Martins, Pedro Silva

Working Paper

Can targeted, non-cognitive skills programs improve achievement? Evidence from EPIS

IZA Discussion Papers, No. 5266

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Martins, Pedro Silva (2010) : Can targeted, non-cognitive skills programs improve achievement? Evidence from EPIS, IZA Discussion Papers, No. 5266, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/46045

DISCUSSION PAPER SERIES

IZA DP No. 5266

Can Targeted, Non-Cognitive Skills Programs Improve Achievement? Evidence from EPIS

Pedro S. Martins

October 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Can Targeted, Non-Cognitive Skills Programs Improve Achievement? Evidence from EPIS

Pedro S. Martins
Queen Mary University of London,
CEG-IST and IZA

Discussion Paper No. 5266
October 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5266
October 2010

ABSTRACT

Can Targeted, Non-Cognitive Skills Programs Improve Achievement? Evidence from EPIS*

EPIS is an original and large private-sector program aimed at improving student achievement and eroding early school leaving at Portuguese state schools. The program first screens students to focus only on those more likely to perform poorly; and then conducts a number of small-group sessions aimed at improving the non-cognitive skills (e.g. study skills, motivation, self-esteem) of the selected students. Our quasi-experimental evidence of the effects of EPIS is drawn from rich longitudinal student data and the different timings in the roll-out of the program, both within and across schools. The results indicate that the program reduced grade retention by at least 10 percentage points and did so in a cost effective way.

JEL Classification: I20, J08

Keywords: student achievement, program evaluation, matched school-student data

Corresponding author:

Pedro Martins
School of Business and Management
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom
E-mail: p.martins@qmul.ac.uk

* I thank Mário Centeno and Diogo Simões Pereira for many discussions and Diogo Rodrigues for information about data. I also thank comments from José Manuel Canavarro, Francis Green, Bernadette Kamleitner, José Ferreira Machado, Eduardo Marçal Grilo, Helena Martins, Alfonso Miranda, Júlio Pedrosa and conference/seminar participants at Banco de Portugal, University of Minho and Institute of Education (University of London). The views and results presented here are my own and not necessarily those of EPIS. Any errors are also of my responsibility only.

1 Introduction

Investment in human capital is widely regarded as a critical avenue towards higher levels of economic wellbeing (Krueger & Lindahl 2001, Heckman & Carneiro 2003). However, substantial skill gaps are still documented around the world, including in developed countries. Moreover, such skill gaps (e.g. schooling attainments of not more than lower secondary education) correlate reasonably strongly with children's socio-economic or minority status and most likely impair intergenerational mobility and the promotion of equal opportunities.

Governments have responded to these skill gaps by launching a number of programs that seek to enhance the schooling achievement of disadvantaged pupils. Most of these programs involve some form of remedial education, where students that perform poorly receive extra tuition, revision sessions, computer-aided learning, etc. However, the causal effects of many of these programs have been shown to be small (Machin et al. 2004), mixed (Jacob & Lefgren 2004) or even insignficant (Leuven et al. 2007). In some other cases, the effects are relatively sizeable (Lavy & Schlosser 2005) but not as cost effective as alternatives based on more systemic changes to the education system (e.g. teacher incentives). Similar findings have been obtained for the case of developing countries (Glewwe & Kremer 2006), where most students would be regarded as disadvantaged by developed-country standards, despite notable exceptions (Banerjee et al. 2007).

While most remedial programs focus on cognitive skills (e.g. revising class room material), economists have recently began devoting attention to the non-cognitive dimensions of the behaviour of children and youngsters (Heckman & Rubinstein 2001, Heckman et al. 2006). Such non-cognitive traits involve aspects of one's personality that may be more difficult to define (at least for an economist) but that may be equally important in terms of shaping one's success at school and after: motivation, discipline, tenacity, self-esteem, self-control, confidence, patience, etc. In this context, one wonders if school interventions based on non-cognitive skills may be more effective than the more traditional remedial programs, as the ones listed above.

This paper addresses this issue by presenting empirical evidence about the effects of EPIS, an original program that seeks to improve achievement and reduce retention and early school leaving of lower-secondary students by strengthening their non-cognitive skills. This program is in operation in Portuguese state schools since 2007 and is run by a private, not-for-profit

organisation. On top of the emphasis placed on non-cognitive skills, EPIS is original due to its screening of all the more than 15,000 students in each treated cohort in the 85 schools that participate in the program. Such screening ensures that the resources invested by the program - mostly the program staff time, typically devoted to small-group or one-to-one sessions with students - are spent only on the youngsters that are likely to need the intervention the most. Indeed, although many education programs have been shown to have small or insignificant effects, 'inputs specifically targeted to helping weaker students may be effective', in particular 'if they address specific unmet needs in the school' (Banerjee et al. 2007).

However, such screening raises obvious issues in terms of the identification of the causal effects of the intervention. We address such concerns by drawing on different features of EPIS: particularly rich longitudinal data on all students (treated and non-treated) - which also allow us to present a number of interesting results about the determinants of student achievement (Lazear 2003, Rivkin et al. 2005, Aaronson et al. 2007, Lavy 2009, Martins & Walker 2006); and differences in the roll-out of the program across and within schools. The second feature includes a number of schools that were not treated in the first two years of the program but for which the same detailed information on student characteristics was also collected and made available for research throughout the entire period.

Our evidence indicates that, unlike many remedial programs, EPIS had a significant, positive effect in terms of improving the achievement levels of treated students. The probability that a student fails 25% or more of their modules (our main measure of achievement, which typically entails grade retention when it occurs in the third and last term of the school year) falls by at least 10 percentage points (and as much as 30 percentage points in some specifications).

However, we also find that the effects of the program on specific modules that may be presumed to be more intensive in cognitive skills (e.g. maths) tend to be smaller, even if still significant. This result may highlight the limitations of an approach based almost entirely on non-cognitive skills; it may also reconcile our results with those of the only other case-study of non-cognitive skills we are aware of (Holmlund & Silva 2009), which find insignificant results. In any case, the effects of the EPIS program on retention appear to be cost effective.

The structure of the paper is as follows: Section 2 describes the main characteristics of the program studied in the paper and some information on the Portuguese education system;

Section 3 presents the data used in the paper, a matched school-student panel data set, and some descriptive statistics; Section 4 describes the main results; while Section 5 presents several extensions and robustness analyses. Finally, Section 6 concludes.

2 The EPIS program

Several countries exhibit high levels of early school leaving. The Portuguese case is particular conspicuous in this respect. According to Eurostat figures for 2007, 39.1% of 18-24 year olds have at most a lower secondary school degree and are not enrolled in any training. This is in huge contrast to the equivalent figure for the 27-country European Union, which is 15.5% (see Figure 1). This is particularly noteworthy given the persistently high wage differentials between schooling levels (Martins & Pereira 2004).

On top of the above, many in this 18-24 cohort will not have completed the lower secondary school level (school leaving age is 15). In fact, retention and school leaving rates in state schools are about 15% in each of the three years of the lower secondary cycle: such rates are 17.9% (7th grade), 11.6% (8th grade) and 14.2% (9th grade), according to Education Department reports for the 2007/08 school year.[1]

The EPIS program seeks to address some of the skill gaps that inevitably follow from the statistics above, in particular the low student achievement and eventual dropping out of many youngsters attending lower secondary school in Portuguese state schools. These goals are pursued through the adoption of an original approach in at least two different respects. Most importantly, EPIS aims to strengthen the non-cognitive skills of students in opposition to a more standard remedial approach, based on cognitive skills. As mentioned before, there is growing awareness about the potential effects of such non-cognitive skills - motivation, discipline, tenacity, self-esteem, self-control, confidence, patience, etc - in terms of different socio-economic outcomes (Heckman & Rubinstein 2001, Heckman et al. 2006).

A second novel dimension of the EPIS program concerns its adoption of a targeted intervention. Specifically, EPIS spends considerable effort in identifying the 7th- and 8th-grade pupils (typically 13-15 year-olds) most at risk of failing their year and/or dropping out. This selection is achieved through at least two rounds of screening, when students are interviewed

[1] While in part this *status quo* of low achievement follows from the low levels of schooling in the overall population, many in the country regard the low attainment of the young as an important barrier towards economic growth. This is likely to be particularly true as globalisation prompted the relocation of many low-wage industries from Portugal to developing countries.

individually by EPIS staff, and also a number of meetings with teachers, headteachers and parents for further information. A detailed questionnaire is conducted for each student, which leads to the assignment of the student to one of three different levels of concern with respect to the students' likelihood of poor achievement. Students that hit the highest level of concern are then selected into the program (on average about one third of all 7th and 8th grade students in each school). Students that hit intermediate levels of concern are referred to their teachers for further monitoring. The remaining students are disregarded by the program, except that their grades are also recorded, up to their graduation from the 9th grade or until they leave school.

Once the students' parents agree on the participation of their children (approximately 95% do), then a specific set of intervention methods is designed for each student. These methods will depend on the individual non-cognitive, behavioral or other issues that are flagged during the screening process. The intervention will therefore implement one or more different approaches, namely individual techniques (motivational discussions, self-control, problem-solving techniques), and group techniques (study methods, social competences training, management of criticism, anxiety self-control).

