Technical Research Report

Sistema COC

A study of the effects of Sistema COC on student achievement in Brazil

20 March, 2019

Authors
Guido Gatti, MA
Kenneth Lee, PhD
Gustavo Alexandre Reis, MDes

Pearson

Table of contents

Pearson

Executive summary

Overview of Revel

Sistema COC is a learning system for Brazilian private school students attending kindergarten through 12th grade. Its mission is to prepare students for assessments administered by the National Institute for Educational Studies and Research (INEP) of the Brazilian Ministry of Education. The most important of these assessments is the ENEM, the main Brazilian college entrance exam. In 2001, COC introduced the Simulated ENEM (SimENEM) program, which allows schools and students to demonstrate their preparedness for the ENEM. The SimENEM is offered free of charge to all students in Brazil. A large number of students take the SimENEM, both from Sistema COC schools and from schools using other systems, including public schools.

Sistema COC integrates a cross-curricular instructional approach — covering literacy, math, natural science and social science — with print, digital, mobile and assessment assets. It also includes professional development and organizational tools for teachers. These aim to make sure resources are implemented in the best way possible, both inside and outside the classroom, and that they have the best possible impact on student achievement.

Intended outcomes

Here's what success looks like for Sistema COC, in terms of school, teachers' and learner experiences of using the system.

Outcomes related to access and experience

1. Learners have a positive learning experience.

Sistema COC incorporates a number of digital tools and platforms intended to provide a positive learning experience. Examples include COC Play, which lets learners access playful multimedia content that supports teaching materials; COC Portal, which provides personalized access to tools and materials, making the school routine easier, and Mangahigh, which uses creative and highly interactive games to make learning math more fun.

Outcomes related to competence/standard of achievement

2. Schools demonstrate improved performance in the national standardized exams

Sistema COC is based around a cross-curricular approach that reflects the four themes of the ENEM assessment: language, math, natural science and human science. The course materials and digital applications supplied as part of the system focus on the content covered in the ENEM exam.

Developed as part of Sistema COC, the SimENEM (SimENEM) program is offered annually to all students across Brazil, free of charge. The program helps schools and students show they are prepared for the real ENEM exam.

3. Learners perform better on the national SimENEM standardized exam

Sistema COC is designed to prepare learners for the ENEM exam and, as the name suggests, the SimENEM is designed to simulate the ENEM. This is true of the content and format of the questions offered in the SimENEM, but also the test-taking environment, rules and procedures. By preparing learners to perform better in the SimENEM, Sistema COC also prepares learners to perform better in the ENEM.

To judge whether Sistema COC achieves this goal, we can look at learners' recent SimENEM scores.

Research aims and research questions

This study examines how Sistema COC school students' achievement compares to students in other private and public schools not using Sistema COC. It analyzes and compares achievements using student scores on the 2018 SimENEM exam.

The research question we set out to answer is as follows.

1. Do high school grade students exposed to the COC learning system for one school year demonstrate higher achievement in the national COC SimENEM exam in language arts, math, human science and natural science compared to similar students not exposed to COC?

This question relates to intended outcome 3. – Learners perform better in the national SimENEM standardized exam.

Pearson

Key findings

In the context of this study, we can make two comparative statements about the efficacy of Sistema COC. Comparative statements are based on research that includes strict controls, but not strict enough to be deemed causal evidence. The research team equated comparison groups by matching examinees older than 18, on the previous year's (2017) SimENEM scores, grade level, number of people living in the household and years attending the same school[1].

- COC students statistically outperformed other similar non-COC public school attendees in both Language and Natural Science by a large margin (21 and 17 percentile points or 0.56 and 0.44 standard deviations, respectively). There were no significant differences in Math and Human Science.
- There were no significant differences in performance between COC students and non-COC private school students in Language, Math, Human Science and Natural Science.

