channels such as television and other mass media instead of more traditional lesson-based methods.

References

- Batty, M., Collins, J. M., Odders-White, E. (2015). Experimental evidence on the effects of financial education on elementary school students' knowledge, behavior, and attitudes. *The Journal of Consumer Affairs*, 49(1), 69-96.
- Becchetti, L., Pisani F. (2011). "Financial education on secondary school students: The randomized experiment revisited." *EconomEtica Working Paper* 34.
- Benartzi, S., Thaler, R. (2001). "Naive Diversification Strategies in Defined Contribution Saving Plans". *American Economic Review*, 91(1), 79-98.
- Berg, G. Zia, B. (2013). "Harnessing emotional connections to improve financial decisions. Evaluating the impact of financial education in mainstream media". *Journal of the European Economic Association*, 15(5): 1025-1055
- Bosshardt, W., Watts M. (1990). "Instructor Effects and Their Determinants in Precollege Economic Education". *Journal of Economic Education*, 21(Summer): 265–276.
- Bover, O., Hospido, L., & Villanueva, E. (2018). "The impact of high school financial education on financial knowledge and choices: evidence from a randomized trial in Spain.", *Documentos de Trabajo*, N.º 1801
- Brugiavini, A., Cavapozzi, D., Padula, M., Pettinicchi, Y. (2018). "On the effect of financial education on financial literacy: evidence from a sample of college students", *Journal of Pension Economics and Finance*. 1–9
- Bruhn, M., de Souza Leao, L., Legovini, A., Marchetti, R., Zia, B. (2016). "The impact of high school financial education: Evidence from a large-scale evaluation in Brazil." *American Economic Journal: Applied Economics*, 8(4), 256-295.
- Cole, S., Paulson, A., Shastry, G.K. (2014), "High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses". Harvard Business School Working Paper 13-064.
- Collins, J. Michael. (2013). "The Impacts of Mandatory Financial Education: Evidence from a Randomized Field Study." *Journal of Economic Behavior and Organization* 95: 146–58.
- Drexler, A., Fischer G., Schoar A. (2014). "Keeping It Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics*, 6(2): 1-31.
- Fernandes, Daniel, John G. Lynch Jr., Richard G. Netemeyer (2014). "Financial Literacy, Financial Education and Downstream Financial Behaviors." *Management Science*, 60 (8): 1861–83.
- Ferretti, R., Rubaltelli, E., Rumiati, R. (2011). "La mente finanziaria. Economia e Psicologia al Servizio dell'Investitore", Il Mulino, Bologna.
- Frisancho, V. (2018). "The Impact of School-Based Financial Education on High School Students and their Teachers: Experimental Evidence from Peru.", IDB working papers series N° IDB-WP-871
- Ganzach, Y. (2000). "Judging risk and return of financial assets". *Organizational Behavior and Human Decision Processes*, 83(2), 353-370.

Gill, A. Bhattacharya, R. (2015). "Integration of a financial literacy curriculum in a high school economics class: Implications of varying the input mix from an experiment". *Journal of Consumer Affairs*, 49 (2): 472–487.

Gigerenzer, G. (2015). "Risk savvy: How to make good decisions". Penguin.

Hambleton, R. K., van der Linden, W. J. (1997). *Handbook of modern item response theory*. New York: Springer-Verlag.

Hospido, L., Villanueva, E., Zamarro, G. (2015). "Finance for all: the impact of financial literacy training in compulsory secondary education in Spain.", *Documentos de Trabajo*, n. 1502, Banca de Espana.

Kaiser, T.; Menkhoff, L.. (2017). "Does financial education impact financial literacy and financial behavior, and if so, when ?", *The World Bank Economic Review*, 31(3), 2017, 611–630.

Kahneman, D. (2011). "Thinking, fast and slow". Macmillan.

Koundinya, V., Klink, J., Deming, P., Meyers, A., Erb, K. (2016). "How Do Mode and Timing of Follow-up Surveys Affect Evaluation Success?". *Journal of Extension*, 54(1), n1.