These one-on-one or small-group interventions are implemented by EPIS staff (*mediadores*) that work full-time in the program. They are based permanently at schools (except for training spells) in offices made available there. Many staff are recent graduates in psychology or education sciences; in the case of three school districts, the program is delivered by tenured teachers that do not have teaching duties (*horários zero*) at the schools to which they are attached on account of demographic or other reasons. EPIS staff also seek to keep in frequent contact with the relatives and teachers of the participant students and, in some cases, their local councils, to monitor the progress of each intervention pupil as closely as possible. EPIS staff not only meet their students individually or in very small groups but they also do so on a relatively frequent basis (e.g., every two weeks). During the first meetings, staff and students agree on goals with respect to each student's future performance. Meetings typically do not overlap with normal classes. This rules out negative effects on class attendance and also minimizes any stigma that may occur if students were asked to leave a class to attend an EPIS session as that could lead to identification by peers.[2]

[2]On the other hand, some anecdotal reports indicate that students regard it as 'cool' to be a participant in the program. See Pereira et al. (2008) for detailed information about the program.

Furthermore, the program places emphasis on the stability in the staff-student relationship, as that is regarded to strengthen the effect of the treatment. There are currently 65 EPIS staff working in schools and each holds a portfolio of a total of 50 to 100 students in one or two schools. All staff take part in several training sessions before they are allocated to schools and during the program.

On an administrative level, the EPIS program is run by a private, not-for-profit organisation of the same name (EPIS stands for 'Entrepreneurs for Social Inclusion'). Its funding comes from approximately 90 of the largest companies based in Portugal, each paying an annual contribution of 25,000 euros since 2006. (A small share of these funds are spent on additional, related initiatives, including a program for management skills for headteachers and the promotion of entrepreneurial skills in students.) In addition, a scientific council formed by education experts from academia helped in the design of the program and meets regularly to offer comments and suggestions to senior EPIS staff.[3]

The Education Department and the school districts where the interventions are taking place also offer additional, generally non-financial, support, namely in terms of facilities at schools and coordination with headteachers (and program staff in a small number of cases). EPIS also received the public endorsement of the Portuguese head of state and several newspapers and TV stations agreed to advertise the program on a *pro bono* basis, to raise the public awareness and profile of the program, thus facilitating its implementation in schools, in terms of its acceptance by teachers, parents and students.[4]

The program was implemented from the 2007/08 school year in all the lower-secondary schools in ten different school districts ('concelhos') across the country, resulting in a total of 85 schools and over 15,000 7th and 8th grade students. These correspond to almost 10% of all 7th and 8th grade students in the entire country. Although this is already a sizable coverage, the program can be scaled up relatively quickly, given that the set-up investments have already been made and variable costs stem mostly from the wage bill for staff based at schools.

[3]The author of this paper has been an unpaid member of this scientific council since its inception, in 2006.

[4]See `http://www.epis.pt` (in Portuguese) for more detail about the program and its initiatives, including links to `YouTube` statements from students that participate in the program.

3 Data

The EPIS program adopted a sophisticated IT system that keeps detailed records of students. In particular, there are individual, longitudinal records about all students from each school that takes part in the program, including those that are not assigned to the treatment group, provided they are enrolled in the 7th or 8th grade in the 2007/08 school year. Most information, including the students' grades, is recorded at the term (quarter) frequency (there are three terms per school year: September to December, January to March and April to June). The data include several demographic variables about each student, such as gender and age, and different profile characteristics, in particular those perceived to be proxies of socio-economic or psychological issues that may affect progression. Interviews with teachers result in additional information from previous school years of each student, including earlier spells of retention.

All information is longitudinal, which creates a matched multilevel panel, in which the levels concern the student, the class (group of students that take the modules together), the EPIS staff and the school: there are unique and time-invariant identifiers for each unit in each one of those four levels. Crucially, the EPIS data also include information about the timings of the treatment for each treatment-group participant.[5]

3.1 Descriptive statistics

As mentioned above, the program began in 2007/08 with 7th and 8th grade students. In 2008/09 the first cohort of students were still under analysis and/or treatment but they then spanned the 7th grade (students retained), 8th grade (students that progressed from the 7th grade and students retained in the 8th grade) and the 9th grade (students that progressed from the 8th grade). Finally, in the current 2009/10 school year, most students correspond to those enrolled in the 7th grade in the first year of the program, which will now be attending the 9th grade in most cases. While the 2007/08 and 2008/09 data concern three terms (1st, 2nd and 3rd), the 2009/10 data used here concerns only the 1st term, creating a total of seven terms.

Given its richness, the data can be described in different ways. Table 1 focuses on student-quarter pooled data, of which there are a total of 93,901 observations, corresponding to a total

[5]The data set used in this paper is available upon payment of a small fee from `http://www.epis.pt`.

of 15,307 different students.[6] The mean age is 14.5 and 49.7% are girls. 4.5% are enrolled in lighter, vocational programs ('CEF'). 30.8% of the observations concern students that are (eventually) treated by the program (i.e. observations either before or during treatment of treated students), while 4.8% of the observations concern students that begin their treatment in that quarter. 18.7% student-quarter observations correspond to treated students observed once their treatment has already started.

Importantly, we find that 31.6% of the student-term observations correspond to three or more failed modules (or 25% or more of all modules), which would typically lead to a retention if that number of retentions concerned the 3rd term result. Grades range from 1 to 5, in which 1 and 2 correspond to a fail.[7] 12% of all student-term observations fail 50% or more of the 12 modules. Finally, there is a very large percentage of fails in Maths, over 40%, and a smaller but still large percentage of fails in Portuguese, 27.3%. Information on the other ten modules is not available on a module-by-module basis.

By contrast, Table 2 concerns cross-sectional data from all students as observed in a specific time period. We select the second quarter (i.e. January-March of the 2007/08 school year) as the program had not started in most schools at this stage. Average age is 13.9 and 49.9% of students are girls. 4,548 students out of 15,307 are subject to treatment and, on average, such treatment begins in quarter 3.5. 33.7% of students failed 25% or more of their modules in that quarter. Average class size is 22.9 while average school size (7th and 8th grades only) is 251 students. The average month number of the interview that led to the start of the intervention was 7.8 (in which 1 is January 2008, 12 is December 2008 and so on), and the average month when the specific intervention plan was put in practice was 12.9 (i.e. virtually January 2009).

The previous tables pool data from treated and non-treated students. Table 3 presents descriptive statistics separately for each group, again referring to the second term of the 2007/08 school year. We find that non-treated students are younger than treated students (13.7 vs 14.4 years of age), more female (51% vs 46%) and exhibit much better achievement

[6]Students from the five schools that were not treated in the first two years of the program are excluded from this descriptive analysis and most regression results. We will return to these schools in Section 4.2, as they turn out to be very useful for identification purposes.

[7]The curriculum in the 3rd cycle includes: two languages (Portuguese and a foreign language - typically English), two social sciences (typically history and geography), maths, physics/chemistry and natural sciences, visual arts, technology/ICT, physical education, ICT, and moral/religious education and non-disciplinary areas, in a total of 12 modules. See European Commission (2007) for more details on the Portuguese education system and its lower secondary level.

levels (19.4% fail 25% or more modules vs 67.4% in the case of treated students). Similarly large gaps in performance are documented for the other measures of achievement, including fails in maths and Portuguese.

Another dimension of interest concerns the ten school districts where the EPIS program is active. Table 4 presents descriptive statistics about some of the main dimensions of interest of the program on a district-by-district basis. This table makes clear the considerable heterogeneity across districts in terms of the sizes and numbers of schools and also the numbers of students. The latter variable ranges from 109 students in only one school (Aljezur district) to 3,423 students in 17 schools (Matosinhos). The percentage of interventions also varies considerably, from 16.7% (Setúbal) to 56.9% (Aljezur); however, their timing is less dispersed as it is in all cases between the third and fifth quarters, except in one instance - Paredes, the first school district where EPIS was implemented - when the average quarter falls between the second and third quarters. Differences in gender and age across school districts are very small, while achievement levels (overall retention and fails in maths or Portuguese) vary reasonably, more or less proportionately to the percentage of interventions.

Finally, we present a number of figures that report additional descriptive statistics from the EPIS program. For instance, Figure 2 describes the number of students and their fail rates for which data are available in each term of each grade and school year. This figure makes clear an interesting phenomenon about the evolution of grades across the three terms of a school year: grades tend to be much lower (i.e. more cases in which there are 25% or more fails) in the first and second terms when compared to the third and final (and so decisive) term. This may be related to the fact that it is the third term result that will determine if the student progresses, although that decision involves some discretion that can be exercised by teachers. Moreover, the third term grade is also supposed to reflect the overall performance of the student over that term and the previous two terms, a feature that makes it more surprising to observe such a pronounced fall in fail rates from the second to the third term. Furthermore, similar findings are documented for the specific cases of maths and Portuguese - Figure 3 - and the 50%-or-more fails and actual retention levels - Figure 4.