Discussion

The results suggest that, after an additional school year, students in Sistema COC schools can be expected to perform better than their peers in public schools in the language and natural science areas of the SimENEM, and at least as well as other private school students in all areas of the SimENEM.

Limitations in the study design and methods, due to limits on the available data and restrictions on its use, only supported comparative efficacy statements and not causal statements. Another limitation is the extent to which the results generalize to the wider population of Brazilian school students, specifically public school students. The sample of students available for the primary analyses supporting the efficacy statements was only 1.8% of all students taking the 2018 SimENEM, making generalizability a concern.

1. The efficacy statements that come from the results of this study are considered comparative rather than causal for four main reasons:
   - Differences in Brazilian private and public schools cannot be equated through matching.
   - Half the COC examinees attended COC schools before 2017, and thus were exposed to the study treatment prior to baseline testing, precluding the matching or equating of study groups prior to treatment.
   - One of the matching variables, the number of people living at home, had group mean differences larger than 0.25 standard deviations after matching for the COC to public school comparison on language.
   - Examinees who reported they used COC but were public school students were omitted from the study, because all COC schools are private schools, which makes these attendees' reports inconsistent.

Next steps

While this study provides a sense of how student achievement compares between Sistema COC schools, non-Sistema COC private schools and public schools, it tells us nothing about the cause of the observed differences. To explain why Sistema COC students perform better in language and natural science tests than public school students, for example, we would need to investigate more closely how the system operates within schools. It seems reasonable to hypothesize that, among Sistema COC schools, those implementing the system with greater fidelity would achieve better results. This could be an area for future research. It is also unclear how well the SimENEM predicts performance in the ENEM. A necessary next step would be to further validate the SimENEM exam.

Pearson

Introduction

Background

The INEP of the Brazilian Ministry of Education is in charge of evaluating kindergarten to 12th grade (K–12) educational systems and the quality of education in Brazil. Schools in Brazil are monitored and their effectiveness judged by their students' performance in INEP assessments. The mission of Sistema COC, a learning system for K–12 Brazilian private school students, is to prepare students for these assessments.

The ENEM is the most important of the INEP assessments. Created in 1998 as an alternative to higher education entrance examinations, the ENEM is the main Brazilian college entrance exam, since 2009, and students with the highest scores win a place in the best universities in the country. For this reason, it is essential that school systems successfully prepare their students for the ENEM. To demonstrate preparedness for the ENEM, the SimENEM (SimENEM) program was developed by COC and introduced in 2001. Furthermore, the SimENEM is advertized throughout Brazil and is offered annually free of charge to all students across Brazil. A large sample of COC students take the SimENEM, as does a large proportion of non-COC private and public school students. This research attempts to compare COC and non-COC examinee performance in the 2018 SimENEM exam.

In Brazil, high school students, including those in private schools, are concentrated in urban areas, as is the population in general (86%). There is a large difference in the performance in the ENEM between private and public schools. In the 2015 ENEM, private high school examinees (n = 293,184) outperformed public school examinees (n = 1,151,664) by 0.9 to 1.0 standard deviations in the four subject area tests (language, math, natural science, human science). Also of interest, the majority (75%) of examinees had no school type listed (n = 4,309,690), and private school examinees outperformed this group by 0.64 to 0.87 standard deviations. Currently, about 6% of private schools in Brazil use COC, and there are several other education systems used by private schools.

Description of Sistema COC

Sistema COC is a learning system for Brazilian private school students attending kindergarten through high school. COC was first offered to K–12 students in 1978. Pearson acquired COC in 2010. In 2017, there were 456 COC schools across 294 cities supporting 167,600 students.

COC integrates a cross-curricular (literacy, math, natural science, social science) instructional approach with print, digital, mobile and assessment assets, as well as professional development and organizational tools for teachers. Ongoing professional development is aimed at providing the best implementation of these resources, both inside and outside the classroom, for the greatest impact on student achievement.

