Policy Research Working Paper 9085

Addressing High School Dropouts with a Scalable Intervention

The Case of PODER

Ciro Avitabile
Janina Cuevas
Rafael de Hoyos
Julian C. Jamison

WORLD BANK GROUP
Education Global Practice
December 2019

Policy Research Working Paper 9085

Abstract

Working with the Mexican Ministry of Education, this study piloted a scalable program to reduce high school dropout rates by focusing on socio-emotional skill development and mathematics tutoring. The intervention was evaluated through a randomized field experiment with more than 5,000 youths at 20 upper secondary schools in Mexico City. An intention-to-treat analysis finds some evidence that exposure to the Opportunities and Development to Avoid Risks Program increases socio-emotional skills, but no evidence that it improves math outcomes or future attendance. Likely explanations for these null results include low take-up and other process factors, which are document qualitatively, as well as heterogeneous treatment effects. In particular, an inverse-probability-weighted matching model is suggestive of an effect whereby some students participate actively in the program and drop out of school less often, while other students choose not to participate when given the option and actually drop out more as a result.

This paper is a product of the Education Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at cavitabile@worldbank.org.

Addressing High School Dropouts with a Scalable Intervention: The Case of PODER

Ciro Avitabile[*], Janina Cuevas[+], Rafael de Hoyos[♦], and Julian C. Jamison[*,#]

Keywords: education policy; enrollment; cognitive behavioral therapy; RCT; Mexico
JEL: C93, D91, I21, I25, O15

The paper has been screened to ensure that no confidential information is revealed. The design of the PODER project benefited from discussions with Julia Quinn and John Wolf at the University of Chicago Urban Lab, Alan Safran and Barbara Algarín at SAGA, Toni di Vittorio and Jorge Rodríguez at Youth Guidance and participants at the seminar "Strategies to reduce high school dropouts in Latin America" organized by the Mexican Ministry of Education (SEP) and the World Bank. We are especially indebted to Marcela Silveyra at the World Bank and Estafanía Moleres at SEP for project management during the implementation of the project. This project would not have been possible without the continuous support from Rodolfo Tuirán at SEP, Mariana Gutiérrez at UNAM, Ricardo Cantoral at IPN and Tony di Vittorio and Jorge Rodríguez at Youth Guidance. The views expressed here are those of the authors alone, and do not necessarily reflect the opinions of the World Bank.

[*] University of Surrey and World Bank
[♦] XABER and ITAM
[+] UNICEF
[#] University of Exeter

I. Introduction

One of the strongest predictors of future success is educational attainment, which is especially important in the context of low- and middle-income countries where variance is highest. There are many drivers of continued schooling – or lack thereof – from income constraints and institutional structures to the behavior of parents and teachers. Given the laudatory accomplishment among many countries in promoting near-universal primary education, the battleground has often become secondary schools: this is where the marginal student is likely to drop out. Yet in that context the trade-off between remaining and leaving is determined in large part by the students themselves (de Hoyos, Rogers and Székely 2016), both for legal reasons (in some developing countries attendance is no longer compulsory) and practical ones. This is a highly influential decision to be made by an adolescent with a still-developing brain, a weak link between choices and consequences, a well-known preference for the present (Sutter et al. 2013), and a strong susceptibility to peer effects (Gaviria and Raphael 2001). Dropping out of school, for instance to take a low-skill job and earn some cash, is tempting in the moment but difficult to reverse, and it can have lasting impacts on wages, mobility, and welfare more broadly (Oreopoulos 2007).

What types of interventions are likely to have the greatest effect on this type of decision-maker? More broadly, is it possible to achieve not just improved decisions, but improved decision-making? Can psychological and socioemotional skills be taught? If so, do they translate to better outcomes, especially among vulnerable and at-risk populations? Although we cannot fully address all those questions in this paper, we provide some initial answers in the context of students transitioning from lower secondary to upper secondary school in the metropolitan area of Mexico City. We rigorously evaluate a program that implemented a carefully designed module based on cognitive behavioral therapy, which was meant to encourage pupils to think before acting, to imagine the downstream consequences of their behavior, to manage their emotions, and to identify more closely with their future selves. This module was complemented by the design of a math tutoring intervention as a strategy to improve the academic readiness of at-risk students. Both components are targeted at students in the first year of upper secondary school, exactly when dropout rates tend to spike, in order to take advantage of the fact that 'real-time' interventions are more salient.

With the objective of reducing upper secondary dropout rates, the Mexican Secretariat of Public Education (SEP for its acronym in Spanish), in close collaboration with the World Bank, designed and implemented the pilot program PODER[1] in 20 public high schools in Mexico City. The design of PODER was inspired by the effectiveness of the Chicago-based program "Becoming a Man (BAM)" and acknowledges that reducing upper secondary dropouts is a complex task involving remedial actions to close skills gaps that, in many cases, have been accumulating since infancy. Two skills gaps are particularly relevant in predicting dropouts: math and socio-emotional skills. Therefore, PODER has two pillars – a workshop to improve socio-emotional skills via the use of cognitive behavioral therapy (CBT), and math tutoring, both voluntary.

[1] PODER stands for "Programa de Oportunidades y Desarrollo para Evitar Riesgos" or program of opportunities and development to avoid risks. All the information of PODER can be found in the program's website: http://programapoder.com/.

Unlike BAM, the CBT modules and the math tutoring under PODER were delivered by trained volunteer teachers in order to potentially allow the scalability of the intervention. After one year of implementation, our results show that PODER had a marginal positive effect on one out of two measures of students' socio-emotional skills and no effect on math test scores and the probability of enrolling in the second year of upper secondary. The general lack of effects is possibly explained by a very low compliance rate, with 70 percent of the students in the treatment group attending zero CBT sessions and practically none of the students attending math tutoring. Despite the low compliance rates, more than 1,000 students attended at least one CBT session. Under the provisional assumption that the attendance decision is explained entirely by observables, we show that PODER had a positive and significant effect on the probability of enrolling in the second year of upper secondary (11^{th} grade) among students who attended at least two sessions. For students attending 5 or more CBT sessions, the probability of enrolling in 11^{th} grade increases by 19 percentage points. However, among treatment students who attended zero sessions, the probability of enrolling in 11^{th} grade is reduced by 9 percentage points, possibly due to stigmatization (as suggested by qualitative evidence). There is no evidence that attending more CBT-based workshops improves math test scores.

In addition to the primary impact evaluation, we also performed a process evaluation regarding the on-the-ground implementation of the program by teachers, school directors, and administrators. Some of the complexities that were encountered, such as the timing of sessions – especially given that participation was voluntary – and the individual-level randomization, help to explain the mixed results described above. We believe that this complementary work (process evaluation), attending to the details of the context, constitutes a substantive and methodological contribution in itself. Another contribution, in addition to being one of the first experimental evaluations of a behavioral approach to reduce dropouts in a developing country, is that we focused on scalability of the intervention. Rather than using counselors who arrived with a background and expertise in psychology or social work, which would never have been possible to implement across a significant number of upper secondary schools in Mexico, the PODER program used existing teaching staff. Naturally the absolute effectiveness of this approach is lower, but the cost-effectiveness is potentially much higher, and it provides evidence on the difficulties of implementing, at scale, such a program in a developing country context.

The paper proceeds as follows: Section II provides a brief overview of the literature, a description of the upper secondary education system in Mexico and the PODER program. Section III includes a description of the evaluation sample and program implementation. The empirical strategy and data are presented in Section IV, while Section V presents the results. Finally, Section VI concludes with discussion, interpretation, and policy implications.

II. Previous Literature, Context and Intervention

2.1 Previous Literature

Both demand and supply related factors can contribute to explain teenagers' decision to drop out of school. Previous studies for Mexico have explored different reasons behind this decision, including limited financial resources, lack of information about future returns associated, and

limited relevance of the academic content for job market outcomes. This evidence has not only contributed to the country specific policy debate, but more in general improved the understanding of an important schooling outcome. Between 2008 and 2011 SEP ran the "*Aligning Learning Incentives*" pilot program which assigned monetary incentives to students, teachers and administrators linked to improvements in math test scores. As shown in Behrman et al (2015) the program was successful in improving student math test scores, but it had no effect on upper secondary (or "EMS" for its acronym in Spanish) dropout and graduation rates. More recently, Dustan (2018) and de Hoyos, Attanasio and Meghir (2019) provide rigorous evidence that two scholarship programs - "*PROBEMS"* and "Prepa Si" – had, on average, a precisely estimated null effect on proxies for high school graduation.[2] While these results seem to suggest that on average liquidity constraints do not affect the decision to drop out, results in Atkin (2016) and de Hoyos, Attanasio and Meghir (2019) suggest that the opportunity cost of schooling might be particularly high for specific sub-groups. In particular, Atkin (2016) finds that the export manufacturing boom that Mexico experienced in the 1990s generated an abundance of new low-skill formal job opportunities which substantially raised the probability of quitting for youths on the dropout margin at age 16. Avitabile and de Hoyos (2018) evaluate the impact of an information intervention, called *Percepciones*, that builds on Jensen (2010) and provides students with information on the monetary returns to high school and higher education completion. The *Percepciones* pilot had no impact on graduation rates but it increased test scores, especially among girls and students above a minimum achievement level at baseline. The evidence suggests that these reasons can only marginally explain the within and between school variation in the dropout rate.