Figure 5 describes the number of interventions by quarter and the average school year and age of the students that are involved in each intervention timing. As indicated before, most interventions start in the third and fourth quarters (ie April to June 2008 and September

to December 2008). In the first case, students' school year ranges between the 7th and 8th, while in the second case their grade ranges between the 8th and 9th. Similarly, in the first case their average age is 14 while in the second case it is already 14.5.

4 Results

We estimate the main effects of the introduction of the EPIS program from student achievement equations, inspired on a linear probability model framework. Specifically, we estimate equations based on three-term differences (i.e. $\Delta y_{it} = y_{i,t} - y_{i,t-3}$) of the 25%-or-more-fails dichotomous variable (y_{it}), as follows:

$$\Delta y_{it} = \beta_1 EPIS_{it} + X'_{it}\beta_2 + \alpha_i + \tau_t + u_{it}. \quad (1)$$

In most specifications, Δy_{it} is a variable referring to student i in term t that takes value one if the student failed three or more modules (out of a total of 12 modules) in term t but not in term $t-3$. If the student does not change her status (i.e. failing three or more modules in both t and $t-3$ or not failing three or more modules in both t and $t-3$), the value of the dependent variable is 0. Finally, if the student improves from failing three or more modules (out of a total of 12 modules) in term $t-3$ but not doing so in term t, then the dependent variable will take value -1.

$EPIS_{it}$ is a dummy variable equal to one if student i begins her EPIS intervention in period t (and zero otherwise), α_i is a student fixed effect and τ_t is a comprehensive time fixed effect (a dummy variable for each academic year/term/school year combination). This very detailed set of time effects follows from the evidence reported in Section 3 of marked time effects, in particular a steep decline in fails from the first and second terms to the third term.[8] Finally, X is a vector of time-varying control variables, including age dummies, a dummy variable for first-sit/resit status and a dummy variable indicating the type of program (one if vocational).

Now turning to the results, panel A of Table 5 considers a streamlined version of equation 1 which excludes all control variables except for the time and student fixed effects. The first

[8]There is also evidence that upper secondary school national exam marks fluctuated considerably over some of the years we cover here (Martins 2010). Some of those fluctuations may also have occurred at the lower secondary level and our detailed time effects should take them into account.

result (column A) also excludes student fixed effects and finds that, in a pooled cross-section analysis, students subject to the EPIS program are 9.2 percentage points less likely to fail than other students. This result is robust to the inclusion of student fixed effects on top of the time effects mentioned above as the EPIS coefficient again exhibits a large and statistically significant magnitude: -9.1 percentage points (column B).

These two results suggest that the EPIS program has an important effect in terms of cutting (likely) retention - or, more precisely, in terms of reducing quarterly grades that involve 25% or more fails. Moreover, when considering actual retention (a dependent variable that is equal to the difference in 25% or more fails variable for terms 1 and 2 but considers the difference in the actual decision regarding progression in the case of term 3), the effect increases, to -10.5%. The difference between the two estimates suggests that the EPIS program has additional effects on top of the reduction in the number of failed modules that prompted the discretion awarded to teachers to decide on the retention of borderline cases. For instance, EPIS students may have exhibited progress in behavioural domains that can be taken into account by teachers. On the other hand, the EPIS effect appears to be weaker in terms of reducing the prevalence of more extreme cases of (likely) retention, namely when students fail 50% or more of their modules in a term. In this case, the coefficient falls to -4.3 percentage points (column D), although it remains highly significant.

At least part of these results could be driven by the selection of students into treatment, a hallmark of the program itself. Indeed, there are still important elements of time-varying heterogeneity that may need to be acknowledged as they could correlate with treatment status. This may matter despite the fact that we already control for time-invariant (observed and unobserved) heterogeneity. For instance, some students will have been retained and therefore will be resiting the same modules in the following year, which is likely to facilitate their progression (students that are retained have to resit all modules, even those they have passed in their first sit). If such students were entering the EPIS program at the same time as they are repeating a year and this were not controlled for, then one may mistakenly attribute to the program an effect that should instead be imputed to the repetition of the year. In particular, if resiting a year facilitates progression, then our estimates of the EPIS effect on retention would be biased downward.

In order to address this potentital problem, Panel B of Table 5 presents the results of the

same specification adopted above (in Panel A) but now including several time-varying control variables: a set of age dummies, a dummy variable (Vocational) equal to one if the student is attending a vocational course (CEF), typically only available for students aged 15 and above (who will therefore typically be repeating a grade), and a dummy variable (Retention) equal to one if the student is resitting the same school year as in the year before. Column A indicates that these time-varying variables can play an important role in the estimation of the EPIS effect, as suggested above, as the coefficient falls considerably to -4.6 percentage points. However, most EPIS coefficients from specifications with control variables prove to be quite close to the result in the counterpart specification without such controls, although generally smaller. For instance, in the case of column B, panel B presents a coefficient of -9.0 percentage points (significant at the 0.1% level), which compares with -9.3 percentage points in panel A. The fact that the time-varying controls make such little difference to the estimates of the EPIS effects (except in the specification without student fixed effects) is reassuring in terms of the assumption of random treatment conditional on time-invariant differences made in difference-in-difference panel data methods.

The results on the control variables are also interesting in themselves, namely in terms of the understanding of the determinants of achievement at the school level (Lazear 2001). For instance, we find that students taking the more vocational-oriented programs are less likely to fail. It is unclear if this is due to the selection of students into these programs or a causal effect of the vocational programs themselves. Both effects are estimated with precision and exhibit large magnitudes; in the latter, perhaps more likely, case, the reduced fail rate may be driven by a better match between the syllabus and the students interests and/or easier requirements to pass.

Overall, these results indicate that the EPIS program has a negative effect upon the probability of retention (or, more generally, of failing modules), of about 9 to 10 percentage points. Putting that number in context by considering the average retention level across all student-observations, in the second quarter of the 2007/08 school year, 67.4% (see Table 3), the effect amounts to approximately at least one sixth of the mean level. If the benchmark figure was the retention levels in the last quarter, the relative magnitude of the effect would be even higher. On the other hand, the effects upon very low achievers (with 50% or more fails) are somewhat smaller.

4.1 Cumulative effects?

Given the relatively large magnitude of the effect documented above, we turn next to the question of its timing. In particular, we want to know if the effect arises soon after the start of the intervention on each student or if the effect is instead a more gradual, perhaps monotonously increasing phenomenon and becomes significant only some time after the intervention began. This analysis may also help in terms of clarifying the scope for Hawthorne effects (Levitt & List 2009), when individuals change their behaviour because they are subject to some form of monitoring, not necessarily because of the treatment itself. For instance, a case in which the EPIS effects occur immediately but then do not grow over time would be far more consistent with Hawthorne effects (and far less interesting from a policy point of view) than a situation in which the immediate effects are small or zero while the longer-term effects are larger.

We consider the empirical merits of this alternative interpretation by adapting equation 1 and extending the range of indicators of the start of treatment:

$$\Delta y_{it} = \beta_1^a EPIS_{i,t} + \beta_2^a EPIS_{i,t-1} + \beta_3^a EPIS_{i,t-2} + X'_{it}\beta_4^a + \alpha_i^a + \tau_t^a + u_{it}^a. \quad (2)$$

All variables take the same meaning as before; while $EPIS_{i,t-1}$ is a dummy variable equal to one if the student joined the program in the previous quarter and $EPIS_{i,t-2}$ is a dummy variable equal to one if the student joined the program two quarters before. This set up allows one to estimate any cumulative effects from the reform - they will arise if the effects of treatments that take place less recently are stronger than those that take place more recently.[9]

Table 6 presents the results. We find that most specifications indicate stronger, negative effects when the intervention started earlier in terms of the comparison in achievement. For instance, column B indicates that the effect from an intervention that started two quarters before is -36.9 percentage points, while that effect drops to -16.1 and -12.3 percentage points when the intervention started one quarter before or in the same quarter, respectively. All coefficients are statistically significant at the 0.1% level.

[9]The evaluation analysis conducted by EPIS staff follows a similar set up, although not distinguishing between different quarters of introduction of the program, and finds an increase in the probability of approval (non-retention) of 13.7 percentage points. Other differences with respect to the approach adopted here concern the control for retention status and other time-varying variables and for time fixed effects.

4.2 Quasi-experimental evidence

An evaluation exercise of this type must pay attention to the profile of the outcome of interest over time, in particular just before the intervention begins. The well-known Ashenfelter's dip stylised fact (Ashenfelter 1978, Heckman & Smith 1999) highlights the potential effect of selection in undermining the before-period comparison value used for estimation purposes. In our case, it could be that pupils enrolled into the program are those that are experiencing particularly poor grades due to transitory shocks. Once the shocks disappear, their grades would resume their earlier trajectories and their probability of retention would consequently also fall. However, this hypothesis would obviously challenge the causal interpretation of our findings so far. Indeed, it would mean we were regarding the improvement in grades subsequent to the shock as the effect of the program when in fact it was due (at least in part) to a regression to the mean phenomenon.