There are four different levels of COC that schools can choose to adopt – Intelligent, Brilliant, Genius and Licensed. Intelligent School is the introductory level, which the majority of schools adopt. Brilliant School is the next level up and provides additional professional development. Genius School provides greater support for teachers beyond the Brilliant School and is part of a larger network of COC schools organized to discuss topics related to the learning system. Licensed schools are offered the highest level of support and also carry the official COC-branded name.

How did research inform the design?

To help teachers instill the knowledge necessary to succeed in the ENEM assessment, the curricula and resources of Sistema COC incorporate numerous evidence-based principles in their design. While it is beyond the scope of this report to review the whole of Sistema COC and its evidence base, this section will review a set of core principles that underpin the pedagogical approach of Sistema COC from early childhood education through high school.

Formative assessment and metacognition

Assessment is often used to evaluate student achievement at the end of the learning process. However, assessment can also be used during the learning process to monitor progress and adjust teaching and learning according to results.

Using assessment to provide ongoing guidance during the learning process — commonly called formative assessment — can benefit learning, particularly when formative assessment is accompanied by timely, informative feedback that students can use to improve (Clark, 2012; Wiliam, 2011). Formative assessment can take many forms, from low-stakes quizzes to classroom activities. Sistema COC emphasizes formative assessment as a core component of its pedagogical approach and provides a variety of tools to support teachers in using formative assessment regularly in their teaching.

Pearson

Just as formative assessment can guide learning, so too can students' reflections on their own knowledge, thinking and progress toward learning goals. This 'thinking about thinking' (or metacognition) is an important component of students taking active and effective control of their own learning (Paris and Winograd, 1990).

Sistema COC provides metacognitive support through the inclusion of self-evaluation tools that teachers can administer to prompt students to engage in metacognition. While students may not always be completely accurate in assessing their own understanding (Serra and Metcalfe, 2009), the inclusion of self-evaluation tools nevertheless provides a means through which teachers can prompt students to engage in potentially beneficial metacognitive processes.

Interdisciplinarity

While traditional disciplines are typically taught in isolation, interdisciplinary approaches to education attempt to bring content from multiple disciplines together to encourage integration of disciplinary concepts and ways of thinking. For example, a unit of study can be organized around a central theme or problem that asks students to draw on content from multiple disciplines. Interdisciplinary learning is thought to promote higher-order thinking skills and synthesis of concepts to form more sophisticated conceptual understandings than might be achieved by studying each discipline in isolation (Jacobs, 1989; Mansilla, 2010).

Interdisciplinarity is a core component of Sistema COC's approach. For example, early childhood education materials include Explore Collection books organized around central interdisciplinary themes, and interdisciplinary projects are proposed at the end of each elementary school book.

Contextualization

Academic content can be taught in the abstract, removed from the real-world contexts to which it applies. However, research has shown that students' interest, persistence and achievement can be increased when they see the connection between the content and their interests, lives or a larger social purpose (Dweck, Walton and Cohen, 2014). For example, one study found that elementary school students' motivation and achievement in learning mathematical order of operations was improved by embedding an educational computer game in the context of space exploration, which presumably is of interest to many young students (Cordova and Lepper, 1996). The benefits of linking content to real-world contexts aligns with the theory that motivation depends in part on students seeing subjective value in their studies (Wigfield and Eccles, 2000). Sistema COC attempts to apply these benefits by continually demonstrating the practical applicability of content to students' lives. For example, the high school textbooks begin each chapter with a common everyday fact or problem and provide exercises that encourage students to relate the content to their daily lives.

In line with the general goal of demonstrating the relevance of content to the real world, Sistema COC takes a humanistic approach to science education. Humanistic approaches situate science instruction in the historical and social contexts in which science has been developed and is applied. While the goals of humanistic science education extend beyond simply demonstrating the relevance of science to the real world, this nevertheless is one of its benefits. Both qualitative and quasi-experimental research has found evidence that humanistic science teaching can improve students' attitudes toward science and can support the development of critical thinking skills and the ability to apply science to everyday events (Aikenhead, 2003).