The implicit assumption common to the interventions discussed above is that EMS school dropouts are primarily a phenomenon associated with students' context or efforts, with little or no relationship with the supply of education services. Evidence presented in de Hoyos, Rogers and Székely (2016) and de Hoyos, Estrada and Vargas (2018) shows that EMS dropout is the outcome of a structural problem that starts earlier in the education trajectory of youth and it is related to income inequality and an education system that is not capable of compensating for highly unequal initial conditions. Students with low achievement levels in 6th grade show a significantly higher probability of not graduating from upper secondary. The increasing awareness that unequal initial conditions can dramatically shape future academic trajectories has shifted the focus of practitioners and academics towards targeted interventions that address existing gaps in cognitive and non-cognitive skills. Fryer and Howard-Noveck (2017) provide one of the latest examples of how a high dosage tutoring model for reading significantly increased school attendance.

The evidence of successful programs such as "Becoming A Man" (BAM) in Chicago (Heller et al. 2017) and "Pathways to Education" in Toronto (Oreopoulos, Brown and Lavecchia 2017) that address gaps both in socioemotional and cognitive skills for high-risk populations in urban areas suggests that bundled interventions have the potential to be highly effective in contexts where the gaps are not exclusively academic.

[2] "*PROBEMS*" targeted at poor urban students in upper secondary education, while "Prepa Sí" provided untargeted scholarships for public school students in Mexico City.

2.2 The Upper Secondary Education System in Mexico

Mexico, like other middle-income countries, has reached almost universal enrollment rates in primary and lower secondary (grades 7 to 9), but its education system still faces important challenges, especially in high school (grades 10, 11 and 12). For instance, more than 35 of every 100 students who enroll in high school will never graduate. Among those who graduated from high school in 2015, more than half got insufficient achievement levels in math according to the national standardized test, PLANEA. Many of the students dropping out or finishing high school but with insufficient skills come from poor or marginalized households, replicating or even exacerbating existing inequalities. Therefore, upper secondary education dropouts and low achievement levels have important implications for Mexico's long-term economic growth and income disparities.

The upper secondary education or "educación media superior (EMS)" system in Mexico has 4.9 million students, typically between 15 and 18 years old, in grades 10th, 11th and 12th. EMS is offered by four different providers: 1) the federal government (accounting for 21.8 percent of total enrollment), 2) the state governments (47.4 percent), 3) publicly-financed autonomous universities (12.3 percent), and 4) private schools. EMS offers three types of degree programs: general, which prepares students for higher education; technological preparing students both for the labor market and higher education; and technical, which emphasizes technical and vocational education.

According to official statistics from the National Institute for the Evaluation of Education (Instituto Nacional para la Evaluación de la Educación, INEE) in 2015, only 67.3 percent of students graduated on time, usually three years after enrolling in EMS. On-time graduation rates vary across types of degree programs, with general schools showing the highest rate (69.2 percent), followed by technological schools with rates very close to the national average (67 percent) and technical schools showing the lowest (54.1 percent).

Table 1 compares the characteristics of 12th grade students from the four different providers of upper secondary using information from the context questionnaire of the national standardized test PLANEA. Students attending schools run by the federal government, the universe from which the sample of PODER was drawn, display levels of math achievement that are in between the levels observed by those attending schools run by states and autonomous universities. This is consistent with socio-economic indicators (parents' education, books and computer at home), students in federal schools come from better-off households compared to students in state-run schools but are relatively poorer than students in EMS linked to autonomous universities.

Table 1: Characteristics of EMS by service providers, %

	Federal	State	Autonomous	Private
Share of total enrollment	21.8	47.4	12.3	18.6
Students with insufficient math achievement level	62.6	73	52	61
Students that have a job	35.1	37.8	37.5	33.3
Students with scholarship	46.4	47.3	36.7	20.9
Mother with higher education	7.3	4.2	11.7	22.6
Father with higher education	9.2	5.2	14.4	24.2
More than 25 books at home	66.5	60.3	69.9	77.9
PC at home	70.2	56.5	73.2	83.4

Sources: Context survey PLANEA 2016. The sample includes students who answered the 12th grade context survey; INEE, Panorama Educativo 2016 and National results of PLANEA 2016-2017.

EMS schools run by the federal government tend to be large with an average of 1,159 students usually split in morning and evening shifts. Federal EMS schools located in the metropolitan area of Mexico City (CDMX from now onwards), are particularly large with an average of 2,200 students with most of them having morning and afternoon shifts. The EMS system is characterized by some strict promotion criteria. Students must pass five out of eight disciplinary subject areas and practical modules per semester. Otherwise they must repeat the semester. Students who fail three or fewer subject areas can enroll in the next semester, but they must attend and pass leveling courses during a time window to be determined by each school on a case-by-case basis. In addition, students must satisfactorily complete all their subject areas and modules—completed after 6 semesters when students show no delay—within, at most, 10 semesters after enrolling in EMS. Otherwise they lose the right to re-enroll and therefore the chance of having a high school diploma. Partly because of the strict promotion rules, grade repetition rates are high, 15.6 percent in 2015.[3]

According to the 2009 Survey of EMS Dropouts (Encuesta Nacional de Deserción en la Educación Media Superior), subject repetition is the second most common reason—after financial constraints—mentioned as a reason for early dropout.[4] In 2016, 15.5 percent of enrolled students dropped out of upper secondary, on average. These average repetition rates translate into an on-time graduation rate of 60 of every 100 students enrolled in EMS. More than half of total dropouts in EMS take place during the first year, with a large share of them concentrated during the first semester. A dimension that shows a large correlation with dropouts is household income. In 2012, according to the "National Evaluation Council", CONEVAL, 59.5% of young people between 15 to 18 years old, the standard age to be enrolled in upper secondary education, who were not enrolled in school were concentrated in the poorest 50%.

[3] Students who fail three or more subjects for two consecutive semesters have to repeat the entire grade.

[4] Jacob and Lefgren (2009) use a plausibly exogenous variation in retention generated by a test-based promotion policy in the United States to show that retaining low-achieving eighth grade students in elementary school substantially increases the probability that these students will drop out of high school.

A second factor highly correlated with dropouts is the pre-EMS academic performance. Given the strict promotion criteria, it is not surprising that pre-EMS low achievement levels—in primary or lower secondary—is one of the most important predictors of upper secondary dropout rates. For instance, using a longitudinal data set following students between 6th and 12th grades, de Hoyos, Estrada, and Vargas (2018) show that an improvement of one standard deviation in test scores at the end of 6th grade increases the probability of on-time high school graduation by 10 percentage points.

Overall, low achievement levels, particularly in math, are one of the most important challenges in EMS. According to the results of PLANEA, a standardized test applied to those finishing EMS in 2017, 66.1% of the students in 12th grade had an "insufficient" achievement level, 23.3% obtained "regular", 8% of all graduates had a "good" achievement level and only 2.5% achieved an "excellent" level. EMS students display not only low basic cognitive skills, but the evidence suggests that they have low levels of socio-emotional skills. According to the data from the socio-emotional module of PLANEA 2016, 20% of students declared feeling stressed very often and 28% felt they cannot control important issues of their lives.

2.3 The PODER Pilot Program

With the objective of reducing EMS dropout rates, SEP, together with the World Bank and two local universities, designed the pilot program "Programa de Oportunidades y Desarrollo para Evitar Riesgos (PODER)". The design of the program acknowledges that reducing upper secondary dropouts is a complex task involving remedial actions to close skills gaps that, in many cases, have been accumulating since infancy.

The program assumes that youth at risk of dropping out are those with significant skills gaps in two areas, math and socio-emotional skills, and therefore applies remedial actions to try to reduce those gaps. Through a CBT-based workshop and math tutoring, PODER aims to improve students' socio-emotional and math skills, respectively, hence reducing their probability of dropping out. The specific approach undertaken in our project is based most heavily on two previous papers: Blattman, Jamison, & Sheridan (2017) and Heller et al. (2017). Both develop a version of cognitive-behavioral therapy, which is one of the most evidence-based treatments in psychology and adapt it to vulnerable but non-clinical populations of young men – the former in Liberia, the latter in inner-city Chicago. Both show remarkable successes – the former for anti-social behavior and violence, the latter for crime and school enrollment. We diverge from them in terms of including mixed genders and especially in terms of using non-specialist counselors to lead the group sessions: volunteer teachers who are briefly trained in CBT techniques and the chosen curriculum, as opposed to carefully selected staff whose sole job and expertise revolves around psychological interventions for personal development.

PODER was designed between January and July 2016 and targets at-risk first year EMS students in public schools run by the federal government located in CDMX. Like BAM, student participation in PODER is voluntary, does not bear any consequences for the students' academic progress and it is implemented outside school hours. As opposed to BAM, PODER's CBT-based sessions and math tutoring are implemented entirely by school teachers whose participation is

voluntary and not financially rewarded. As will be explained in more detail below, these features of PODER—reflecting budget constraints faced by SEP—made its implementation really challenging.