Here we examine this hypothesis in more detail drawing on a natural experiment related to the staggered introduction of EPIS across schools. In particular, in a specific district (Amadora) only part of those schools were intervened in the first two years of the program - recall the statistics reported in Table 4 and discussed in Section 3.1. This happened because it was not possible to obtain enough EPIS staff to implement the program in all schools due to timing and planning constraints. It was then decided to focus on schools whose students had lower achievement levels, even if practical issues, including schools' openness towards the EPIS program, also affected the selection process. However, detailed data on student achievement and background was obtained from all schools in the district, including those where the intervention was delayed.

We draw on this event to estimate a difference-in-difference matching (DDM) model. In the first stage, we assemble a matched sample of treated students, from schools that participated in the program (i.e. treated students in 'treated' schools) and comparable students from schools that did not participate in the program. In order to ensure that the two samples are as comparable as possible, we focus on students based in other schools in the same district or in a neighbouring district (Odivelas).[10] In the second stage, we use that matched sample to estimate the effects of the program. In other words, we use students from non-treated schools that are (very) similar to treated students from treated schools as counterfactuals to estimate

[10] Matching the two samples proved less satisfactory when considering the full sample of treated schools but the qualitative results - available upon request - are unchanged.

the effect of the EPIS program.

Tables 7 and 8 report the results from the matching exercise, based on the estimation of a propensity score and the imposition of a common support between the two samples. The very long list of matching variables allows one to argue that the scope for unobserved heterogeneity across the two samples to drive the results is limited. Moreover, the test of the equality of the means of each variable across the two samples is not rejected in all cases, which is evidence of the quality of the matching.

Finally, when estimating the effects using the new sample, we found results that are consistent with the main findings from the full sample of treated schools - see Table 9. Panel A (based on current term effects as in Panel B of Table 5 and equation DID1) indicates coefficients that are always negative, in three cases significantly so (at least at the 5% level), and of similar magnitudes than in the main sample, even if on the lower bound. Moreover, when also considering the effects from the introduction of the program in the previous quarter or the quarter before, the point estimates of those earlier quarters tend to increase considerably (in absolute value), as in the equivalent results based on the full sample. Although the precision of these estimates falls in some specifications, this can be attributed to the smaller sample sizes.

Overall, we take these estimates to support a causal interpretation of the main results in this paper. Students that enter the EPIS program exhibit significant declines in their fail rates. The results are robust to the comparison of students subject to the program's interventions with very similar students but that are not involved in the program at the same time because their schools only joined EPIS later.

5 Extensions

5.1 Alternative specification

Here we consider a new form of equation 1 in which the dependent variable is equal to one if the student fails (i.e. fails 25% or more of her modules in a given quarter):

$$y_{it} = \beta_1^b EPIS_{i,t}^* + X_{it}'\beta_2^b + \alpha_i^b + \tau_t^b + u_{it}^b. \tag{3}$$

This specification is based on a new version of the EPIS variable, which is now equal

to one from the moment the program is introduced. This allows one to compare the mean achievement levels before and after the student is exposed to the program.

Table 10 presents the results. We find in Panel A that, except in the case that excludes student fixed effects (column A), EPIS status decreases the probability of failing. The magnitude of the effects is again very similar to the results documented in our benchmark specification. Moreover, the results are again particularly robust to the addition of vocational program and retention status controls. For instance, the EPIS effect moves from -9.1 percentage points in the specification B of Panel A to -9.3 in the same specification but with the additional two control variables (plus age dummies).

5.2 Additional robustness tests

To test the robustness of the results to other data sources, we conducted a simple difference-in-differences analysis based on national exams data. These exams are compulsory for 9th grade students, provided they are not retained (see Martins (2010) for more details on the data used and the structure of the examinations). Our goal is to assess the extent to which school-level grades and number of exams are affected by the EPIS program.

We do find evidence of a decline in the national exam marks and an increase in the number of exams. However, point estimates are not always very precise (results available upon request), which we explain taking into account the relative small number of students involved that reached the national exam level. In this case, the results are consistent with the main findings in the paper, in that the reduction of fails and retention would increase the number of students entitled to sit the national exams. Moreover, to the extent that these marginal students are below the mean achievement levels in their schools, the school-level results would fall, as they do in the data.

In another analysis, we consider two additional dependent variables: the students' achievement in maths and in Portuguese. The dependent variable in this case is the three-term difference in a dummy variable equal to one if the student fails maths (or Portuguese). The results - presented in Table 10 - suggest that the effects of these two specific modules are smaller than in the more encompassing cases of across-the-board fails considered before. These effects can be explained by the possibly different nature of maths and Portuguese, namely in terms of their greater intensity in cognitive skills.

We also examined if the effects of the program differ across students depending on the workload of their EPIS staff (the 'mediadores'). In fact, the standard deviation of the load across the 63 staff in mid 2008 is 30.5, which compares with an average load of 72.3 students, thus highlighting some scope for dispersion in the intensity of treatment across schools. We then split the sample between students of low-load agents (up to 90% of the mean) and students of high-load agents (110% or more of the mean). Finally, we run our benchmark specification separately for each group of students. Perhaps somewhat surprisingly, our results (not shown but available upon request) indicate that the effects of the program appear to be stronger for students under high-load EPIS staff than for students under low-load colleagues. These results suggest that staff and/or their load are not allocated randomly.

Finally, we check whether the main results differ depending on the students' gender. Table 12 replicates our main analysis but splitting the entire sample between boys and girls. The results indicate that the EPIS effects are present in the two cases. However, the program appears to be somewhat stronger in the case of girls. For instance, when considering the dependent variable based on 25% or more fails, the effect of EPIS for boys is -7.5 percentage points while the effect for girls is -10.7 percentage points.

5.3 Cost-benefit analysis

Here we offer a simple back-of-the-envelope cost-benefit analysis of the program.We consider the (wage) cost of each EPIS staff (approximately 25,000 euros per year), the main variable cost of the program, and then multiply it by the total number of mediators (65, even if some of those are teachers paid by the Education department). We then divide that product by the number of non-retained students that are estimated to have been able to progress as a consequence of the program intervention (10% to 30% of 5,000). This results in a cost of approximately 1,000 to 3,250 euros per student that otherwise would have failed her year.

These figures should be compared with estimates of the cost of one year of schooling, which can be estimated at around 3,000 euros per year (e.g. if the Education Department paid tuition fees charged by private schools, when there are no public schools available in a specific region). In this case, the program would be - in the worst possible case - just about cost effective. However, when taking into account any other gains for the students that are not retained, not to mention the externalities of education (Moretti 2004, Martins & Jin 2010),

the program moves decidedly into the cost effective range. The same would apply the longer lasting are the program effects.

6 Conclusions

The substantial dispersion of schooling levels in workforces around the world have prompted governments to launch a number of programs that seek to enhance academic achievement levels, in particular of disadvantaged pupils. Most of these programs involve some form of remedial education, where students that perform poorly receive extra tuition targeted at cognitive skills. However, at the same time that the causal effects of many of these programs have been shown to be small at best, there is increasing awareness of the role of non-cognitive dimensions (motivation, discipline, self-esteem, etc) in terms of pupil achievement (Heckman & Rubinstein 2001, Heckman et al. 2006).

This paper addresses the potential of interventions that focus on non-cognitive skills by presenting empirical evidence about the effects of the EPIS program. This is an original, large-scale program that seeks to improve achievement and reduce retention and early school leaving by strengthening students' non-cognitive skills. So far, the program already targeted 15,000 lower-secondary students in 85 Portuguese state schools. EPIS is original also due to its screening of all the students in the participating schools, ensuring that the resources invested by the program - mostly the program staff time, typically devoted to small-group or one-on-one sessions - are only spent on the students that are likely to need the intervention the most.

We explore the rich longitudinal data on all students (treated and non-treated) and the differences in the roll-out of the program across and within schools to identify the effects of the program upon student achievement. In particular, we exploit the fact that, for a subgroup of schools, the intervention stage of the program was deferred due to operational constraints, even if data and diagnostics were collected as in other schools. Our evidence indicates that, unlike many remedial programs, EPIS had a significantly positive effect in terms of improving the achievement levels of treated students. The probability that a student is retained falls by at least 10 percentage points. This effect increases to up to 30 percentage points when considering cumulative effects. Moreover, our results also indicate that the program is cost effective.

References

Aaronson, D., Barrow, L. & Sander, W. (2007), 'Teachers and student achievement in the Chicago public high schools', *Journal of Labor Economics* **25**, 95–135.

Ashenfelter, O. C. (1978), 'Estimating the effect of training programs on earnings', *Review of Economics and Statistics* **60**(1), 47–57.

Banerjee, A. V., Cole, S., Duflo, E. & Linden, L. (2007), 'Remedying education: Evidence from two randomized experiments in India', *Quarterly Journal of Economics* **122**(3), 1235–1264.

European Commission (2007), The education system in Portugal, Eurybase report.

Glewwe, P. & Kremer, M. (2006), *Schools, Teachers, and Education Outcomes in Developing Countries*, Vol. 2 of *Handbook of the Economics of Education*, Elsevier, chapter 16, pp. 945–1017.