Summary of the learner outcomes for Sistema COC

Sistema COC is a K–12 cross-curricular learning and assessment system that also offers professional development and organizational tools for teachers. The four main learner outcomes defined for COC are:

1. Learners will have a positive learning experience.
2. Schools will demonstrate better performance in the Prova Brasil (5th and 9th grade) and ENEM (high school graduates) national standardized exams.
3. There will be better student performance in the national SimENEM standardized exam.
4. COC partner schools will maintain high student flow – students progressing through to the next grade.

In the present study, we will be addressing learner outcome #3.

The present study

The purpose of the study is to collect and analyze data to determine if COC high school students demonstrate higher achievement than similar private and public school students who attend schools that don't use COC.

Research questions

The study will address the following research question:

- Do high school grade students exposed to the COC learning system for one school year demonstrate higher achievement in the national COC SimENEM exam in language arts, mathematics, human science and natural science compared to similar students not exposed to COC?

The analytic sample for the (primary) analyses that specifically address this research question includes only those students who were 18 at the time they registered for the SimENEM. The reason for focusing on 18-year-olds is that, legally, these are the only students in the sample who can consent to sharing the data needed to merge their 2018 SinENEM test scores with their 2017 test scores, critical for creating well-matched comparison groups. For students younger than 18, this type of merging was not possible. As a supplemental analysis including younger students, all available examinees' scores from the 2018 SimENEM were statistically compared between COC users and other examinee groups.

Implementation of product in study

To begin this section, it should be noted that probably half the COC examinees attended COC schools before 2017, and thus were exposed to COC before baseline testing. This precludes the matching or equating of study groups prior to treatment. For this reason, the efficacy statements that come from the results of this study are considered 'comparative' rather than causal. Another thing to note is that the efficacy statements (based on the matched analyses) offer the comparative effect of COC from a single school year of study, at the end of the students' high school years, and after taking the SimENEM at least once before. For these reasons, the estimated comparative effect of COC on students' preparedness for the ENEM given in this report may be considered conservative. The matched analyses make the study groups equal on prior year SimENEM, and as the analytic sample only includes examinees over 18, the full effect across high school years cannot be estimated.

Implementation of Sistema COC

A Sistema is an integrated learning system incorporating curriculum design, teacher support and training, print content and courseware, technology platforms, assessments and other services. COC is a such a system, provided to primary and secondary private schools in Brazil.

COC is offered in the following three stages:

- Segment 1: Educação Infantil: age 3–5, kindergarten
- Segment 2a: Anos Iniciais, Elementary 1: age 6–10, primary education, 1st to 5th grade
- Segment 2b: Anos Finais, Elementary 2: age 11–14 secondary education, 6th to 9th grade
- Segment 3: High school and PV (pre-vestibular or pre entrance exam): age 15–17, 10th to 12th grade

Students and teachers receive teaching materials (textbooks, activity books and notebooks) covering, according to the curriculum, culture, art, environment, science, math, reading/language arts (in Portuguese) and technologies. COC provides two options for schools: they may choose either the blue collection or yellow collection. The difference is that the blue is designed for 30 weeks of classes and the yellow is designed for 24 weeks.

The high school and PV (Segment 3) resources are delivered four times per academic year. During the school year, students receive 10 books if the schools choose the blue collection. For the yellow collection, students get eight books. In both cases, each book comes in two volumes (one for math and natural sciences and the other for language and human sciences). They focus on the content covered by the ENEM exam.

The two volumes of each book contain chapter theory and sequentially numbered activities (modules). The activity modules offer:

- Application exercises, which the teacher works through in class
- Extra exercises, which the teacher works through in class when there is time, otherwise they become part of the proposed task
- A study guide at the end of each module, which refers to the theoretical content studied before the exercises
- Proposed task – a series of individual student learning activities
- Differentiated skill-related exercises

The study guide, besides supporting the student's daily study at home, also supports the work of the teacher in preparing class instruction. Exercise resolutions are also available in COC Play/COC Tools.