SEP and the World Bank partnered with two local public universities to design the CBT-based sessions and math tutoring. The socio-emotional intervention was designed by the Faculty of Psychology at the "Universidad Nacional Autónoma de México" (UNAM) and it combines conventional cognitive behavioral therapy with "youth engagement" strategies.[5] The math tutoring was designed by the Department of Educational Mathematics at the "Instituto Politécnico Nacional" (IPN). Two versions of the socio-emotional workshop were developed by UNAM, one consisting of 10 CBT-based sessions of one hour each to be covered in a total of 10 weeks during the first semester of EMS; the second version covered the same themes, but it was delivered in 20 CBT-based sessions of one hour each to be covered during the first and second semesters of EMS. UNAM developed two separate manuals, one for workshop facilitators (trained teachers) and one for the students. It also developed and implemented a 25- and 35-hours face-to-face training sessions for workshop facilitators responsible of delivering the 10- and 20-sessions of CBT, respectively. PODER included only one version of math tutoring consisting of 10 weekly sessions of three hours each. IPN developed the math tutoring and a three-weeks training strategy for tutors—two weeks of online training and three days face to face. The math training for tutors is based on the social epistemological theory which has a paradigm that rests on the use of contexts accounting for social and cultural characteristics of the students to build meaningful mathematical knowledge.[6]

First-year EMS students at risk of dropping out were identified via an algorithm. The variables included in the algorithm had to comply with two criteria: (1) be highly correlated with the probability of dropping out, and (2) be available to all students before the start of the academic year. Two variables fulfilled these criteria, the math test score in the EMS entry examination applied to all candidates in the metropolitan area of Mexico City (known as COMIPEMS, for its acronym in Spanish)[7] and the grade point average of lower secondary. Considering budget constraints and the objective of implementing a scalable intervention, the CBT session and the math tutoring were delivered by teachers already in the system. Teachers participated voluntarily in PODER and received training and a certificate by UNAM or IPN, respectively, but no monetary incentives.

III. Evaluation Sample and Program Implementation

3.1 Evaluation Sample

[5] Youth engagement strategies are activities to create trust among the participants of the CBT session and between them and the facilitator. The youth engagement strategies were developed by Youth Guidance, the NGO implementing BAM in Chicago. Staff from Youth Guidance trained a group of psychologists from UNAM on how to implement the youth engagement strategies.

[6] The intervention, including all the manuals and guidelines of PODER can be found in https://programapoder.com/

[7] The characteristics of the COMIPEMS test can be found here: https://www.comipems.org.mx/. For a more general description of the COMIPEMS test see de Janvry, Dustan and Sadoulet (2017).

To identify the impact of PODER, we relied on a randomized control trial, assigning some students to one of two treatments or to a control group. To improve statistical power within a very tight budget, the randomization was done in two steps, first at the school level and then, within schools, at the student level. In the first step, 20 schools were randomly selected among the population of Federal EMS schools in CDMX. Annex A shows the comparison of characteristics of the 20 schools included in PODER versus the universe of Federal EMS schools in CDMX. In general, PODER schools are similar to the rest of federal schools in CDMX, with a few exceptions, such as the number of students graduated during the school year 2015-2016 and the students with scholarships. In both variables, PODER schools have, on average more graduated students and more students with scholarships.

During the second step, first-year EMS students that were part of the 20 percent with the highest risk of dropping out—according to the algorithm identifying at-risk students within each of the 20 schools—were randomly assigned to two treatment groups and control. For all the first-year EMS students in each of the 20 schools in the evaluation sample, we create school-specific ranking of students' risk of dropping out based the algorithm described above. The results of the process to identify at-risk students are shown in Annex B.

Once we have the students ranked according to their probability of dropping out, we define the number of students who are potentially eligible for being part of *PODER* after knowing the number of teachers who expressed interest and availability to act as counsellors for the *Cognitive Behavioral Therapy* (CBT) sessions. Based on the experience of *Becoming a Man* (Heller et al. 2016), we aimed at assigning two groups of maximum 15 students to each teacher. Therefore, if in one school there were 10 self-selected, counsellors, 300 students would become eligible and assigned to either type of treatment, and 150 would be assigned to the control group, both assignment at random. In total, our sample includes 5,274 students in the three groups: 10 CBT sessions (2,109), 20 CBT sessions (1,407) or a control (1,758). Annex C shows the list of EMS schools participating in PODER, together with the number of students selected and the facilitators and tutors needed in each school.

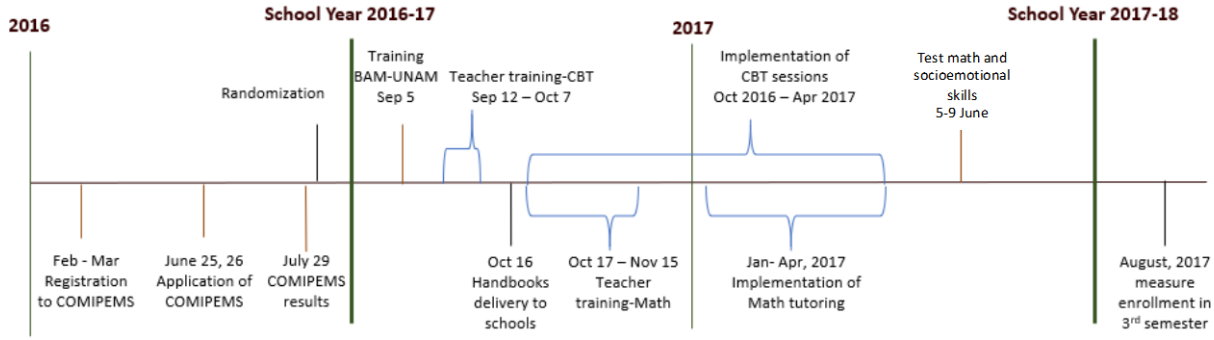
Annex D shows the characteristics of students in the top 20 percent of the distribution of risk of dropping out (eligible to be included in our evaluation sample) versus other students in the 20 PODER schools. Not surprising given the targeting criteria, students included in the PODER sample belong to households with relatively lower socio-economic status: lower parent's education, fewer books and household assets. PODER students are also more likely to be enrolled in the evening shift.

3.2 Program Implementation

PODER started its first CBT-based workshop in October 2016. The start of program implementation was preceded by many key steps: design the intervention, prepare the training and learning materials, train the teachers, perform the randomization, and inform students and their parents about the program. The stages followed in PODER's implementation are summarized in Figure 1. The design of the intervention by UNAM and IPN, in collaboration with the World Bank and SEP, took place between January and July 2016. Ninth grade students about to enter EMS

registered to sit the EMS entry examination, COMIPEMS, between February and March, candidates took the test on late June and the results were made public on July 29, 2016. With the results of COMIPEMS and the Grade Point Average (GPA) of lower secondary, the World Bank team assigned each of the first-year EMS student's in our selected schools a probability of dropping out using the algorithm described above (see Annex B). Training was implemented via a cascade mode. UNAM psychologists, trained on CBT received a two-days training by staff from "Youth Guidance" on how to undertake "youth engagement" strategies prior to the implementation of the CBT. UNAM psychologists then delivered between 25 and 35 hours of training, depending on the treatment arm, including CBT and "youth engagement" to volunteer teachers between September 5 and October 7. On October 16, the handbooks for teachers and workbooks for students participating in the CBT sessions were distributed in the 20 participating schools. The training by IPN for math tutors took place between October 17 and November 15, already well into the first semester. The evaluation's outcome variables, math test scores, socio-emotional skills and enrollment into the second year of EMS, were collected at the end of school year 2016-17 (test scores and socio-emotional skills) and at the beginning of school year 2017-18 (enrollment).