Linciano, N., Gentile, M., Soccorso, P. (2016). "Report on Financial Investments of Italian Households. Behavioural Attitudes and Approaches-2016 Survey".

Lührmann, M., Serra-Garcia M., Winter J. (2015). "Teaching teenagers in finance: Does it work?", *Journal of Banking & Finance*, 54 (2015) 160–174.

Lusardi, A., Mitchell, O. S., Curto, V. (2010). "Financial literacy among the young". *Journal of consumer affairs*, 44(2), 358-380.

Lusardi, A., O.S. Mitchell. (2014). "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature*, 52 (1): 5–44.

Machado, J. A. F., Parente, P. M., Silva, J. S. (2015). qreg2: Stata module to perform quantile regression with robust and clustered standard errors. *Statistical Software Components*.

McKenzie, D. (2012). "Beyond Baseline and Follow-up: The Case for More T in Experiments." *Journal of Development Economics* 99 (2): 210–21.

Meier, S., Sprenger C. (2007). "Selection into Financial Literacy Programs: Evidence from a Field Study". Federal Reserve Bank of Boston Discussion Paper 07–5.

Miller, M., J. Reichelstein, C. Salas, B. Zia. (2015). "Can You Help Someone Become Financially Capable? A Meta-Analysis of the Literature." *World Bank Research Observer*, 30 (2): 220–46.

Puri, Manju, Robinson, David T., (2007). "Optimism and economic choice," *Journal of Financial Economics*, vol. 86(1), pages 71-99, October

Romagnoli, A., Trifilidis, M. (2013). "Does financial education at school work? Evidence from Italy" (No. 155). Bank of Italy, Economic Research and International Relations Area.

Walstad, W.B., Rebeck K., MacDonald R.A. (2010). "The Effects of Financial Education on the Financial Knowledge of High School Students." *Journal of Consumer Affairs*, 44 (2):336–57.

Willis, L.E., (2011). “The Financial Education Fallacy.” *American Economic Review: Papers and Proceedings*. 101 (3):429–34.

TABLES AND FIGURES

	Controls	Treated	Total
Vocational	65	176	241
Technical	60	76	136
Liceum	34	114	148
Total	159	366	525

Class	Controls	Treated	Total
I	34	23	57
II	25	69	94
III	42	72	114
IV	34	122	156
V	24	80	104
Total	159	366	525

Table 3. Respose rates by question, test phase and treatment condition

Table 3. Respose rates by question, test phase and treatment condition							
	Controls			Treated			
Question	Phase 0	Phase 1	Diff Phase1-Phase 0 C (a)	Phase 0	Phase 1	Diff Phase1-Phase 0 T (b)	Diff in diff. (a)-(b)
1	0.24	0.25	0.01	0.17	0.56	0.39	0.39
2	0.17	0.26	0.09	0.11	0.57	0.46	0.38
3	0.33	0.51	0.18	0.35	0.78	0.44	0.26
4	0.88	0.94	0.06	0.95	0.93	-0.02	-0.08
5	0.63	0.73	0.10	0.69	0.82	0.13	0.03
6	0.44	0.41	-0.03	0.42	0.60	0.18	0.21
7	0.36	0.42	0.06	0.33	0.31	-0.02	-0.08
8	0.53	0.58	0.05	0.63	0.82	0.19	0.14
9	0.24	0.39	0.15	0.29	0.47	0.18	0.03
10	0.73	0.79	0.06	0.79	0.90	0.11	0.06
11	0.47	0.50	0.04	0.54	0.70	0.16	0.12
12	0.48	0.50	0.01	0.41	0.67	0.26	0.25
13	0.50	0.51	0.01	0.54	0.59	0.05	0.05
14	0.46	0.28	-0.18	0.41	0.37	-0.04	0.14
15	0.28	0.28	0.01	0.27	0.31	0.05	0.04
16	0.86	0.84	-0.03	0.87	0.89	0.02	0.04
17	0.84	0.75	-0.09	0.85	0.87	0.02	0.10
18	0.35	0.36	0.01	0.38	0.53	0.15	0.14
19	0.65	0.58	-0.07	0.78	0.81	0.02	0.09
20	0.29	0.27	-0.02	0.28	0.39	0.11	0.13
21	0.10	0.10	0.00	0.11	0.11	0.01	0.01
22	0.45	0.53	0.08	0.45	0.63	0.18	0.10
23	0.66	0.66	0.00	0.71	0.81	0.10	0.10
24	0.42	0.55	0.14	0.55	0.69	0.14	0.01
25	0.32	0.40	0.08	0.36	0.48	0.11	0.04
26	0.47	0.51	0.04	0.62	0.69	0.07	0.03
27	0.30	0.34	0.04	0.25	0.36	0.11	0.08
28	0.64	0.69	0.05	0.77	0.80	0.03	-0.02
29	0.33	0.33	0.00	0.33	0.48	0.14	0.14
30	0.64	0.56	-0.08	0.71	0.72	0.01	0.09
Number of students	159	159		366	366		