Heckman, J. & Carneiro, P. (2003), Human capital policy, NBER Working Papers 9495.

Heckman, J. J. & Rubinstein, Y. (2001), 'The importance of noncognitive skills: Lessons from the GED testing program', *American Economic Review* **91**(2), 145–149.

Heckman, J. J. & Smith, J. A. (1999), 'The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies', *Economic Journal* **109**(457), 313–48.

Heckman, J. J., Stixrud, J. & Urzua, S. (2006), 'The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior', *Journal of Labor Economics* **24**(3), 411–482.

Holmlund, H. & Silva, O. (2009), Targeting non-cognitive skills to improve cognitive outcomes: Evidence from a remedial education intervention, IZA Discussion Paper 4476.

Jacob, B. A. & Lefgren, L. (2004), 'Remedial education and student achievement: A regression-discontinuity analysis', *Review of Economics and Statistics* **86**(1), 226–244.

Krueger, A. B. & Lindahl, M. (2001), 'Education for growth: Why and for whom?', *Journal of Economic Literature* **39**(4), 1101–1136.

Lavy, V. (2009), 'Performance pay and teachers effort, productivity and grading ethics', *American Economic Review* **99**, 1979–2011.

Lavy, V. & Schlosser, A. (2005), 'Targeted remedial education for underperforming teenagers: Costs and benefits', *Journal of Labor Economics* **23**(4), 839–874.

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

Lazear, E. P. (2003), 'Teacher incentives', *Swedish Economic Policy Review* **10**(2), 179–214.

Leuven, E., Lindahl, M., Oosterbeek, H. & Webbink, D. (2007), 'The effect of extra funding for disadvantaged pupils on achievement', *Review of Economics and Statistics* **89**(4), 721–736.

Levitt, S. D. & List, J. A. (2009), Was there really a Hawthorne effect at the Hawthorne plant? An analysis of the original illumination experiments, NBER Working Paper 15016.

Machin, S., McNally, S. & Meghir, C. (2004), 'Improving pupil performance in English secondary schools: Excellence in cities', *Journal of the European Economic Association* **2**(2-3), 396–405.

Martins, P. (2010), Individual teacher incentives, student achievement and grade inflation, CEE Discussion Paper 0112.

Martins, P. & Jin, J. (2010), 'Firm-level social returns to education', *Journal of Population Economics* **23**(2), 539–558.

Martins, P. S. & Pereira, P. T. (2004), 'Does education reduce wage inequality? Quantile regression evidence from 16 countries', *Labour Economics* **11**(3), 355–371.

Martins, P. S. & Walker, I. (2006), Student achievement and university classes: Effects of attendance, size, peers, and teachers, IZA Discussion Paper 2490.

Moretti, E. (2004), 'Workers' education, spillovers, and productivity: Evidence from plant-level production functions', *American Economic Review* **94**(3), 656–690.

Pereira, A. I., Goes, A. R., Ferreira, A. J., da Silva, C. F., Rijo, D., Barros, L. & Nossa, P. (2008), *Trabalhar para o sucesso escolar: Manual para técnicos*, Texto Editora/EPIS.

Rivkin, S. G., Hanushek, E. A. & Kain, J. F. (2005), 'Teachers, schools, and academic achievement', *Econometrica* **73**(2), 417–458.

Figures

Figure 1: **Percentage of 18-24 year olds not in training and with lower secondary schooling at most**

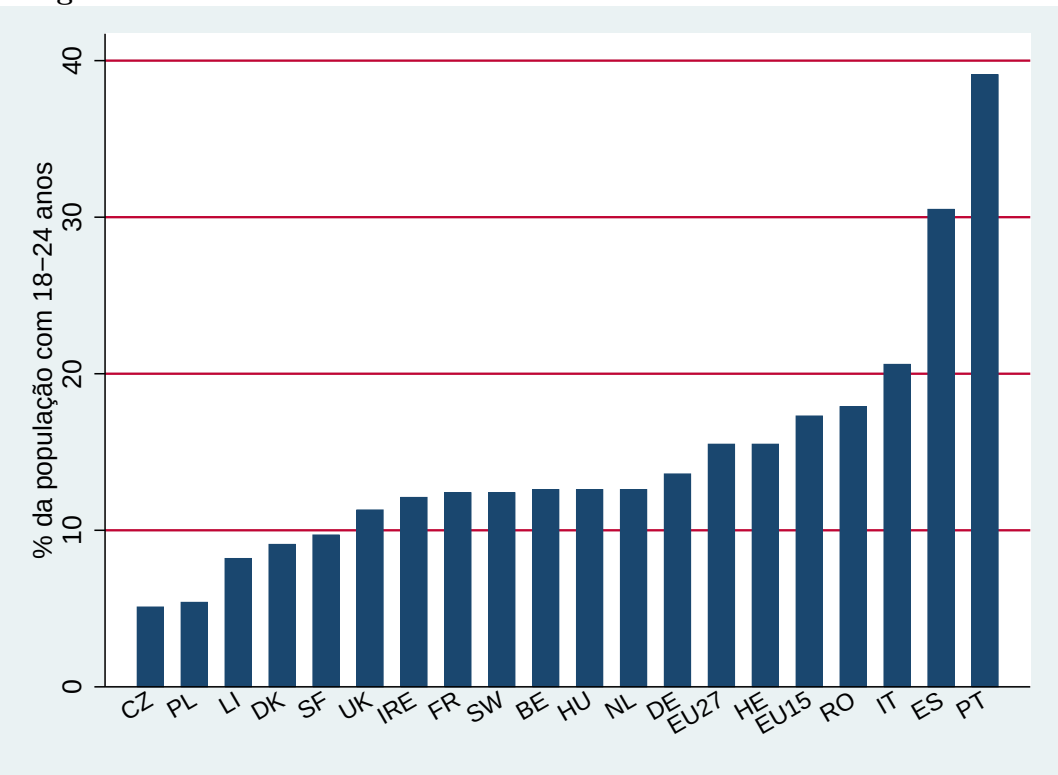
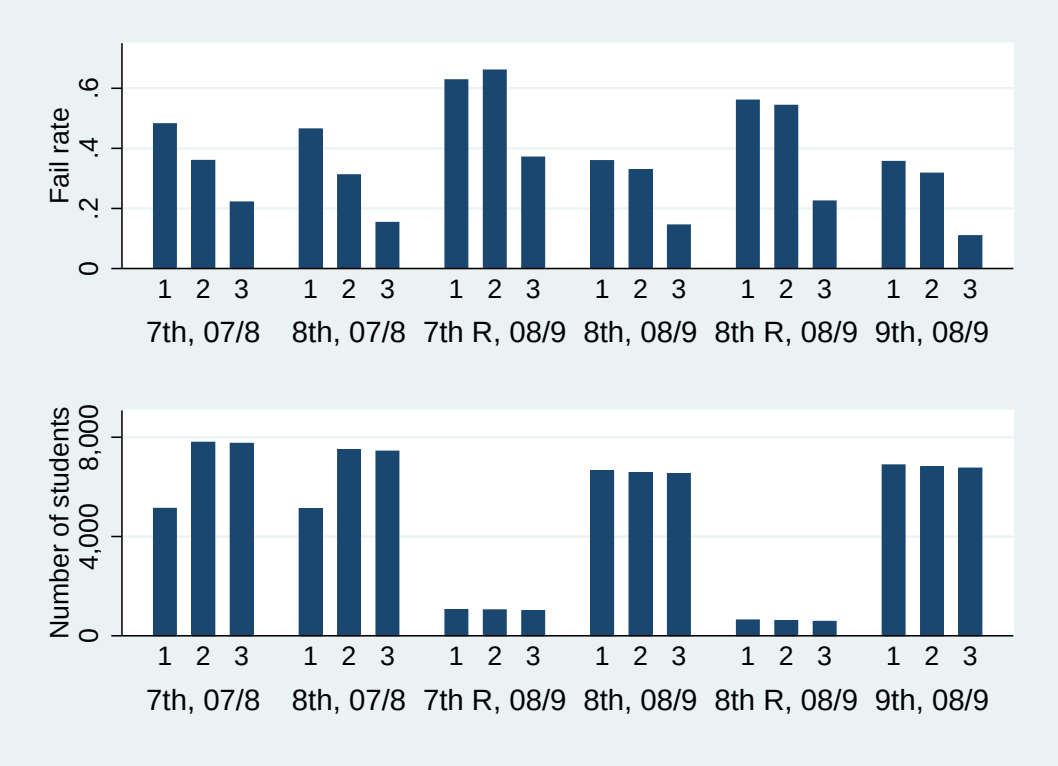
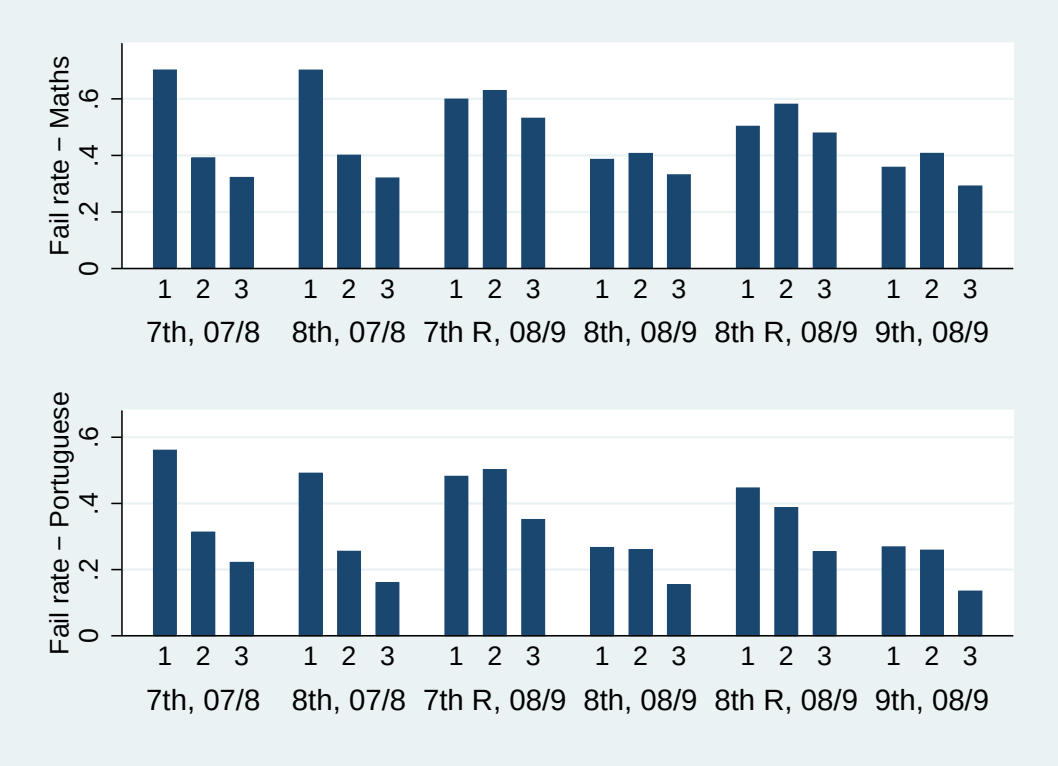