Figure 1: Timeline of PODER's implementation



The take-up rates in the program were disappointingly low (see Figure 2). In the case of the CBT-based workshop only 3 of every 10 students in the treatment group attended at least one session. Attendance rates to the math tutoring were practically zero. According to a process evaluation carried out as a companion document of the PODER impact evaluation (see Cuevas and Silveyra 2018), four shortcomings explain the low attendance rates:

1. *The program start late*. The CBT sessions started in mid-October—as oppose to early September as it was originally planned—when some students had already dropped out of school. Even though PODER took place in Federal schools, Federal SEP does not have a

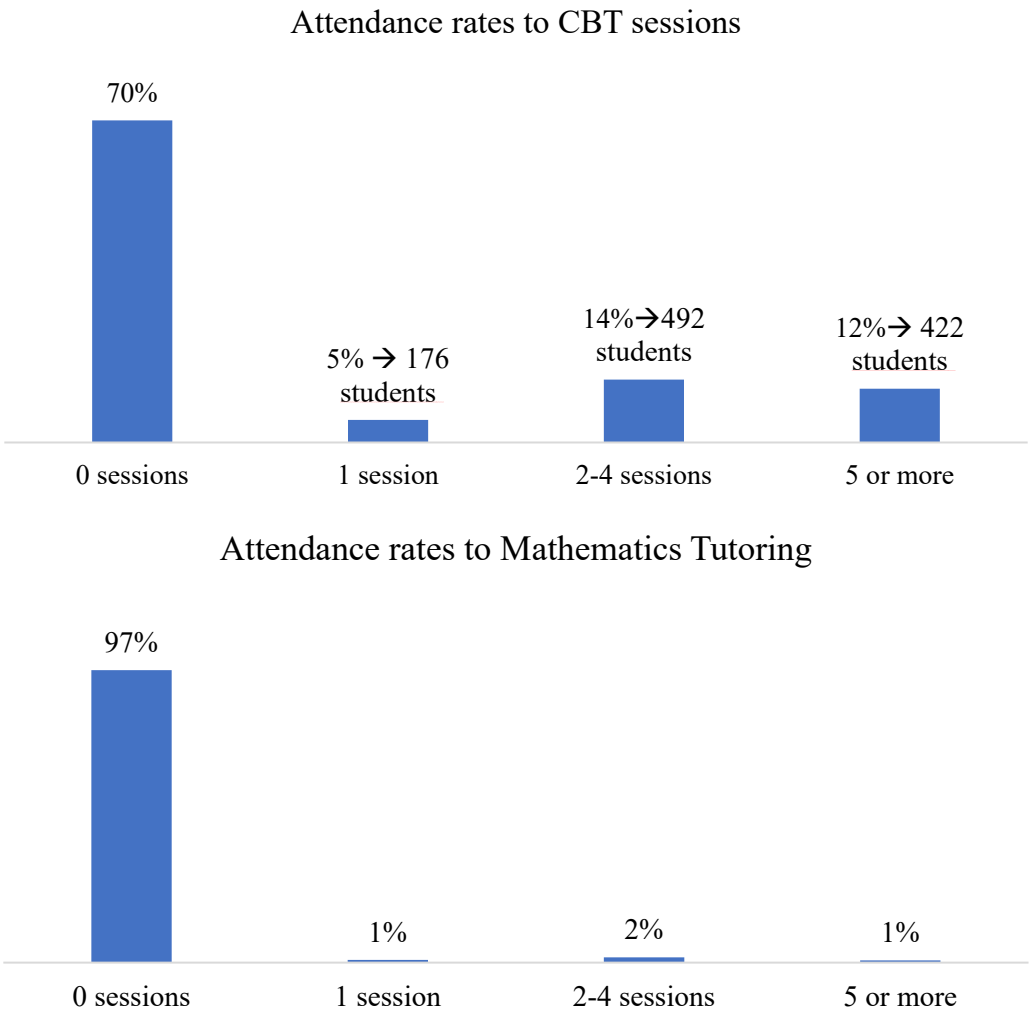
centralized school registry compiling the information of all new students enrolled in EMS. Due to teachers' limited availability to participate as tutors, the math tutoring could not start during the first semester, therefore displaying an even lower participation or attendance rate (very close to zero).

2. *The original project design did not include a communication strategy*. There was a substantial level of confusion regarding the program's objectives and content among principals, teachers and students. Lack of information about the program's objective and limited familiarity with counselling and psychological support were quickly filled by stigma, labeling PODER as a program to help "problem students". Stigmatization was a serious problem affecting attendance rates and the ultimate effects of the program.
3. *Use of out-of-school hours*. The CBT workshops and math tutoring were planned to be held during out-of-school hours. This was particularly challenging for students who had to arrive earlier at school or stay longer since many of the students in our sample, especially those in the afternoon shift, combine school with work.
4. *Lack of incentives*. The design of PODER was based on the voluntary participation of students and teachers. Lack of incentives combined with a strategy that rely on implementation during out-of-school hours was, perhaps, the single most important factor behind the low participation rates.

All these complex implementation challenges were compounded by the student-level randomization creating treatment and control groups within each school. This factor increased the role played by stigmatization and introduced further complications to the implementation of the Program, particularly in a large school setting. As it is shown in Figure 2, the take-up rates for the math tutoring were close to zero, therefore, *de facto*, PODER consisted in a CBT-based intervention aimed at improving socio-emotional skills, improving math test scores and reducing school dropouts.

Despite the significantly low take up or compliance rates, the qualitative evidence presented in Cuevas and Silveyra (2018) shows that among those who had direct experience of the intervention there is a strong appropriation and willing to continue with the CBT-based workshops. During the fieldwork, interviews, workshops and satisfaction surveys were conducted and teachers expressed repeatedly that there is a need of training and coaching on how to develop socioemotional skills, not only for the students, but also for themselves. From the teachers' point of view, there is a need of training and intervention designs in this issue that has not been filled by current training programs. Teachers and students identified PODER as a support and guideline to better understand their peers and improve the communication between teachers and youth. Not much can be said about the perception of teachers and students regarding the math tutoring since attendance was practically zero. Many school directors suggested that PODER be incorporated to the curriculum for it to be implemented during the conventional schooling hours or allow schools to decide when and under which format to implement the program.

Figure 2: Attendance rate to the CBT sessions and math tutoring



IV. Econometric Model and Data

4.1 Econometric Model

In order to estimate the Intention to Treat Effect (ITT) of PODER, we estimate the following equation:

$$(1)\ Y_{ij} = \alpha + \beta_1 CBT10_{ij} + \beta_2 CBT20_{ij} + \gamma' X_{ij} + \delta\mu_j + u_{ij}$$

Where Y_{ij} denotes the outcome of student i in school j ; $CBT10_{ij}$ takes the value 1 if student i in school j is assigned to the treatment group with 10 CBT-based sessions, 0 otherwise; while $CBT20_{ij}$ takes the value 1 if student i in school j is assigned to the treatment group with 20 CBT-

based sessions, 0 otherwise. X_{ij} denotes a set of baseline characteristics of student *i* in school *j*. In the baseline specification, X_{ij} includes a dummy for whether the student is attending the evening shift, gender, the lower secondary school GPA, whether they had taken part in the COMIPEMS assignment twice, the math and Spanish average. Since the randomization has been conducted within school, and the number of students potentially eligible defined in accordance with the number of available tutors, all the regressions account for school fixed effects (μ_j). In eq. (1) β_1 and β_2 are the parameters of interest, since they capture the effect of being assigned either to the 10 sessions or the 20 sessions group.

As discussed in section 2.3, attendance to the CBT-based sessions was very low. The limited compliance makes the ITT effects not particularly informative, therefore we complement it with the Average Treatment Effects (ATE). Among students assigned to the two treatment groups, the decision to attend a certain number of sessions might be related to student observable and unobservable characteristics. In order to account for selection, we use two alternative strategies. In the first one, we assume that students' decision about the number of sessions to attend only depends on their observable characteristics. Therefore, under the assumption that the selection is based on observable characteristics, we use matching methods to estimate the following multi-valued treatment effect equation below:

$$(2)\ Y_{ij} = \alpha + \pi_0(0Sess)_{ij} + \pi_1(1to4Sess)_{ij} + \pi_2(5plusSess)_{ij} + u_{ij}$$

$0Sess_{ij}$ takes the value 1 if student *i*, who was originally assigned to any of the treatment groups, took 0 sessions, 0 otherwise; $1\text{to}4Sess_{ij}$ takes the value 1 if student *i*, who was originally assigned to any of the treatment groups, took between 1 and 4 sessions, 0 otherwise; $5plusSess_{ij}$ takes the value 1 if student *i*, who was originally assigned to any of the treatment groups, took 5 sessions or more, 0 otherwise. The omitted category in eq. 2 is represented by individuals in the control group. The parameters π_0, π_1 and π_2 measure the ATE of taking 0, 1 to 4 sessions and more than 5 sessions. When looking at the effects of CBT sessions, we focus on categorical variables, since for policy purposes it is more interesting to identify whether there is a threshold effect rather than the effect of an additional session. To increase the probability of matching student *i* with a student who is observationally equivalent, we do not control for school fixed effects, thus allowing for potential pairs to be assigned in different schools.

4.2 Baseline Data

The main sources of data at baseline are the administrative records from the COMIPEMS assignment process, which allows us to observe the full ranked list of schooling options requested; the score in the admission exam (measuring math and Spanish language); the GPA in lower secondary; plus the information contained in the socio-demographic survey filled out at registration, which provides us with variables such as gender, age, parental education and occupation, assets, and more.

Table 2 presents the baseline characteristics of students selected to participate in the pilot, separately for those in the two treatments and the control arms. On average the three groups display similar characteristics, with only 3 tests (out of 66) failing to reject the equality between one of the

treatment arms and the control group. Only 8 percent of the students in the control group report having received an income-based scholarship in lower secondary. On average students in the sample provide 50 percent of the correct answers in the COMIPEMS admission exam, and with only 40 percent of the correct answers in math. Overall, we interpret the results in Table 2 as evidence that, the randomization was performed well, and the groups are balanced at baseline. The lower secondary GPA varies in a range between 6 (the passing mark) and 10, and therefore an average of 7.2 supports the evidence that students in the sample are on average not performing well. When it comes to their preferences over high school options, on average students in the control group report 10 preferences (over a maximum of 20). There is a relatively large share of students who took the COMIPEMS admission exam for the second time (28 percent). Nevertheless, more than 70 percent of the students are assigned to the third or higher preference.