Sample	Score type and phase	N. of students	Mean	p50	Standard deviation	p5	p95	Min	Max	
C	Phase 0 raw score (a)	159	0.47	0.47	0.12	0.27	0.63	0.10	0.73	
	Phase 1 raw score	159	0.49	0.50	0.15	0.27	0.70	0.13	0.97	
	Phase 0 estimated score (c)	159	-0.21	-0.17	0.79	-1.45	1.05	-2.56	1.89	
	Phase 1 estimated score	159	-0.44	-0.51	0.83	-1.62	0.73	-2.19	1.97	
T	Phase 0 raw score (b)	366	0.50	0.50	0.12	0.30	0.70	0.20	0.93	
	Phase 1 raw score	366	0.62	0.63	0.14	0.40	0.83	0.17	0.97	
	Phase 0 estimated score (d)	366	0.09	0.10	0.83	-1.31	1.39	-1.90	2.29	
	Phase 1 estimated score	366	0.19	0.21	0.83	-1.21	1.58	-1.99	2.11	
	Diff. (b)-(a) (2)	525	.028 (.027)	.033 (.028)						
	Diff. (d)-(c) (2)	525	.301 (.196)	.261 (.212)						

(1)-The table reports descriptives on raw and estimated scores achieved by students in phase 0 and 1 of the test. Scores are defined by the share of correct answers out of the 30 questions included in the questionnaire (2) Mean and median differences across samples in phase 0 are tested by regressing phase 0 scores on the treatment dummy. For the test on mean difference an OLS method is used, for that on medians we use the qreg2 procedure in stata. Clustered standard errors at class\section level are in parenthesis * denote that the estimated parameter is significant at 10 per cent, ** at 5 per cent, *** at 1 cent..(

Explanatory variables	Mean	Standard deviation	P value (2)	N. observations
Student' self assesment (3)	3.54	1.39	0.597	525
Evaluation of a risky action (3)	6.08	1.55	0.215	525
Propensity to risky actions (3)	2.68	1.84	0.732	525
Grades in math	6.59	1.24	0.286	265
Gender	0.51	0.50	0.419	265
Nationality	0.88	0.33	0.702	265
Proof length (in minutes)	52.93	10.84	0.637	525
Time interval elapsed between Phase 0 and Phase1 of the test (in days)	87.74	19.50	0.574	525

(1)-The table reports descriptives and testing on some explanatory variables used in the regression. (2) The p values represent significance levels obtained by regressing via OLS each explanatory variable on the treatment dummy. Standard errors are clustered at class\section level. Student self-assessment, evaluation of risky actions and the propensity to risky actions are all measured through a scale going from 1 (the minum level) up to 7 (the maximum level). These variables are assumed to be continuous. For testing the differences on the nationality and gender we resort to a linear probability model.