Alongside the high school segment, there is PV, which provides ENEM preparatory course material. The PV module offers exercises aligned to the most recent ENEM exams and is updated every year. There are three main options offered:

- PV 1000 extensive is designed to last one academic year
- PV 500 semi-extensive is designed to be covered in six months or one year in fewer classes
- PV 300 intensive is designed for two months and is additional to PV 1000 and PV 500. The aim is to help students study for specific universities that don't use ENEM as an entrance exam. These universities offer their own entrance exams, usually two months after the ENEM.

The evaluation of learning in COC schools for various grade levels includes:

- 1st to 9th grades: assessments related to bimonthly content and simulated exams (Avaliação nacional)
- 9th to 12th grades and PV: SimENEM

COC also provides digital platforms to support teaching and learning:

COC PLAY

Through COC Play, students access all available products and services, from support materials to digital content. COC Play brings together everything they need in one place. Students can access customized environments and engaging content to support the teaching materials, such as exercise resolution videos and content such as videotapes, support materials and journalistic articles. The student also has online support from COC teachers, can use different educational solutions, and has access to digital content.

COC Tools

COC.Tools is the platform that enables school administrators to access relevant information to guide them in their daily demands, such as syllabus content, scheduling suggestions, pedagogical calendars, and price lists. Educators have electronic books, teaching materials, access to pedagogical purpose, on-line tutoring for teachers, continuing education, and tools peculiar to printed material and its guidelines.

MyClass

Classbuilder is a digital tool that helps the teacher build and personalize lesson presentations. The teacher can interact with the contents by means of multiple tools, such as paint brush, highlighter, arrows, eraser and, geometrical forms, as well as through animations and videos.

Pearson

COSMO

COSMO is an online digital platform offering individualized interactive instruction and assessment from mobile devices.

Specific digital learning applications include:

COC Academy – continuing education and professional development for administrators and teachers
3D Rooms – to immerse students in realistic and fascinating environments
Electronic books – all COC text material in PDF/Flash format.
Reading and Company – evaluation and reports to diagnose the individual development of students' reading abilities.
Simulated Exams
Events – COC network-wide training events to improve performance
TV COC – live broadcasting of COC events, videos on continued education, orientations and tutorials
Idiomas – tools to support bilingual teaching
Educational Journalism – lesson plans for news related to the competences and abilities necessary for the ENEM exam.
Virtual Lab – digital science lab
Mangahigh – interactive math games
Interactive Digital Board – multimedia resources through a touch interface

Pearson

Method

A quasi-experimental design was used to match individual 18-year-old students in the same grade on 2017 SimENEM test scores, on the number of people living at home and on the number of years attending the same school. The reason for focusing on 18-year-olds is that, legally, these are the only students in the sample who can consent to sharing the data needed to merge their 2018 SinENEM test scores with their 2017 test scores. For students younger than 18, this type of merging was not possible.

A research design such as this, that closely matches students on prior scores and other demographic variables, would typically provide causal evidence. Many of the examinees, however, attended COC schools before 2017, and thus were exposed to the study treatment (COC) before baseline testing. This precludes the matching or equating of study groups prior to treatment. For this reason, the efficacy statements are considered 'comparative' rather than causal. Something to note here is that the efficacy statements may be considered to offer conservative estimates of the comparative effect of COC on students' preparedness for the ENEM. The matched analyses make the study groups equal on prior year SimENEM, and as the analytic sample only includes examinees over 18, the full effect across high school years cannot be estimated.

As a supplemental analysis including younger students, all available examinees' scores from the 2018 SimENEM were statistically compared between private school students using COC and two other student groups, creating two comparisons of interest. The other two groups are:

- Private school students not currently using COC
- Public school students

It is important to compare private and public school students separately for two reasons:

1. True COC partner schools are private, as Sistema COC is not made available to public schools.
2. The difference between achievement test scores for private and public school students is substantial.