Table 2: Baseline characteristics by treatment status

	Control Group		T1-C	T2-C	F Test	Obs
	Mean	SD				
Male	0.618	0.486	0.008	0.010	0.808	5274
Share correct answers COMI -general	50.213	11.806	0.235	-0.063	0.413	5174
Share correct answers COMI verbal	54.154	18.987	-0.196	-0.764	0.394	5174
Share correct answers COMI Spanish	53.485	18.764	0.392	0.133	0.750	5174
Share correct answers COMI Math	39.786	17.537	0.244	-0.470	0.412	5174
Share correct answers COMI Geo	52.676	18.138	0.664	0.233	0.422	5174
Share correct answers COMI History	47.267	17.535	-0.228	-0.271	0.856	5174
Share correct answers COMI Chemistry	47.300	18.222	0.187	0.300	0.866	5174
PNBIO	51.407	18.790	0.913*	0.045	0.150	5174
Sec. School Average	7.166	0.764	-0.021	-0.013	0.677	5174
Took COMIPEMS twice or more	0.278	0.448	0.008	-0.009	0.550	5274
Math Average 2nd year Lower Sec	2.872	1.629	0.013	-0.054	0.475	5174
Spanish Average 2nd year Lower Sec	3.412	1.783	-0.020	-0.106*	0.209	5174
Income based Scholarship in Lower Sec	0.080	0.272	-0.002	-0.004	0.928	5174
Mother with Primary	0.176	0.381	-0.001	-0.012	0.653	5108
Father with Primary	0.143	0.351	0.007	0.019	0.378	4799
Mother with Lower sec	0.378	0.485	0.002	0.027	0.247	5108
Father with Lower sec	0.404	0.491	-0.015	0.002	0.580	4799
Mother with Upper Sec	0.201	0.401	-0.004	-0.007	0.889	5108
Father with Upper Sec	0.240	0.427	-0.015	-0.016	0.507	4799
Mother with Higher Ed	0.231	0.422	0.002	-0.009	0.766	5108
Father with Higher Ed	0.199	0.399	0.020	-0.011	0.104	4799
Number of Preferences	10.478	4.298	0.194	0.324**	0.104	5174
Assigned first option	0.189	0.392	-0.024**	-0.003	0.102	5274
Assigned second option	0.102	0.303	-0.004	-0.006	0.827	5274
Assigned third option or higher	0.703	0.457	0.027*	0.010	0.159	5274
Phone at home	0.719	0.450	0.016	0.004	0.543	5174
Wash. Machine	0.852	0.355	0.008	0.011	0.655	5174
Fridge	0.946	0.226	0.005	0.007	0.654	5174
Microwave	0.665	0.472	0.035**	0.017	0.072	5174
Cable	0.465	0.499	0.008	-0.035*	0.047	5174
Tablet	0.501	0.500	-0.006	-0.021	0.503	5174
Number of books	3.478	1.529	0.050	0.039	0.572	5174

4.3 Outcome Measures

In order to measure the impact of the intervention, we focus on two sets of outcomes: socioemotional skills, and academic outcomes including math test scores, probability of enrollment in 11th grade, and "student engagement" as defined below. Through the CBT-based sessions, PODER aimed at influencing a large array of socioemotional skills, including: a) intrapersonal skills (such as assertiveness, perseverance, emotional regulation), b) stress management (self-control), and c) problem solving.

In June 2017, among those students who had enrolled in the second semester, we administered a multiple-choice questionnaire that students had to complete autonomously under the supervision of two enumerators, and a math standardized test and. The questionnaire included two instruments to elicit socioemotional skills: the PLANEA instrument designed by INEE in collaboration with SEP and the BarOn Emotional Quotient Inventory: Youth Version. Questions on socioemotional outcomes are asked as part of the PLANEA standardized test applied at the end of upper secondary. The instrument has the advantage that is possible to isolate some of the skills targeted by the program: perseverance, feelings' management and decision making. The BarOn aims at capturing five broad domains: (intrapersonal and interpersonal skills, stress management, adaptability, overall feeling) and 15 specific skills in total. We focus on intrapersonal skills, management, and adaptability, since PODER aimed at influencing a subset of skills within these three broad categories. The math standardized test was developed by taking items from previous editions of the national assessment "ENLACE" for 9th grade.

Nationwide, about 25 percent of the students who enrolled in 10th grade in the academic year 2015—the year prior to the implementation of PODER—did not enroll in 11th grade in the following academic year and 28 percent re-sat the COMIPEMS examination. We measure whether PODER affected the probability of enrolling in the third semester of upper secondary (11th grade) and the probability of re-sitting the COMIPEMS exam. Among EMS students in CDMX, not enrolling in 11th grade does not necessarily mean that the student decided to disengage from the education system. Some students not enrolling in 11th grade might do so to re-enroll in the centralized assignment system, COMIPEMS, to have a chance of getting a place in an academic option that is either more selective or more aligned with their preferences. In this case, it is unclear whether the decision not to enroll in 11th grade will lead to negative consequences for the student, therefore we analyze the effect of PODER on 11th grade enrollment, re-enrollment to COMIPEMS and a measure of "student engagement" combining both variables (either enrolling in 11th grade or, conditional on not enrolling, re-sitting the COMIPEMS).

V. Results

5.1 Intention to Treat Effects (ITT)

In order to simplify the analysis, we focus on summary measures of the socioemotional skills by looking at the first component, proxied through a Bartlett score, obtained by Exploratory Factor Analysis of the PLANEA and the BarOn scales, respectively. The Bartlett score is normalized using the mean and the standard deviation of the control group. Columns (1) to (4) in Table 3 present the estimates of the intention-to-treat effects on the PLANEA principal component. Columns (5) to (8) present the results for the BarOn principal component. Columns (1) and (5)

estimates the effect of the two treatment types (10- and 20-sessions of the CBT workshop) controlling only for school fixed effects. Columns (2) and (6) show the results for the baseline specification discussed above, with full set of controls. In Columns (3) and (7) we present the results combining the two treatment types. Since only those students who appeared in the school records in April 2017 were invited to the survey measuring socio-emotional skills, there might have been self-selection. For this reason, in columns (4) and (8) we present the results accounting for the possible self-selection into the survey for applying two-stage Heckman procedures, with teacher fixed effects as exclusion restriction. The underlying rationale is that specific teachers might have been more active than others in keeping children in school until the survey was taken.

Table 3: ITT on Socio-emotional skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Std Bartlett score - PLANEA				Std Bartlett score - BarOn			
10 Sessions	0.178***	0.165**			0.051	0.047		
	(0.067)	(0.068)			(0.059)	(0.060)		
20 Sessions	0.166**	0.175**			-0.035	-0.025		
	(0.074)	(0.074)			(0.065)	(0.066)		
Combined 10 and 20 sessions			0.169***	0.168***			0.017	0.027
			(0.060)	(0.061)			(0.053)	(0.054)
Inverse of Mills Ratio				-0.057				0.009
				(0.181)				(0.157)
Observations	1279	1241	1241	1203	1466	1422	1422	1382
Mean Dep. Control Group	-0.000	-0.001	-0.001	-0.001	0.000	-0.000	-0.000	-0.000
SD Dep. Control Group	1.000	1.004	1.004	1.005	1.000	0.998	0.998	0.999

Notes: columns 1 to 4 show the intention to treat (ITT) effects of PODER on the Bartlett score of an Exploratory Factor Analysis (EFA) capturing socio-emotional skills (SES) based on the PLANEA instrument. Columns 5 to 8 show the same effect but using the BarOn instrument. Columns 3 and 7 combines, in a single group, students in treatment 1 and 2. Columns 4 and 8 controls for potential self-selection into the test. All specifications include school fixed effects.

The effect on the PLANEA score is large and statistically significant, about 0.17 and 0.18 standard deviations (sd), for the 10 and the 20 sessions respectively. Both effects, separately or combined are statistically significant at conventional statistical level and are robust to potential selection into the exam. When we look at the impact on the BarOn inventory, we find an effect that is small for both treatment groups and statistically not significant. When the two treatment effects are combined, the resulting effect is about 0.03 sd after accounting for possible selection into the exam (column 8 in Table 3).

Table 4 reports the intention to treat effects of PODER on math test scores, enrollment in 11th grade, re-taking COMIPEMS 2017 or a measure of school engagement combining both 11th grade enrollment or re-taking COMIPEMS. Overall, we find a negative but small and statistically insignificant effect on math test scores and the probability of enrolling in 11th grade. The effect is

statically not significant when we measure the effects of the 10-session or 20-session CBT workshops separately (odd columns) or if we combine the two treatment types (even columns). We find positive but statistically insignificant effects when we look at the probability of re-entering the COMIPEMS entry examination (columns 5 and 6). When we combine both variables into a single indicator that proxies for "Student Engagement", we find a non-statistically significant effect (columns 7 and 8).