Figure 2: **Fail rate and number of students, by academic and civil years and term**



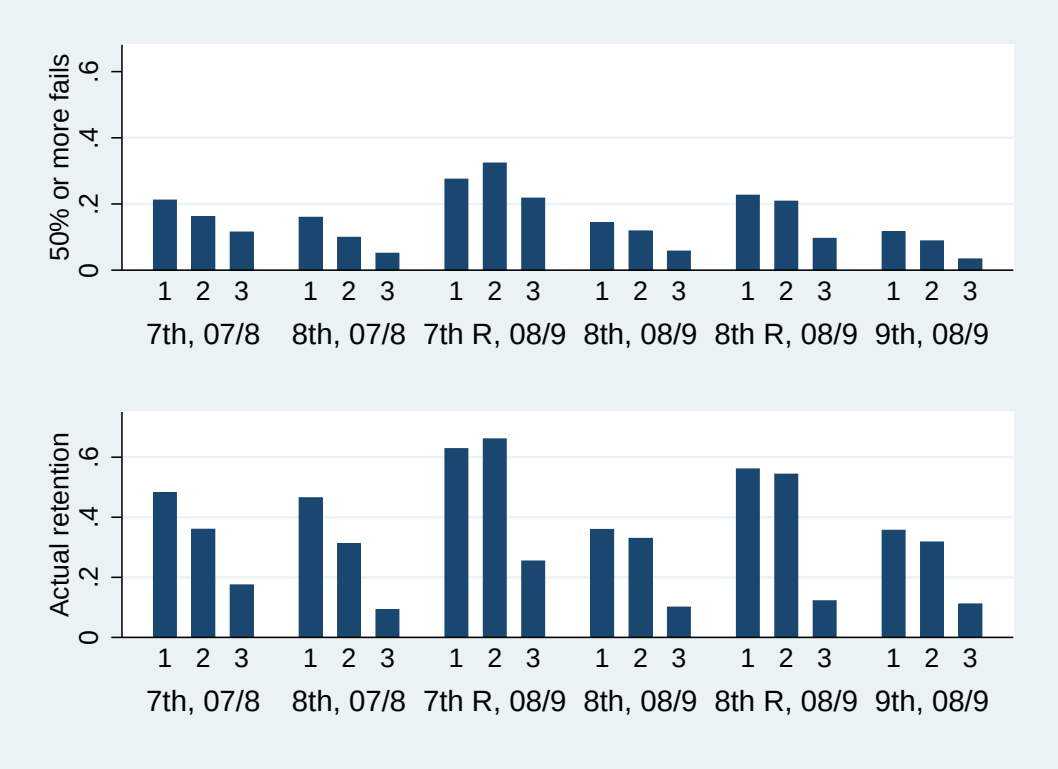
Notes: '7th, 07/8' concerns students enrolled in the 7th grade in the school year 2007/2008; '8th, 07/8' concerns students enrolled in the 8th grade in the school year 2007/2008; '7th R, 08/9' concerns students enrolled in the 7th grade in the school year 2008/2009 that are repeating the grade; and so on. '1', '2', and '3' concerns the term (quarter): the first term runs from September to December, the second term from January to March and the third term from April to June.

Figure 3: **Fails in Maths and Portuguese, by academic and civil years and term**



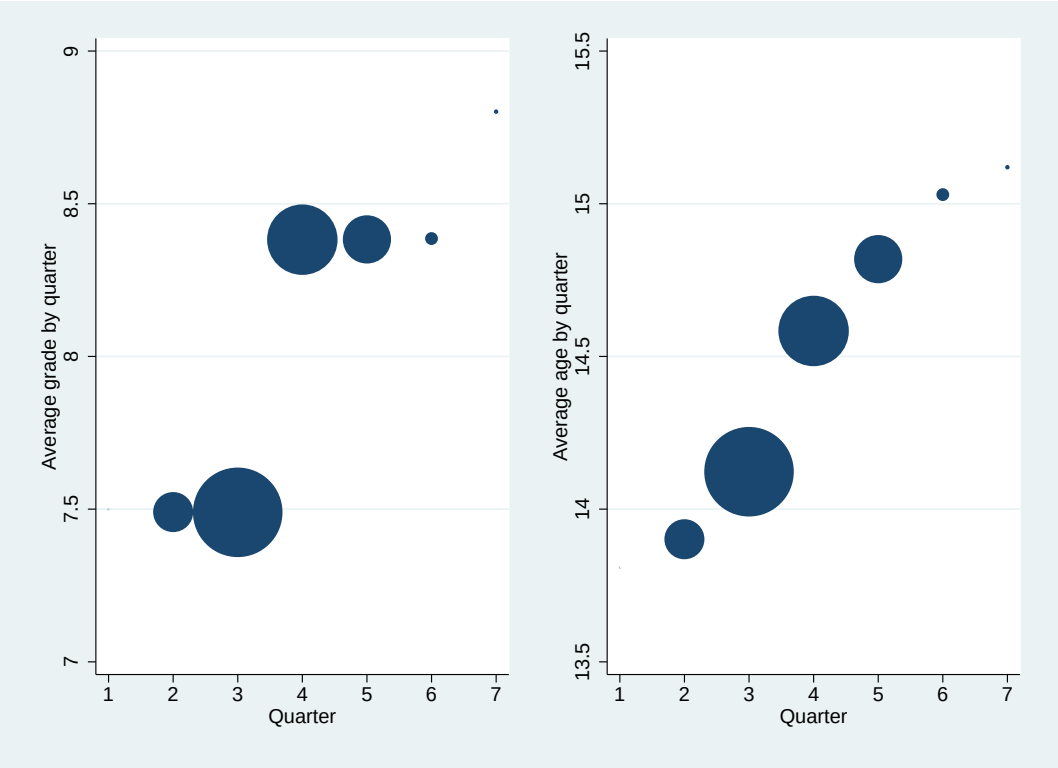
Notes: 'Fail rates' refer to a result of 1 or 2 (out of 5) in a term. Only if the student obtains a result of 1 or 2 in the third term will the student fail the module. See notes to Table 2 for more details.

Figure 4: **50% or more fails and retention, by academic and civil years and term**



Notes: 'Actual retention' concerns students that fail to progress. This variable is equal by construction to 'Fail rate' in terms 1 and 2 but may differ in term 3 as the actual rules for progression vary from school to school and involve some discretion by teachers. See notes to Table 2 for more details.

Figure 5: **Grade, age and interventions started, by year and term**



Notes: The size of each circle indicates the number of students in each category. Quarter 1 corresponds to the first term of the 2007/08 academic year; quarter 2 corresponds to the second term of the 2007/08 academic year; and so on until quarter 7 (first term of the 2009/10 academic year). See notes to Table 2 for more details.