In addition to group status (the independent variable), other demographic student information was collected from each examinee (mandatory on the online registration form) and included in the statistical model to control for the effects of these factors. This information included:

- Grade level
- Whether students had practiced for the ENEM exam before
- Student rating of own math skills,
- Student rating of own Portuguese skills
- Number of people living at home
- Number of years attending the same school

Participants

The overall study sample consisted of high school students from all over Brazil, who took the 2018 SimENEM (n = 13,653). The analytic sample for the primary analyses, which matched students, included 248 students (COC = 163, non-COC private = 48, public school = 37). These examinees had taken the 2017 SimENEM and were 18 at the time of registration. Only those students fitting these criteria were included in these analyses.

The analytic sample for the secondary analysis, comparing COC students (n = 7,788) to non-COC students (public school = 2,892, non-COC private = 1,928) in the 2018 SimENEM included 12,608 students in all. The remaining 1,045 examinees reported they used COC but were public school students. This is inconsistent, and likely due to public high school graduates in their post-grad study using COC, high school students in schools using unsanctioned COC materials, or simply a misunderstanding of the question. Because of this uncertainty, we dropped these students from the analysis.

It is possible that the exclusion of these students, at least in part, is due to factors related to their achievement level. If this is the case, then they are not missing at random and their exclusion may bias the results of the COC to non-COC public school comparison. In estimating the true effect of COC in comparison to traditional public school students, this issue amounts to a trade-off between the potential bias in the results from excluding these students we are uncertain about, and the bias that may be introduced by including them as they do not appear to be typical. The research team can say that, when including all six factors in the statistical model mentioned in the previous section, the examinees reporting that they were public school COC students outscored the traditional or non-COC public school students by 0.129 standard deviations in the total SimENEM score. This means that, if the COC public school respondents could be included in the analysis, their presence would not have a substantial or very meaningful impact on the results.

It should also be noted that substantially more COC students take the SimENEM than non-COC students because they are encouraged to do so. It is offered at COC partner schools, which is convenient for them. In general, private school students are about 20% of the high school population. In the 2018 SimENEM however, they make up 70% of the examinees. There also exists a large achievement gap between public and private high school students, in the order of 30 percentile points (as seen in the 2015 ENEM). The 2018 sample of public school examinees, though still lower performing, were much closer in achievement in the 2018 language SimENEM with a 17 percentile point difference. This would suggest that the public school examinees were those from the upper range of the distribution. The examinees from the two private school groups, COC and non-COC, performed similarly in the 2018 SimENEM with a difference of 7 percentile points.

Data collection

All data collected for the study comes from the 2017 and 2018 SimENEM exams and the registration form that accompanied the 2018 exam. The 2018 SimENEM scores are the outcome measure of interest for this research effort. The SimENEM scores measure student academic performance. The SimENEM is advertised throughout Brazil, is open to all and is offered free of charge at hundreds of COC partner schools. Those interested in taking the 2018 exam simply needed to complete a registration form online. The 2018 registration form included several mandatory fields that provided examinee data necessary for the planned analyses. The SimENEM and registration form are explained in detail in the Measures section of this report.

The 2018 SimENEM online registration opened June 18 and closed on August 10, with over 33,000 students signing up to take the exam. The SimENEM testing took place on two consecutive Sundays, August 26 and September 2. In 2017, over 9,000 examinees took the SimENEM. Testing took place on two consecutive Sundays, August 27 and September 3. A total of 13,653 examinees took the 2018 SimENEM. Note that the school year in Brazil begins in February and ends in early December.