Table 3: ITT on Math Score, Student enrollment in 11th grade, re-taking COMIPEMS 2017 and School Engagement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math Score		Student enrollment		Re-taking COMIPEMS		Student Engagement	
10 Sessions (T1)	0.013		-0.023		0.001		-0.021	
	(0.128)		(0.016)		(0.010)		(0.016)	
20 Sessions (T2)	-0.019		-0.012		-0.000		-0.013	
	(0.140)		(0.018)		(0.010)		(0.018)	
T1 and T2 combined		-0.043		-0.023		0.004		-0.021
		(0.114)		(0.015)		(0.009)		(0.015)
Observations	1443	1400	5273	5026	5274	5027	5274	5027
Mean Dep. Control Group	5.681	5.711	0.515	0.531	0.094	0.091	0.594	0.607
SD Dep. Control Group	2.167	2.174	0.500	0.499	0.292	0.287	0.491	0.489

Notes: columns 1 to 3 show the intention to treat (ITT) effects of PODER on the probability of enrolling in 11th grade. Columns 4 to 6 show the same effect but on the probability of re-taking COMIPEMS entry examination. All specifications include school fixed effects.

5.2 Average Treatment Effects

Under the assumption that among students in the treatment groups the decision to attend a certain number of sessions depends exclusively on observable characteristics, we use matching to isolate the average treatment effects (ATE) of being exposed to different "doses" of CBT-based sessions on the socioemotional and education outcomes, vis-à-vis being assigned to the control group. We follow Cattaneo (2010) and specify four values for the treatment: 0 denoting the value assigned to students in the control group, 1 for students in the treatment group with 0 sessions, 2 for students in the treatment group who attend between 1 and 4 sessions, and 3 the value assigned to students who took 5 or more CBT-based sessions. The first stage is modeled using a multinomial logit with a flexible specification. The predicted probabilities are then used to estimate eq. (2) via a regression-adjusted inverse probability weighting (RAIPW). The graph that plots the density distributions of the propensity scores for the different exposures to the program, presented in Annex E, support the hypothesis that individuals from the different groups are relatively similar in

terms of observable characteristics. The results of the RAIPW are presented in Table 5.[8] A higher intensity of CBT sessions is not associated with statistically different impacts on socioemotional well-being or math test scores. Effects remain positive, but not statistically significant when we look at the PLANEA score, there is no clear pattern for the BarOn inventory, but all coefficients are not statistically significant.

Table 4: ATE on Outcome Measures

	(1) PLANEA	(2) BarOn	(3) Math score	(4) Enrolled in 11th grade	(5) Re-take COMIPEMS	(6) Student Engagement
0 sessions	0.097	0.001	0.029	-0.090***	0.025**	-0.065***
	(0.077)	(0.065)	(0.139)	(0.016)	(0.010)	(0.016)
1 to 4 sessions	0.143	-0.106	-0.061	0.074***	-0.038***	0.026
	(0.088)	(0.073)	(0.154)	(0.025)	(0.014)	(0.024)
5 or more sessions	0.148	0.062	0.001	0.190***	-0.068***	0.121***
	(0.093)	(0.080)	(0.174)	(0.028)	(0.014)	(0.027)
Mean Dep. Control Group	0.008	0.023	5.648	0.522	0.099	0.604
Observations	1119	1282	1261	4632	4633	4633

Notes: columns 1 to 4 show the ATE of PODER on two measures of socio-emotional skills (PLANEA and BarOn), the probability of being enrolled in 11th grade, the probability of re-taking COMIPEMS, respectively. The omitted category in this specification is represented by students who were assigned to the control group.

In the control group, we observe very little correlation between the measures of socioemotional well-being and the results on the math assessment. Therefore, any significant effect of CBT-based sessions on the math results would be concerning for the validity of our identification strategy. Reassuringly, both for students who attended 1 to 4 sessions and those who attended 5 or more, we find a null effect on the math score (column 3). Individuals, originally assigned to the treatment group, who attended 0 sessions displayed a 9-percentage point reduction in the probability of enrolling in 11th grade. Some of these individuals had quit school in order to re-enter the COMIPEMS assignment, as we observe that those who had 0 sessions in either the 10 or 20 session treatment group were 2.5 percentage points more likely to try being re-assigned to a different academic option through the COMIPEMS system. Individuals who had attended between 1 and 4 counselling sessions displayed a 7.4 percentage point increase in the probability of enrolling in 11th grade, while among those who had attended at least 5 sessions there is a 19 percentage points increase in the probability of enrolling (column 4). Conversely, we observe a reduction in the probability of re-entering the COMIPEMS assignment system: there is a 3.8 percentage points reduction among those who had attended between 1 and 4 sessions, and a 6.8 percentage points reduction among those with 5 or more sessions (column 5). When we consider the net effect, as

[8] Alternative matching strategies produce similar coefficients. Results are available upon request.

proxied by the probability of staying engaged, we do find that students in the treatment group who attended 0 sessions were 6.5 percentage point less likely to remain engaged, while those with 5 or more sessions were 12.1 percentage point more likely to remain engaged with the system (column 6).

Overall the matching results seem to suggest a highly heterogeneous impact of the intervention. In most applications, researchers are interested in measuring the effect of having any exposure to the treatment - irrespective of whether it was induced by the intervention's design or not, and the ATE can be interpreted as a rescaling of the intention-to-treat effects. This is not the case for the analysis presented here. As we are interested in isolating the effect on those who should have attended at least one CBT-based session but they did not, the ATEs identify the effects of alternative levels of exposure for those who were meant to be treated, vis a vis those who could not attend any session because of the design. Our results suggest that for some students, being assigned to the PODER pilot represented an incentive to quit the school, since treatment students with 0 sessions are more likely to drop out than those in the control group. While some of the students might not have understood the nature of the intervention and the purpose of the CBT-based sessions, for others the program might have increased the opportunity cost of staying in school, either because the sessions took place far from the regular class schedule[9] or because of the possible stigma associated to the attendance. In fact, the qualitative analysis undertaken as part of PODER's evaluation[10] suggests that, even without attending any session, students formed a negative perception of PODER as it was associated with psychological support for under-performing students. The results for students who attended even a relatively small number of sessions suggest that a relatively mild CBT-based intervention can help them address some of the behavioral issues that could affect their academic outcomes.

The hypothesis of selection on observables is potentially too restrictive. Alternative strategies, that rely on the use of the randomized assignment to overcome the potential selection issues, are not exempt from threats. For instance, the identification of the Local Average Treatment Effect parameters (Angrist and Imbens (1994)) build on three assumptions: 1) the exclusion restriction should be independent of all potential outcomes and potential treatments; 2) the exclusion restriction affects the treatment decision in the same direction for all individuals; 3) the treatment decision and the exclusion restriction are strongly correlated. In our context there are two possible serious threats to assumptions 1) and 2). Marginal students, whose opportunity cost of staying in school is already high, might have higher incentives to drop out to avoid the attendance of PODER sessions. If that is the case, the exclusion is not independent of the potential outcome when analyzing the effect on enrollment and the probability of taking COMIPEMS. If dropout occurs before the survey, the results on socioemotional outcomes might also be plagued by self-selection that the original assignment cannot overcome. The evidence presented in Table 5, with the individuals who were originally assigned to one of the two treatment groups and attended 0 sessions being more likely to drop out than those in the control group, supports this concern.

[9] Evidence from the qualitative evaluation shows that in a selected number of schools, students perceived that the attendance of the CBT-based sessions was compulsory, possibly as a result of inadequate communication with teachers and school management.

[10] See Annex F and Cuevas and Silveyra (2018) for more details.

VI. Conclusions

School dropouts are a major policy concern, and there is increasing interest in scalable non-traditional solutions that go above and beyond the provision of economic incentives and information. Given the robust success of cognitive behavioral therapy not only within psychology (e.g. for phobias and post-traumatic stress), but also as applied to at-risk but non-clinical adolescent populations in Chicago and Liberia, we helped to develop the PODER program to address school dropouts in Mexico. We conducted a randomized impact assessment using voluntarily trained teachers (and voluntary student compliance) in 20 upper secondary public schools in the metropolitan area of Mexico City.

Results as implemented were mixed at best. Among all those assigned to the treatment, socio-emotional (noncognitive) skills appear to slightly improve, but mathematics (cognitive) skills do not – probably due to extremely low attendance at the complementary tutoring sessions – and more importantly neither does enrollment in the following year, which was the primary outcome variable. Although we do not have experimental variation in the intensity of treatment, a matching model does suggest that those who participated did better than those who did not, above and beyond the selection effects which are presumably also present. A large share of the students in the treatment group might have perceived the intervention as irrelevant in terms of benefits, but potentially costly due to the stigmatization of CBT and the logistical issues.

Moreover, it is worth highlighting that the targeting criterion was exclusively based on a measure of academic risk. Since the math sessions effectively did not take place, the disconnect between the targeting criterion and the nature of the intervention might have discouraged a certain number of students. Taken as a whole, the matching results seem to suggest that students who were able to attend at least 5 sessions have benefitted both in terms of socioemotional skills and "student engagement" outcomes.