	I	II	III	IV	V	VI
	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	.116***	0.181***	1.009***	0.196***	.1901***	0.100***
	(.025)	(0.01)	(0.11)	(0.01)	(.008)	(0.02)
N	525	525	525	265	234	1050
R2 (adjusted)	0.250	0.496	0.478	0.513	0.627	0.322

(1) All the OLS regressions are run on the sample of students participating both in phase 0 and 1. Unless otherwise specified, dependent variable is defined by the score, ie the share of correct answers out of the 30 questions included in the questionnaire, obtained by each student in the follow up (phase 1) survey,. Apart from the treatment dummy, signaling that the student attended a course related to the Bank of Italy financial education program,, specification from I to V include baseline (phase 0) score among the controls. Specifications from II to V add to the set of controls teacher fixed effects, school type and class attended, test characteristics, the student's self-evaluation about his own financial knowledge, risky situations and risk propensity (for the definition of the latter two variables see Appendix 2). Clustered standard errors at class section level are in parenthesis * denote that the estimated parameter is significant at 10 per cent, ** at 5 per cent, *** at 1 cent. (2) Explanatory variables include the treatment dummy and the pretreatment score only. (3) Baseline specification. (4) Baseline specification, dependent variable is defined by the estimated score (see in the text). (5) The set of controls include those in the baseline specification and students' grades in math obtained in the 2015-16 school year, gender and nationality. (6) The sample is restricted to those teachers participating with their classes to both the treatment and control sample (7) Panel data specification, see equation (2) in the text, it includes student fixed effects. The estimated parameter refer to interaction terms of POST and the treatment dummies.

	5 Perc.	10 Perc.	25 Perc.	50 Perc.	75 Perc.	90 Perc.	95 Perc.
Treatment	0.103***	0.186***	0.191***	0.200***	0.178***	0.146***	0.153***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N	525	525	525	525	525	525	525

(1) All the quantile regressions are run on the sample of students participating both in phase 0 and 1. The dependent variable is defined by the score, ie the share of correct answers out of the 30 questions included in the questionnaire, obtained by each student in the follow up (phase 1) survey,. Apart from the treatment dummy, signaling that the student attended a course related to the Bank of Italy financial education program,, specification include among the controls the baseline (phase 0) score, teacher fixed effects, school type and class attended, proof characteristics, the student's self-evaluation about his own financial knowledge, risky situations and risk propensity (for the definition of the latter two variables see Appendix 2). Clustered standard errors at class section level are in parenthesis * denote that the estimated parameter is significant at 10 per cent, ** at 5 per cent, *** at 1 cent. Estimations are carried out at the 5, 10, 25, 50, 75, 90 and 95 th percentile of the distribution of the dependent variable.

	I	II	III	IV
	(2)	(3)	(4)	(5)
Treatment	0.052*	0.097***	0.083***	0.090***
	(0.02)	(0.02)	(0.02)	(0.02)
N	525	525	265	234
R2 (adjusted)	0.058	0.254	0.212	0.251

(1) All the OLS regressions are run on the sample of students participating both in phase 0 and 1. Dependent variable is defined by the score, ie the share of correct answers out of the 8 selected questions included in the questionnaire and that are less related to the argument taught during the courses, obtained by each student in the follow up (phase 1) survey,. Apart from the treatment dummy, signaling that the student attended a course related to the Bank of Italy financial education program,, specification from I to IV include baseline (phase 0) score among the controls. Specifications from II to IV add to the set of controls teacher fixed effects, school type and class attended, proof characteristics, the student's self-evaluation about his own financial knowledge, risky situations and risk propensity (for the definition of the latter two variables see Appendix 2). Clustered standard errors at class section level are in parenthesis * denote that the estimated parameter is significant at 10 per cent, ** at 5 per cent, *** at 1 cent. (2). Explanatory variables include the treatment dummy and the pretreatment score only. (3) Baseline specification. (4) The set of controls include those in the baseline specification and students' grades in math obtained in the 2015-16 school year, gender and nationality. (5) The sample is restricted to those teachers participating with their classes to both the treatment and control sample.