Measures

All data collected for the study comes from the 2017 and 2018 SimENEM exams and the registration form that accompanied the 2018 exam. Additional information was collected from the examinee registration form along with 2017 SimENEM scores if the 2018 examinee took the exam in 2017. The 2017 SimENEM scores are used as a matching variable and entered into the statistical model to equate examinees for prior SimENEM performance in the primary analyses. Likewise, the information collected from the 2018 registration form was also used as student level matching variables to equate examinees across the three study groups.

Pearson

It should be noted here that the self-assessment math and Portuguese responses collected from the registration form were not used in the primary matched comparison analyses to match examinees on academic performance. Instead, the actual prior year SimENEM was used, since it is a much more valid indicator for this purpose. The self-assessment responses were used in the secondary analysis of the 2018 SimENEM scores because they were available for all examinees. Only a small proportion of the 2018 examinees took the 2017 SimENEM and were 18 years old, and therefore could consent to the use of their personal identifying information necessary for the matching process.

Simulated ENEM

Since 2001, the SimENEM has been used to offer COC students practice for the very important national ENEM exam. The SimENEM was developed to closely mirror the structure and difficulty of the actual ENEM. The SimENEM is open and free of charge, to all Brazilians, and prizes are offered to students with the highest scores. Non-COC public and private students taking the SimENEM can be expected to be at the top of their classes. This means the three populations are expected to overlap considerably in achievement and provide a situation conducive to matching. Substantially more COC students take the SimENEM than non-COC students. This is because COC students are encouraged to take the test and it is offered at over 200 COC partner schools.

The SimENEM includes 185 multiple choice questions, measuring the same four subject areas as the ENEM. These are:

- Human science (50 questions in the areas of history, geography, philosophy and sociology)
- Language (45 questions about the Portuguese language, literature, foreign language, history of art, physical education, information technology and communication
- Mathematics (45 questions of general math and geometry)
- Natural science (45 questions in biology, physics and chemistry)

Reliability information is not publicly available from the vendor that offers the SimENEM. To address this gap, the research team was able to analyze data made available from the 2015 testing sessions for evidence of internal consistency (asking whether test questions measured similar knowledge). Coefficient alpha was calculated for each subject area test for both third year high school and post-graduate examinees. The results are reported here in Table 1. These internal consistency indices demonstrate the test scores are adequately reliable to compare group means. The research team could not estimate internal consistency for the 2017 and 2018 exams. Only individual scaled scores were made available and individual question responses are necessary to estimate internal consistency.

Pearson

Table 1: 2015 SimENEM internal consistency

Grade level	Subject test	N	Alpha
High school Year 3	Language	2,737	0.846
	Math	2,680	0.834
	Human science	2,851	0.837
	Natural science	2,695	0.730
Pre-vestibular	Language	431	0.753
	Math	412	0.861
	Human science	501	0.855
	Natural science	460	0.787

As expected, the subject area test scores correlated to each other, demonstrating validity evidence for the internal structure of the test. Math and language were the least correlated (r = 0.596). Human science and language scores were most highly correlated (r = 0.799). Not surprisingly, the second weakest correlation was between math and human sciences (r = 0.631). The remaining three correlations ranged between r = 0.670 and r = 0.720.

Additionally, examinees (older than 18) with both 2017 and 2018 SimENEM scores, demonstrated appropriate test-retest reliability. Test-retest reliability investigates if examinees occupy similar positions in the distribution of scores across the two years. The correlation between years ranged from 0.657 for human sciences and 0.769 for math (natural science r = 0.685, language r = 0.683).

There is no information available on how well the SimENEM predicts the actual ENEM, which would require longitudinal tracking of students through the ENEM. This, of course, is a necessary next step for validating the SimENEM exam as a predictor of ENEM performance. However, the above validity and reliability evidence suggests that the SimENEM is a suitable measure of achievement for the purposes of this study.