Overall, the low take-up of the intervention (partly due to the timing of the sessions outside required school hours) prevents us from making any conclusive statement. Furthermore, randomization was carried out at an individual level, rather than across schools (or even classes). Given this, as well as the fact that teachers who knew the students were directly involved, some level of spillover across treatment groups was probably unavoidable. Future work should potentially consider different designs in order to rule this out.

The results pointing to a positive effect of PODER for students who attended at least 5 sessions, in addition to the positive statements from school directors, teachers and students in the qualitative evaluation, suggest that a CBT intervention may yet be effective to reduce school dropouts in Mexico. Given the lack of effects of other interventions aimed at reducing dropouts and the sense of urgency to improve the education and labor market opportunities for youth in Mexico, it seems sensible to continue experimenting with an intervention like PODER. The evaluation of PODER yields valuable lessons to improve the program's implementation and attendance rates. Having a

strong communication strategy to reduce the risk of stigmatization and implementing the CBT workshops and math tutoring during normal schooling hours will be crucial.[11]

Policy makers in Mexico and throughout Latin America need to recognize that secondary school dropouts are the outcome of elementary education systems that are not providing all students with sufficient foundational skills. Basic numeracy, literacy and socioemotional skills are essential to face the academic requirements of upper secondary and, later in life, to secure a formal, well-remunerated and stable job. PODER begins to show that, even among youth with a large deficit in foundational skills at age 15, progress can still be possible.

References

Angrist, J.; and Imbens G. (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62 (2): 467-475.

Adelman, M.; Haimovich, F.; Ham, A.; and Vazquez, E. (2017). "Predicting School Dropout with Administrative Data: New Evidence from Guatemala and Honduras", *World Bank Policy Research Working Paper* #8142.

Atkin, D. (2016). "Endogenous Skill Acquisition and Export Manufacturing in Mexico", American Economic Review, American Economic Association, vol. 106(8), pages 2046-2085, August

Avitabile, C.; and de Hoyos, R. (2018). "The Heterogeneous Effect of Information on Student Performance: Evidence from a Randomized Control Trial in Mexico", *Journal of Development Economics* 135: 318-348.

Barrera-Osorio, F.; Bertrand, M.; Linden, L.L.; and Perez-Calle, F. (2011). "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia", *American Economic Journal: Applied* 3(2): 167-195.

Behrman, J.; Parker, S,; Todd, P.; and Wolpin K. (2015). "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools", *Journal of Political Economy* 123(2): 325 - 364

Benhassine, N.; Devoto, F.; Duflo, E.; Dupas, P.; and Pouliquen, V. (2015). "Turning a Shove into a Nudge? A 'Labeled Cash Transfer' for Education", *American Economic Journal: Policy* 7(3): 86-125.

Blattman, C.; Jamison, J.C.; and Sheridan, M. (2017). "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia", *American Economic Review* 107(4): 1165-1206.

[11] Some of these lessons are already incorporated in new versions of PODER currently implemented in the Mexican state of Nuevo León and in the province of Buenos Aires.

Cattaneo, M. (2010). “Efficient semiparametric estimation of multi-valued treatment effect under ignorability” *Journal of Econometrics*, 155: 138-154.

Cuevas, J. and Silveyra, M. (2018). “Reporte de implementación del Programa PODER”, Documento interno, Banco Mundial.

De Brauw, A; and Hoddinott, J. (2011). “Must Conditional Cash Transfer Programs Be Conditioned to Be Effective? The Impact of Conditioning Transfers on School Enrollment in Mexico”, *Journal of Development Economics* 96(2): 359-370.

de Hoyos, R; Attanasio, O.; and Meghir, C. (2019) “Can scholarships increase high school graduation rates? Evidence from a randomized control trial in Mexico”, *World Bank Policy Research Working Paper* No. 8826. World Bank, Washington, DC

de Hoyos, R.; Estrada, R.; and Vargas, M. (2018), “Predicting Individual Wellbeing Through Test Scores : Evidence from a National Assessment in Mexico”, *World Bank Policy Research Working Paper* No. 8459. No. 8826. World Bank, Washington, DC

de Hoyos, R.; Rogers, H.; and Székely, M. (2016). “Out of School and Out of Work : Risk and Opportunities for Latin America’s Ninis”, World Bank, Washington, DC

de Janvry, A.; Dustan, A.; and Sadoulet, E.; (2017). "Flourish or Fail?: The Risky Reward of Elite High School Admission in Mexico City," *Journal of Human Resources*, University of Wisconsin Press, 52(3): 756-799.

De Witte, K.; Cabus, S.; Thyssen, G.; Groot, W.; and van den Brink, H.M. (2013). “A Critical Review of the Literature on School Dropout”, *Educational Research Review* 10: 13-28.

Deininger, K. (2003). “Does Cost of Schooling Affect Enrollment by the Poor? Universal Primary Education in Uganda”, *Economics of Education Review* 22: 291-305.

Dustan, A. (2018), “Can large, untargeted scholarships increase urban high school graduation rates? Evidence from Mexico City’s Prepa Sı”, Vanderbilt University.

Fryer, R. Jr; Howard Noveck, M. (2017). “High-Dosage Tutoring and Reading Achievement: Evidence from New York City”, *NBER Working Papers 23792*, National Bureau of Economic Research, Inc.

Gaviria, A. and Raphael, S. (2001) “School-based peer effects and juvenile behavior”, *Review of Economics and Statistics*, *83*(2): 257-268.

Heckman, J., Stixrud, J. and Urzua, S. "The Effects Of Cognitive and Noncognitive Abilities On Labor Market Outcomes and Social Behavior," Journal of Labor Economics, 2006, v24 (3,Jul), 411-482

Heller, S.B.; Shah, A.K.; Guryan, J.; Ludwig, J.; Mullainathan, S.; and Pollack, H.A. (2017). “Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago”, *Quarterly Journal of Economics* 132(1): 1-54.

Jensen, R. (2010). "The (Perceived) Returns to Education and the Demand for Schooling", *Quarterly Journal of Economics* 125(2): 515-548.

Lavecchia, A.M.; Liu, H.; and Oreopoulos, P. (2016). "Behavioral Economics of Education: Progress and Possibilities", *Handbook of the Economics of Education* 5(ch.1): 1-74.

Lybbert, T.J.; and Wydick, B. (2018). "Poverty, Aspirations, and the Economics of Hope", *Economic Development and Cultural Change* 66(4): 709-753.

Oreopoulos, P. (2007). "Do Dropouts Drop Out too Soon? Wealth, Health, and Happiness from Compulsory Schooling", *Journal of Public Economics* 91: 2213-2229.

Oreopoulos, P.; Brown, R.S.; and Lavecchia, A.M. (2017). "Pathways to Education: An Integrated Approach to Helping At-Risk High School Students", *Journal of Political Economy* 125(4): 947-984.

Sutter, M.; Kocher M.G.; Glätzle-Rützler D.; and Trautmann S. T.. (2013). "Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior", *American Economic Review*, 103 (1): 510-31.

Wilson, S.J.; and Tanner-Smith, E.E. (2013). "Dropout Prevention and Intervention Programs for Improving School Completion among School-Aged Children and Youth: A Systematic Review", *Journal of the Society for Social Work and Research* 4(4): 357-372.

Annex A: **Differences between PODER schools and Federal Schools in CDMX and Estado de México**

	VARIABLE	PODER schools	SE	Obs.	Mean Dep. Control Group	SD Dep. Control Group
School year 2015-2016	Male students 1yr	-79.485	(454.508)	133	333.798	2040.313
	Female students 1yr	-161.328	(649.19)	133	410.586	2914.601
	Total students 1yr	-240.813	(1103.34)	133	744.384	4953.277
	Male students	54.062	(579.739)	133	589.506	2602.265
	Female students	-0.815	(842.525)	133	728.067	3781.702
	Total students	53.247	(1421.11)	133	1317.573	6378.756
School year 2016-2017	Num. students graduated	186.906*	(58.93)	133	191.266	274.475
	Students w/scholarship	301.837*	(154.498)	133	470.206	703.5
	Admission requests male	-139.405	(621.829)	133	441.896	2791.64
	Admission requests female	-212.772	(817.634)	133	514.921	3670.928
	Admission requests total	-352.177	(1439.09)	133	956.817	6460.869
	New students 1yr	-204.124	(1157.25)	133	832.436	5194.959
	Students enrolled in 1yr	-277.016	(1312.93)	133	897.288	5894.129
	Groups 1yr	-28.957	(76.08)	133	37.953	341.677
	Students enrolled	-22.955	(1751.66)	133	1553.73	7862.4
	Number of groups	-27	(93.522)	133	57.939	419.91
	Male teachers	10.451	(17.844)	133	29.003	80.198

	Female teachers	0.381	(30.445)	133	29.834	136.656
	Total teachers	10.831	(48.1)	133	58.837	215.94

Notes: the values in column 3 "PODER Schools" are the outcome of running a regression of each of the variables in column 2 against a dummy variable for PODER schools and sub-systems fixed effects. Using data from Formato 911-EMS 2016-2017.