Figure 1. Raw score distribution in phase 0 by treatment condition

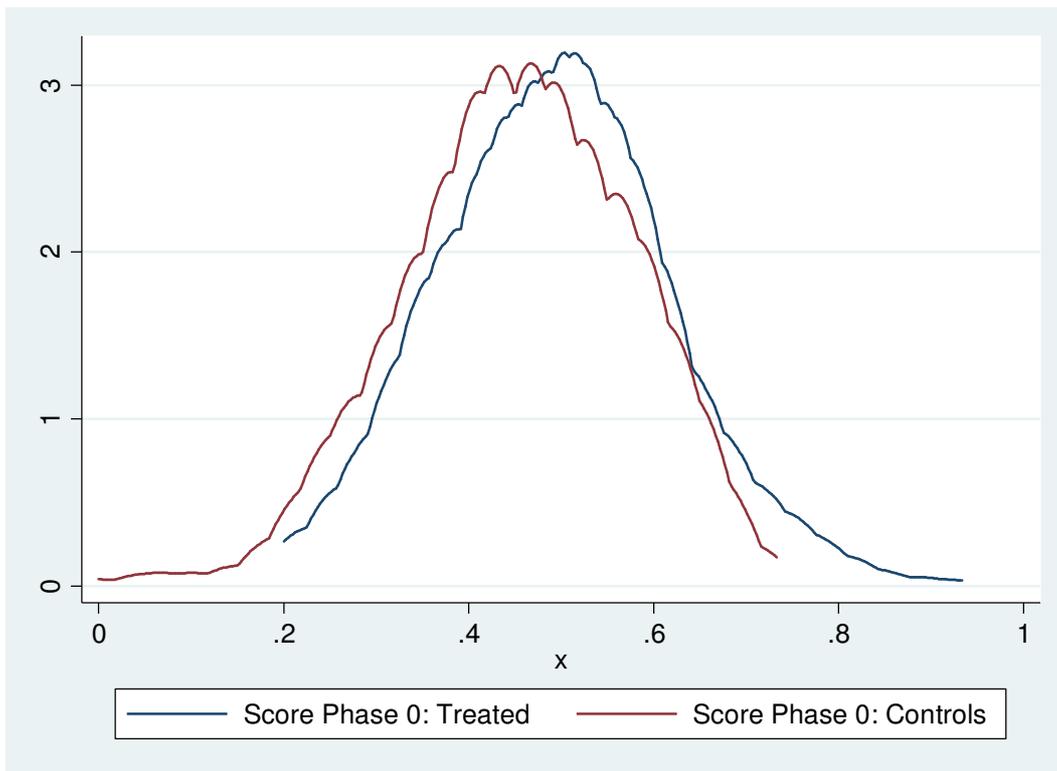
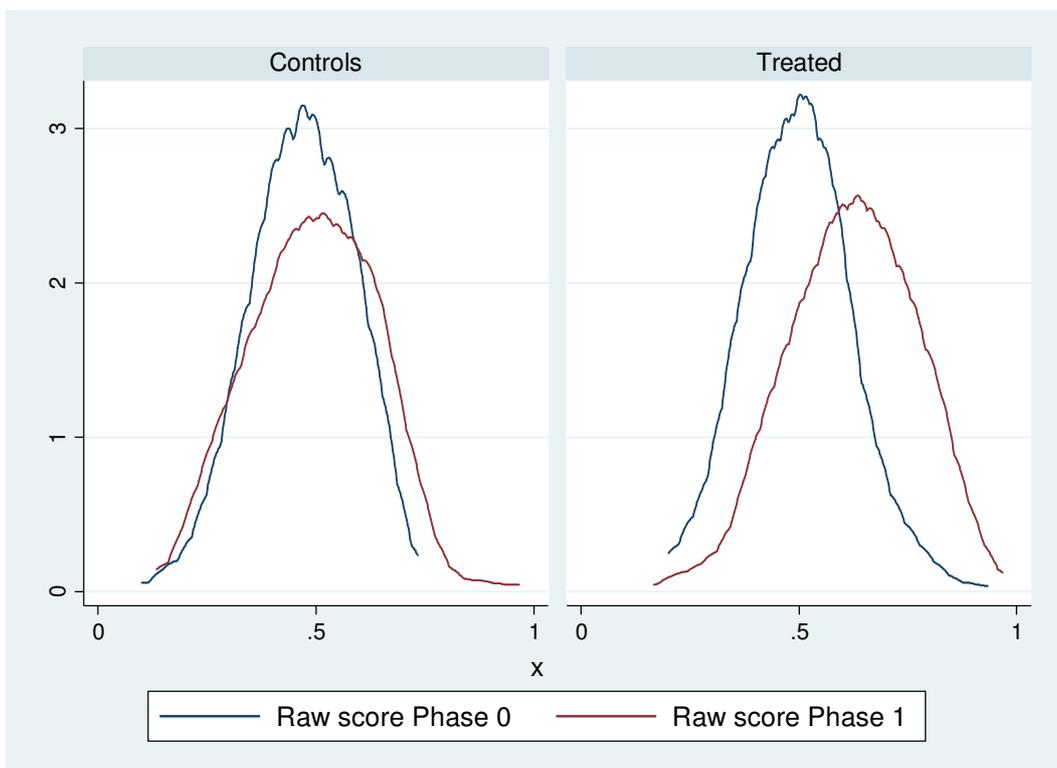