Tables

Table 1: **Descriptive statistics, pooled student-quarter data**

Variable	Mean	Std. Dev.	N
Age	14.463	1.17	92085
Female	0.497	0.5	93901
EPIS	0.187	0.39	93901
Program participant	0.308	0.462	93901
EPIS-quarter	0.048	0.213	93901
Failed 25% or more modules	0.316	0.465	93901
Failed 50% or more modules	0.12	0.325	93901
Retained	0.303	0.46	93901
Failed Maths	0.405	0.491	90941
Failed Portuguese	0.273	0.446	91008
Vocational	0.045	0.208	93901
Quarter	3.916	1.83	93901
Civil year	8.032	0.735	93901
School year	7.651	0.63	93901

Notes: EPIS is a dummy variable equal to one when treatment is ongoing for that student (ie 0 in the quarters when the treatment has not started and 1 when the treatment had started; students that are never treated have all quarters as 0). Program participant is a dummy variable equal to one if student is (eventually) subject to treatment. EPIS-quarter is a dummy variable equal to one only in the quarter when treatment begins for that student. Vocational is a dummy variable equal to one if student enrolled in technical program. Quarter is a specific term-year combination (1 in the first term of the 2007/08 academic year, 2 in the second term of the 2007/08 academic year, and so on up to 7 in the first term of the 2009/10 academic year).

Table 2: **Descriptive statistics, pooled student data**

Variable	Mean	Std. Dev.	N
Age	13.902	1.124	15170
Female	0.499	0.5	15307
Quarter intervention began	3.506	0.871	4548
Program participant	0.297	0.457	15307
Failed 25% or more	0.337	0.473	15307
Failed 50% or more	0.131	0.337	15307
Retained	0.337	0.473	15307
Failed Maths	0.396	0.489	15307
Failed Portuguese	0.285	0.451	15307
Vocational	0.033	0.178	15307
Signalled	0.033	0.18	15307
Retentions	1.388	0.733	13862
Punctuality	0.917	0.277	11726
Parental permission	0.948	0.222	14672
Students per class	22.892	4.207	15307
Students per school	251.161	85.989	15307
Date student interview	7.817	3.745	4551
Date intervention plan	12.947	1.97	4074
Zooming - student	0.638	1.099	15307
School year	7.49	0.5	15307

Notes: The results concern the second quarter (January-March) of the 2007/08 school year. 'Signalled' is a dummy variable equal to 1 if the student has special needs or psychological problems diagnosed previously. 'Retentions' is a dummy variable equal to 1 if the student was retained at least once in the previous three years. 'Punctuality' is a dummy variable equal to 1 if the teacher responsible for the student's class states that the student is punctual. 'Parental permission' is a dummy variable equal to 1 if the student's parent agreed that the student participated in the program. 'Students per class' measures the number of pupils in the same class as the student. 'Students per school' measures the number of pupils in the same school as the student and that are monitored by EPIS. 'Date student interview' and 'Date intervention plan' is measured in months from 1 (January 2008; eg 13 is January 2009); 'Date intervention plan' indicates when EPIS staff determined the specific program that would apply to the student. 'Zooming - student' is a dummy variable equal to one if the student was subject to a second wave of screening. See notes to Table 1 for information on the remaining variables.

Table 3: **Descriptive statistics, pooled student data**

	Treated students			Non-treated students		
Variable	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**
Age	14.449	1.17	4501	13.671	1.019	10669
Female	0.464	0.499	4551	0.514	0.5	10756
Quarter intervention began	3.506	0.871	4548			0
Program participant	1	0	4551	0	0	10756
Failed 25% or more	0.674	0.469	4551	0.194	0.396	10756
Failed 50% or more	0.297	0.457	4551	0.061	0.239	10756
Retained	1.812	0.897	4551	1.182	0.527	9311
Failed Maths	0.674	0.469	4551	0.279	0.448	10756
Failed Portuguese	0.523	0.5	4551	0.184	0.388	10756
Vocational	0.048	0.214	4551	0.026	0.16	10756
Signalled	0.068	0.251	4551	0.019	0.136	10756
Punctuality	0.824	0.381	3927	0.963	0.189	7799
Parental permission	1	0.021	4549	0.925	0.264	10123
Students per class	22.316	4.600	4551	23.136	4.004	10756
Students per school	241.656	83.883	4551	255.183	86.555	10756
Date student interview	7.817	3.745	4551			0
Date intervention plan	12.946	1.969	4072			0
Zooming - student	2.144	0.911	4551	0	0	10756
Proximity assignment	0.975	0.155	4551	0.025	0.156	10756
Screening - student	2.578	0.635	4551	1.206	0.453	9311
Screening - family	1.113	0.423	4551	1.003	0.076	9311
Screening - graffar	1.39	0.52	4551	1.08	0.277	9311
Screening - school	1.449	0.539	4551	1.182	0.403	9311
School year	7.495	0.5	4551	7.489	0.5	10756

Notes: The results concern the second quarter (January-March) of the 2007/08 school year. 'Proximity assignment' is a dummy variable equal to 1 if assigned to enter the program. See notes to Table 1 for information on the remaining variables.

Table 4: **Descriptive statistics, by school district**

School District	No. students	Interventions %	Interventions Quarter	Female	Age	Retained	Failed Maths	Failed Portug.	No. Schools	No. Staff	No. Classes
Aljezur	109	0.569	3.016	0.440	14.291	0.319	0.319	0.297	1	1	5
Amadora	1585	0.339	4.577	0.509	14.559	0.326	0.400	0.302	8	10	77
Amadora (control)	1647	0	NA	0.476	14.492	0.221	0.297	0.237	6	5	74
Matosinhos	3423	0.238	3.258	0.485	14.067	0.201	0.301	0.169	17	10	165
Odivelas	2851	0.336	3.453	0.483	14.372	0.187	0.321	0.226	13	10	132
Paredes	2287	0.275	2.379	0.499	14.066	0.189	0.345	0.224	9	8	99
Resende	224	0.496	3.838	0.473	14.471	0.278	0.333	0.167	1	2	11
Santarém	966	0.176	4.106	0.488	14.054	0.117	0.232	0.121	6	7	45
Setúbal	2722	0.167	4.499	0.492	14.284	0.128	0.302	0.136	10	9	117
Tavira	504	0.313	3.139	0.508	14.326	0.179	0.324	0.235	2	3	23
V. F. Xira	2484	0.285	3.330	0.507	14.155	0.161	0.306	0.145	10	7	116

Notes: All statistics refer to a snapshot in the third quarter of 2007/08. 'Amadora (control)' denotes the set of schools that serve as an additional control group (where the interventions started much later); 'Amadora' denotes the set of schools in that school district where the interventions started at approximately the same time as in the other school districts. 'Interventions' is the percentage of students that are treated. See notes to Table 1 for information on the remaining variables.

Table 5: **Effects of EPIS on fails and retention: Main results**

	A	B	C	D
Dependent variables:	$\Delta Fail$ (25% or more)	$\Delta Fail$ (25% or more)	Retention	$\Delta Fail$ (50% or more)
Panel A				
EPIS	-.092	-.093	-.105	-.043
	(.015)***	(.017)***	(.017)***	(.015)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	45924	45924	45924	45924
R^2	.061	.465	.469	.469
Panel B				
EPIS	-.046	-.090	-.101	-.037
	(.015)***	(.018)***	(.018)***	(.016)**
Vocational	-.323	-.407	-.399	-.283
	(.029)***	(.068)***	(.064)***	(.053)***
Retention	-.563	-.657	-.692	-.452
	(.016)***	(.043)***	(.042)***	(.034)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	45069	45069	45069	45069
R^2	.149	.484	.489	.484

Notes: The dependent variables are, alternatively: the three-term change in a dummy variable equal to one if the student has 25% or more failed modules in the term, which will typically lead to retention if in the third and final term (columns A and B); the three-term change in a dummy variable equal to one if the student has 25% or more failed modules in terms A or B and retained in the third term (column C); or the three-term change in a dummy variable equal to one if the student has 50% or more failed modules (column D). All specifications include detailed time dummy variables (one for each school year/civil year/term combination). The EPIS dummy variable is one only in the term when the programme started for that student. Age dummies are included in all columns. Vocational is a dummy variable equal to one if the student is enrolled in professional degrees (CEF). Retention is a dummy equal to one if the student is resitting a grade. Each observation corresponds to a student/academic year/term combination. Standard errors allow for clustering at the class level. Significance levels: *: 0.05; **: 0.01; ***: 0.001.

Table 6: **Effects of EPIS: Differences over time**

	A	B	C	D
Dependent variables:	$\Delta Fail$ (25% or more)	$\Delta Fail$ (25% or more)	Retention	$\Delta Fail$ (50% or more)
$EPIS_{t-2}$	-.149	-.369	-.376	-.297
	(.092)	(.104)***	(.104)***	(.110)***
$EPIS_{t-1}$	-.072	-.161	-.185	-.117
	(.027)***	(.039)***	(.036)***	(.033)***
EPIS	-.053	-.123	-.139	-.062
	(.016)***	(.019)***	(.019)***	(.017)***
Vocational	-.324	-.401	-.392	-.279
	(.029)***	(.069)***	(.064)***	(.053)***
Retention	-.563	-.657	-.692	-.452
	(.016)***	(.043)***	(.042)***	(.034)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	45069	45069	45069	45069
R^2	.149	.485	.491	.485

Notes: See notes to Table 5. The dependent variables are the same as in Table 5. $EPIS_t$ is a dummy variable equal to one if the program started for that student in that term. $EPIS_{t-2}$ ($EPIS_{t-1}$) is a dummy variable equal to one one (two) term(s) before the term when the program started for that given student. Significance levels: *: 0.05; **: 0.01; ***: 0.001.