Annex B: Dropout predictors

Using the full universe of students who had entered the COMIPEMS in the academic year 2008-2009, and who were supposed to complete high school in Spring 2012, we regress the dummy for taking part in the 2012 ENLACE 12th grade - a proxy for on-time EMS completion – on the lower secondary GPA score and the score in the COMIPEMS admission exam, both measured in the school year 2008-2009.[12]

While the intervention aims at tackling both academic and socioemotional issues that can lead to the decision of dropping out of high school or EMS, at the time of the pilot's design, we had no information on proxies for the socioemotional skills of 9th grade students in Mexico. Once we condition on school fixed effects, GPA and the COMIPEMS admission exams could explain 19% of the total variation in the probability of completing, on-time, upper secondary.

With the OLS estimates of the coefficients on GPA and the COMIPEMS exam in hand, for all the first-year EMS students in each of the 20 schools in the evaluation sample, we construct a predicted probability of not finishing EMS.

Dropout predictors	
COMIPEMS scores	-0.003***
	(0.018)
Secondary average	-0.174***
	(0.018)
Male	0.004
	(0.002)
Observations	186,293
Mean Dep. Control Group	0.518
SD Dep. Control Group	0.500
R-square	0.185

[12] For a discussion on the reliability of ENLACE 12th grade as a proxy for EMS completion see de Janvry, Dustan and Sadoulet (2017) and Avitabile and de Hoyos (2018).

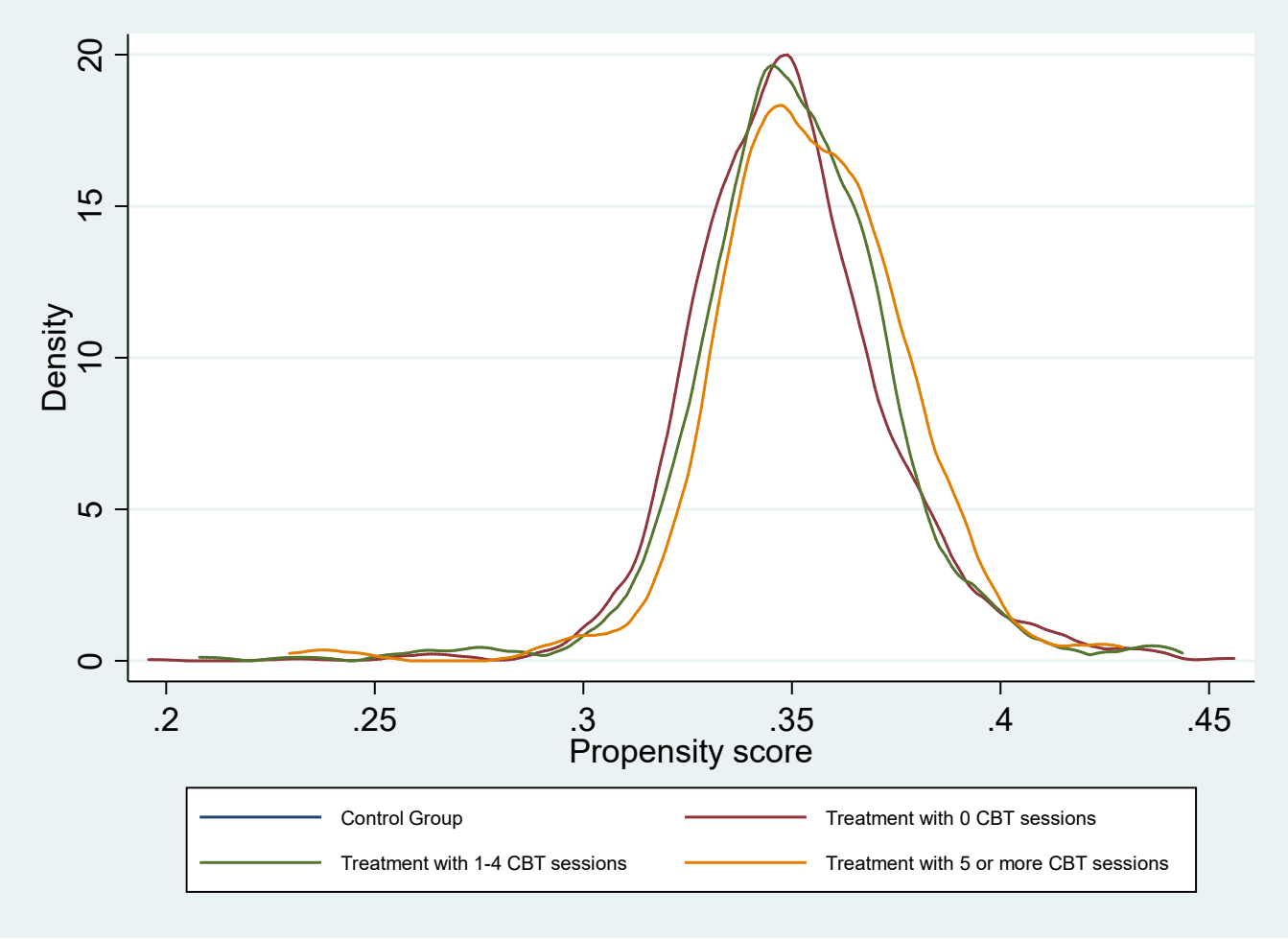
Annex C: Participating Schools, Students and Teachers

School	Num. of students (Control and Treatment)	CBT facilitators	Math tutors
DGB	181	4	4
DGETI A	120	4	4
DGETI B	137	6	3
DGETI C	363	5	8
DGETI D	227	4	5
DGETI E	181	4	4
DGETI F	183	5	4
COLBACH A	605	14	13
COLBACH B	227	15	12
COLBACH C	459	15	9
COLBACH D	491	13	8
COLBACH E	413	5	5
COLBACH F	363	6	10
CONALEP A	137	5	5
CONALEP B	181	5	4
CONALEP C	192	3	4
CONALEP D	226	3	3
CONALEP E	225	3	5
CONALEP F	181	6	5
CONALEP G	182	3	4
Total	**5274**	**128**	**119**

Annex D: Comparison between PODER students and the rest of the students in 1st grade within each school

	PODER students	Mean	SD
COMIPEMS scores	-4.398***	51.795	20.577
	(0.0313)		
Secondary average	-0.776***	7.719	1.423
	(0.023)		
Scholarship Econ Need	0.026***	1.886	0.318
	(0.009)		
Num. of books at home	-0.024*	3.806	1.621
	(0.048)		
Internet at home	-0.000	1.291	0.454
	(0.014)		
Higher Edu Father	-0.037***	0.271	0.445
	(0.013)		
Higher Edu Mother	-0.014*	0.254	0.435
	(0.013)		
Observations	22.132		
Mean Dep. Control Group	0.518		
SD Dep. Control Group	0.500		
R-square	0.185		

ANNEX E: Multi-treatment propensity scores



Annex F: Implementation Lessons

To know more about the opinion of students and teachers about the implementation of PODER, in December 2017 satisfaction questionnaires, interviews and focus groups were applied by UNAM to 32 teachers, 158 students who attended the CBT sessions and 58 students who were part of the treatment group but did not show up to any session. All these students were enrolled in 11th grade during the interviews.

In the individual interviews and focus groups that were held with teachers, it was found that 80 percent of the teachers said the topics developed in the CBT sessions were important and innovative. Almost 100 percent of the teachers expressed their willingness to participate again because CBT sessions allowed them to approach and support adolescents in an innovative way. They considered that the CBT sessions really helped students to develop three skills: emotional self-control, problem solving and decision making. These skills, in the view of teachers, help at-risk students face difficulties that arise in the family, school, with their partners and friends. From the teachers' point of view, these are some of the recommendations to improve the program:

1. Encourage the voluntary participation of students during the sessions. Do not force the participation of the students since it is not a class. Forcing participation could cause students who are reluctant to participate to leave the sessions or attend but without participating.

2. Avoid social pressure to share personal problems. Do not pressure young people to share their experiences, especially if it compromises the well-being of the student. Teachers-counselors are not trained to deal with situations of high emotional distress and could cause undesirable reactions and outcomes.

3. Put into practice the socio-emotional skills discussed during the initial training. Keep active listening, feedback and try to maintain a relaxed attitude during the session to generate a genuine motivation to participate among students.

4. Remember the pillars of cognitive-behavioral theory: use comments and reflections based on the cognitive triad (thought, emotion and behavior), so the student can reflect on their own cognitive triads. The fact that they reflect and become aware of their own ideas is more important and better than "getting them to talk".

5. Open the dialogue with real examples from the teacher-counselors' experiences. Model and adopt those attitudes that the teacher-counselor wants to encourage in the students.

For students, the main reason to attend the CBT sessions was having clear explanations from the teacher including real-life examples and the atmosphere of trust generated by the teacher and classmates. Students who did not attend declared that the schedule was the main obstacle to attend the sessions because they had to arrive earlier to school or leave later than usual. Also, many of them said they perceived PODER as a mandatory program for psychological support for those who were underperforming in school. In other words, since it was linked with the word "therapy", PODER suffered from a considerable degree of stigmatization among students, reducing participation.