Figure 2. Raw score distribution by treatment condition and phase



APPENDIX 1

Questionnaire to evaluate competency in economic and financial subjects

(For each question, there is always only one possible correct answer. Giving more than one answer or failing to provide an answer are not, therefore, considered valid responses)

Section 1 Money

1. Legal tender is characterised by its:
 - a) Traceability
 - b) Validity as a payment instrument to discharge a debt or meet a financial obligation
 - c) Convertibility
 - d) I don't know
2. If you wish to make a purchase using coins, what is the maximum number of coins that a vendor is required to accept?
 - a) There is no limit
 - b) One hundred
 - c) Fifty
 - d) I don't know
3. Up to which amount is it possible to make a payment in cash?
 - a) 3,000 euros
 - b) 1,500 euros
 - c) 5,000 euros
 - d) I don't know

Section 2 Current accounts and payment instruments

4. What is a bank current account?
 - a) A contractual relationship with a bank in which only credit transactions are recorded
 - b) A contractual relationship with a bank in which only debit transactions are recorded
 - c) A contractual relationship with a bank in which financial transactions are recorded
 - d) I don't know
5. A deposit in a current account of 1,500 euros which earns annual interest, net of withholding tax, at a rate of 5 percent, after 12 months will be worth:
 - a) 1,575 euros
 - b) 1,555 euros
 - c) 1,585 euros
 - d) I don't know

6. A debit card allows the holder to make payments:
 - a) Only if he or she has funds available in his or her bank account at the time of the transaction
 - b) Even if he or she does not have funds available in his or her bank account at the time of the transaction
 - c) Credited to the current account the following month
 - d) I don't know
7. You have deposited 5,000 in a current account which pays annual interest at a rate of 20% which is redeposited in the account without being spent. Supposing that there are no taxes to pay, two years after depositing the initial sum in your current account, the total amount will be:
 - a) Less than 7,000 euros
 - b) Equal to 7,000 euros
 - c) More than 7,000 euros
 - d) I don't know
8. You need to open a current account at a bank and you know that you will need to perform 50 transactions in a year. Indicate which of the three contracts proposed offers the lowest cost per transaction:
 - a) Fixed fee of 20 euros to open the account and a fee of 50 cents for each transaction
 - b) No fee to open the account, the first 20 transactions free of charge and the remaining transactions at 2 euros each
 - c) No fee to open the account, the first 10 transactions free of charge and the remaining transactions at 1 euro each
 - d) I don't know
9. You have a current account at a bank through which all your income and expenditure is transacted. On 31 December 2015, the account had a positive balance of 2,000 euros. A year later, the balance stood at 4,000 euros. Which of the following statements is definitely false?
 - a) Between 2015 and 2016, income grew by 3,000 euros and expenditure decreased by 1,000 euros
 - b) In 2016, expenditure stood equal to half of income
 - c) Between 2015 and 2016 income grew more than expenditure
 - d) I don't know