Table 7: **Quality of matching between students in treated and non-treated schools**

Variable	Sample	Treatment group Mean	Control group Mean	p-value
25%-50% fails (1st term)	Unmatched	0.42869	0.34606	0.003
	Matched	0.42869	0.42965	0.962
50%+ fails (1st term)	Unmatched	0.44473	0.56802	0
	Matched	0.44473	0.4305	0.485
25%-50% fails (2nd term)	Unmatched	0.39156	0.34368	0.083
	Matched	0.39156	0.3631	0.153
50%+ fails (2nd term)	Unmatched	0.34768	0.44391	0
	Matched	0.34768	0.34709	0.976
25%-50% fails	Unmatched	0.21941	0.179	0.08
(Interaction 1st and 2nd terms)	Matched	0.21941	0.20556	0.41
50%+ fails	Unmatched	0.26498	0.37232	0
(Interaction 1st and 2nd terms)	Matched	0.26498	0.24853	0.359
Fail in Maths	Unmatched	0.83544	0.82339	0.571
	Matched	0.83544	0.84246	0.642
Fail in Portuguese	Unmatched	0.72405	0.68258	0.107
	Matched	0.72405	0.71946	0.803
Fail in Maths and	Unmatched	0.56371	0.50835	0.05
Portuguese (interaction)	Matched	0.56371	0.56276	0.963
Previous retentions	Unmatched	1.816	1.79	0.599
	Matched	1.816	1.8349	0.599
Female	Unmatched	0.47848	0.43675	0.141
	Matched	0.47848	0.46504	0.512
Year	Unmatched	7.4641	7.4869	0.423
	Matched	7.4641	7.4423	0.285
Age	Unmatched	14.264	14.428	0.019
	Matched	14.264	14.305	0.414
$Age^2/10$	Unmatched	20.495	20.975	0.018
	Matched	20.495	20.612	0.42

Notes:

Table 8: **Quality of matching between students in treated and non-treated schools (cont.)**

Variable	Sample	Treatment group Mean	Control group Mean	p-value
(50%+ fails (1st term))*Age	Unmatched	6.274	8.1311	0
	Matched	6.274	6.0758	0.493
(50%+ fails (2nd term))*Age	Unmatched	4.932	6.3758	0
	Matched	4.932	4.8857	0.868
Psychological or other problems	Unmatched	0.10549	0.05489	0.002
	Matched	0.10549	0.11542	0.441
Parents schooling	Unmatched	3.5325	3.5418	0.828
	Matched	3.5325	3.5366	0.894
Parents job type	Unmatched	4.2262	4.284	0.14
	Matched	4.2262	4.2367	0.722
Parents income level	Unmatched	3.3181	3.4821	0
	Matched	3.3181	3.2906	0.333
Graffar group	Unmatched	1.4878	1.4988	0.722
	Matched	1.4878	1.4718	0.47
Screening group (student)	Unmatched	2.6996	2.7685	0.02
	Matched	2.6996	2.6967	0.896
Screening group (family)	Unmatched	1.1266	1.0835	0.078
	Matched	1.1266	1.134	0.697
Screening group (school)	Unmatched	1.4734	1.4081	0.038
	Matched	1.4734	1.4861	0.592
Class size	Unmatched	22.247	24.007	0
	Matched	22.247	22.324	0.709

Notes:

Table 9: **Effects of EPIS: DDM results**

	A	B	C	D
Dependent variables:	$\Delta Fail$ (25% or more)	$\Delta Fail$ (25% or more)	Retention	$\Delta Fail$ (50% or more)
Panel A				
EPIS	-.044	-.050	-.064	-.076
	(.027)	(.031)*	(.031)**	(.030)**
Vocational	-.264	-.422	-.415	-.186
	(.053)***	(.131)***	(.135)***	(.112)*
Retention	-.468	-.592	-.636	-.460
	(.032)***	(.072)***	(.072)***	(.072)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	4835	4835	4835	4835
R^2	.176	.485	.503	.517
Panel B				
$EPIS_{t-2}$	.0002	-.084	-.141	-.309
	(.105)	(.136)	(.122)	(.156)**
$EPIS_{t-1}$	-.061	-.025	-.078	-.145
	(.040)	(.064)	(.058)	(.056)***
EPIS	-.055	-.059	-.089	-.122
	(.030)*	(.036)	(.036)**	(.034)***
Vocational	-.262	-.419	-.409	-.174
	(.053)***	(.132)***	(.137)***	(.115)
Retention	-.469	-.593	-.638	-.465
	(.032)***	(.072)***	(.071)***	(.072)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	4835	4835	4835	4835
R^2	.176	.485	.503	.519

Notes: See notes to Table 5. The dependent variables are the same as in Table 5. $EPIS_t$ is a dummy variable equal to one if the program started for that student in that term. $EPIS_{t-2}$ ($EPIS_{t-1}$) is a dummy variable equal to one one (two) term(s) before the term when the program started for that given student. Significance levels: *: 0.05; **: 0.01; ***: 0.001.

Table 10: **Effects of EPIS: Alternative specification**

	A	B	C	D
Dependent variables:	25% or more fails	25% or more fails	Retention	50% or more fails
Panel A				
EPIS (*)	.300 (.007)***	-.091 (.007)***	-.099 (.008)***	-.040 (.006)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	93807	93807	93807	93807
R^2	.126	.591	.574	.494
Panel B				
EPIS (*)	-.094 (.009)***	-.093 (.007)***	-.101 (.008)***	-.042 (.006)***
Vocational	-.207 (.016)***	-.466 (.026)***	-.462 (.025)***	-.275 (.021)***
Retention	-.066 (.014)***	-.004 (.021)	-.002 (.020)	.005 (.019)
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	91991	91991	91991	91991
R^2	.262	.6	.583	.5

Notes: The dependent variables are now measured in levels, not in differences as in Table 5. $EPIS_t$ is a dummy variable equal to one if the program started for that student in that term and in all terms after that. See notes to Table 5 for more details. Significance levels: *: 0.05; **: 0.01; ***: 0.001.

Table 11: **Effects of EPIS: Mathematics and Portuguese**

	A	B
Dependent variables:	$\Delta Fail$ Mathematics	$\Delta Fail$ Portuguese
$EPIS_{t-2}$	-.176	-.154
	(.123)	(.169)
$EPIS_{t-1}$	-.051	-.102
	(.039)	(.047)**
EPIS	-.043	-.062
	(.022)**	(.023)***
Vocational	-.509	-.333
	(.084)***	(.082)***
Retention	-.562	-.462
	(.047)***	(.042)***
Student FE	X	X
Time FE	X	X
Obs.	42321	42380
R^2	.494	.454

Notes: See notes to Table 5. The dependent variables are the same as in Table 5. $EPIS_t$ is a dummy variable equal to one if the program started for that student in that term. $EPIS_{t-2}$ ($EPIS_{t-1}$) is a dummy variable equal to one one (two) term(s) before the term when the program started for that given student. Significance levels: *: 0.05; **: 0.01; ***: 0.001.

Table 12: **Effects of EPIS: Differences by gender**

	A	B	C	D
Dependent variables:	$\Delta Fail$ (25% or more)	$\Delta Fail$ (25% or more)	Retention	$\Delta Fail$ (50% or more)
Panel A: Boys				
EPIS	-.042	-.075	-.086	-.035
	(.021)**	(.025)***	(.026)***	(.023)
Vocational	-.309	-.343	-.345	-.331
	(.035)***	(.071)***	(.069)***	(.061)***
Retention	-.551	-.688	-.729	-.444
	(.020)***	(.054)***	(.052)***	(.042)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	22639	22639	22639	22639
R^2	.149	.482	.489	.479
Panel B: Girls				
EPIS	-.052	-.107	-.117	-.039
	(.021)**	(.024)***	(.025)***	(.022)*
Vocational	-.348	-.493	-.470	-.223
	(.037)***	(.096)***	(.090)***	(.085)***
Retention	-.576	-.620	-.646	-.460
	(.022)***	(.057)***	(.056)***	(.056)***
Student FE		X	X	X
Time FE	X	X	X	X
Obs.	22430	22430	22430	22430
R^2	.151	.488	.491	.494

Notes: See notes to Table 5. The dependent variables are the same as in Table 5. $EPIS_t$ is a dummy variable equal to one if the program started for that student in that term. $EPIS_{t-2}$ ($EPIS_{t-1}$) is a dummy variable equal to one one (two) term(s) before the term when the program started for that given student. Significance levels: *: 0.05; **: 0.01; ***: 0.001.