Section 3: Inflation

10. The real value of money is:
 - a) The currency exchange rate
 - b) The purchasing power of a currency
 - c) The cost of holding currency
 - d) I don't know

11. Does uncertainty over the future value of money have an impact on the distribution of wealth among citizens?
- a) No, it does not have an impact on the fairness of wealth distribution
 - b) Yes, it increases the fairness of wealth distribution
 - c) Yes, it decreases the fairness of wealth distribution
 - d) I don't know
12. What categories of people are penalised most in periods of high inflation?
- a) Employees
 - b) Freelance professionals
 - c) No particular category
 - d) I don't know
13. If in 2016 there was a reduction of 4 percent in general price levels while in 2015 there was a reduction of 2 percent, it can be correctly stated that:
- a) Deflation levels have reduced in intensity
 - b) Deflation has intensified
 - c) The price level in the most recent year is half that of the preceding year
 - d) I don't know
14. Supposing that in 2016 your income doubled compared with that of 2015 and the prices of all the goods and services grew by 50 percent. Compared with 2015, the income you received in 2016 will enable you to buy a quantity of goods and services that is:
- a) Greater
 - b) The same
 - c) Lesser
 - d) I don't know
15. Imagine that you will be paid a sum of money in a year. Indicate which sum you would prefer out of the three scenarios proposed:
- a) 1,000 euros in a scenario where the expected annual increase in all prices for the next year is 5 percent
 - b) 1,200 euros in a scenario where the expected annual increase in all prices for the next year is 30 percent
 - c) 800 euros in a scenario where the expected annual increase in all prices for the next year is 1 percent
 - d) I don't know

Section 4: Financial system and types of financial activity

16. What does a bank do?
- a) Collects, holds and lends money
 - b) Collects, holds and prints money

- c) Collects, holds, lends and prints money
- d) I don't know

17. What is a bank loan?

- a) It is a sum of money that the customer lends to the bank which must be returned to the customer with interest
- b) It is a sum of money that the bank lends to the customer which must be returned to the bank with interest
- c) It is a sum of money that the bank lends to the customer who declares that he or she will return it to the bank
- d) I don't know

18. What is the main purpose of the financial system (banks, mutual investment funds, stock market)?

- a) To protect the rights of savers
- b) To reduce social inequalities
- c) To put persons with revenues greater than their expenditure in contact with those who find themselves in the opposite situation (their expenditure is greater than their revenues)
- d) I don't know

19. Which of the following forms of investment allows you to most easily access your cash in the event you suddenly need it?

- a) Current account deposit
- b) Government securities
- c) Stocks and shares
- d) I don't know

20. What happens to the value of a security when interest rates decrease?

- a) It remains the same
- b) It increases
- c) It decreases
- d) I don't know

21. You wish to purchase a commodity with a value of 1,000 euros. You do not currently have the funds available and so decide to apply for a loan. Which of the following three proposals do you think would be most beneficial to you?

- a) Apply for a loan with a one-year repayment period at the end of which the debtor is required to pay back a total of 1,200 euros
- b) Pay 12 monthly instalments of 100 euros
- c) The solutions proposed in a) and b) are the same
- d) I don't know

Section 5: Risk diversification

22. When a saver divides his or her wealth among different financial assets, the risk of losing money:
- a) Increases
 - b) Decreases
 - c) Remains the same
 - d) I don't know
23. Indicate which of these categories of financial assets is usually considered as subject to greater risk:
- a) Stocks and shares
 - b) Bank deposits
 - c) Government securities
 - d) I don't know
24. When a financial asset is considered 'high risk', this means that:
- a) The return on said asset may be considerably lower than expected
 - b) The return on the asset is very low
 - c) The return on the asset may be either much lower or much higher than expected
 - d) I don't know
25. Out of the following strategies, which do you think is the most effective for containing risk in a financial investment?
- a) Do not invest in high-risk products. Invest only in government bonds and/or securities
 - b) Invest in many financial assets with little correlation between the rates of return on each
 - c) Invest in as many financial assets as possible
 - d) I don't know
26. You are to inherit ownership of some companies which is to be divided between you and other heirs. Each heir is entitled to ownership of three different companies with a current total value of 100,000 euros. You will receive the amount deriving from the sale of the three companies in one year. The revenue from the sale of the company in a year will depend on trends in the sector in which it operates and could also be reduced to zero in the event of any serious crisis in said sector (although it could also exceed 100,000 euros in the event that business is particularly good over the year). You are offered three alternatives. Which of the three do you think reduces to a minimum the possibility of losing all or a large part of the value of your inheritance?
- a) A company that produces raincoats, one that produces swimsuits and one that produces sun cream
 - b) A company that produces umbrellas, one that produces sunglasses and one that runs a cinema with an outdoor and an indoor screen
 - c) A company that produces umbrellas, one that produces raincoats and one that produces swimsuits
 - d) I don't know

27. You have purchased two securities, A and B, each with a value of 100 euros at the start of the year. Over the year you will receive 10 euros in interest on A and 20 euros in interest on B. You resell the two securities after a year earning 115 from the sale of A and 105 from the sale of B. Which of the two securities yielded the highest return?
- a) Security A
 - b) Security B
 - c) The return on securities A and B is the same
 - d) I don't know

Section 6: Insurance

28. What is the purpose of insurance?
- a) To eliminate the impact of adverse events
 - b) To limit the impact of adverse events
 - c) To limit adverse events
 - d) I don't know
29. What is the deductible on a civil liability car insurance policy?
- a) The contract clause establishing the amount that the policy holder must contribute when making a compensation claim
 - b) The contract clause establishing the maximum amount of a compensation claim
 - c) The contract clause that limits or excludes the cover of risks
 - d) I don't know
30. What should you do before choosing a civil liability car insurance policy?
- a) Compare the insurance premiums of various companies
 - b) Compare the policy terms and conditions of various companies
 - c) Compare both the premiums and the terms and conditions of various companies
 - d) I don't know

APPENDIX 2

Legend of the main variables
Dependent variable:
y_{st} is the student s raw or estimated score in phase t ($t=0,1$). Raw scores are defined by the share of correct answers out of the 30 questions included in the questionnaire. The estimated scores, proxing for students' unobserved abilities, are obtained using the Item Response Theory approach, typically used in educational research.
Explanatory variables:
TR_s treatment dummy= 1 if s belongs to T , with $b \in \{T,C\}$ and 0 otherwise
$I_{ds}=1$ Dummy =1 if the student s belongs to a class of teacher d and 0 otherwise
SC vector: 3 dummies for school types : Vocational (omitted dummy), technical, lyceum 5 dummies for the classes attended: first year(omitted) ,second, third, fourth and fifth year
PR vector: LENGTH: the length of the test (maximum amount of time assigned to fill in the questionnaire) expressed in minutes INT: the time interval elapsed between Phases 0 and 1, expressed in number of days TP: dummy = 1 if the modality of compilation of the test was on computer and 0 otherwise.
Vector ST: FL: 7 dummies measuring students' self-evaluation about their own financial literacy (=1 corresponding to null (omitted) to= 7 "very high knowledge and competence") RSK1: 7 dummies measuring the evaluation of the degree of risk associated to spending weekly income at the casino (=1 no risk (omitted) = 7 maximum risk) RSK2: 7 dummies measuring the propensity to undertake the action in RSK1 (=1 low or null propensity (omitted) to = 7 maximum propensity to follow that course of action). MTH: Grades in math in the school year 2015-16 (from 1 (the minum to 10 (the maximum)) SEX: Dummy =1 if the student is female NT: Dummy = 1 if the student is Italian