Pearson

2018 SimENEM online registration form

When registering for the 2018 SimENEM, potential examinees completed an online registration form. A total of 33,126 potential examinees completed the registration. This form asked for information necessary to the analyses, including:

- 2018 grade level
- School type (public or. private)
- Currently receiving COC instruction
- Have taken a COC mock exam before. This is actually interpreted as "Have you taken a practice exam?"
- Self-assessment of maths skills, self-assessment of Portuguese skills (rated as 'Best of my class', 'Good student' or 'Not good or bad student')
- Number of people living in the household (1 = live alone, up to 6 or more)
- Start date at current school (kindergarten, ES1, ES2, HS1, HS2, HS3. This information was used to calculate the number of years attending the current school)

Analysis methods

Two orthogonal contrasts were statistically tested for the 2018 SimENEM, the outcome of interest. The performance of examinees reporting they currently use COC was compared to the performance of non-COC private school and public school examinee groups. Comparisons are made between study groups using group mean differences. An ordinary least squares fixed effects model with robust empirical heteroskedasticity-consistent standard errors was employed to statistically test these group mean differences. This type of analysis is the default in the SPSS 25 General Linear Model procedure and the Stata CEM procedure. These contrasts were tested for each of the four content area subtests that comprise the SimENEM. In addition to group status, other examinee information was included in the statistical models to match the group differences for the effects of these factors on 2018 SimENEM scores. This information included:

- Grade level
- Whether students had practiced for the ENEM exam before
- Student rating of own math skills
- Student rating of own Portuguese skills
- Number of people living at home
- Number of years attending the same school

Pearson

Two separate sets of analyses were conducted. The first set, or primary analyses, provide the results that form the basis of the efficacy statements. This set of analyses, first matched individual 18-year-old students in the same grade on 2017 SimENEM test scores, on the number of people living at home and on the number of years attending the same school. These matching variables were also included in the statistical model to further equate pre-existing group differences remaining after matching. We refer to these results as 'doubly-robust', as they include further control for prior group differences. It is these doubly-robust estimates that are intended to form the basis of the efficacy statements. The Stata CEM (Coarsened Exact Matching) procedure was used for these analyses.

It should be noted that many of the examinees in the analytic sample attended COC schools before 2017, and thus were exposed to the study treatment (COC) before baseline testing. This precludes the matching or equating of study groups prior to treatment. For this reason, the efficacy statements are considered comparative rather than causal.

Secondary analyses included all available examinees' scores from the 2018 SimENEM. The statistical models used to test the group mean differences included the six variables mentioned above. These six variables were included in an attempt to adjust the groups for existing differences. In addition to testing group mean differences for each of the four content area subtests, total score differences were also tested. The SPSS 25 General Linear Model procedure was used for these analyses. These analyses are descriptive and do not support efficacy statements on their own.

Pearson

Matched student study analyses

The primary goal of the matching analysis was to assess the impact of being in a school that uses COC on student learning, as measured by student achievement across four disciplines (language arts, mathematics, human sciences, and natural sciences) on the SimENEM. Matching is a nonparametric method of preprocessing data to control for some or all of the potentially confounding influence of baseline factors by balancing or equating the treated and comparison groups. After preprocessing, various methods of analysis can be applied to estimate group differences. Effectively, matching attempts to identify 'virtual twins' between the treatment (COC) and comparison (non-COC) groups where, if successful, the only substantive difference between groups experiencing COC. In this study, the research team tried to achieve virtual twins by matching on prior performance, which is the same measure as the outcome variables, as well as two other factors thought to influence the outcome measure: school mobility and an indicator of socioeconomic status.

Given the vast socioeconomic disparities between Brazilian public and private school instruction, the study takes advantage of two comparison groups:

- Private school students not currently using COC (n = 48 / 248, 19%)
- Public school students not currently using COC (n = 37 / 248,15%)

Examinees from the treatment group, private school students who experienced COC (n = 163 / 248, 66%),
were matched and subsequently compared to the two comparison groups.

Individual COC students were matched to non-COC students using multiple variables:

- Grade level
- 2017 SimENEM subject area scores
- Number of people living at home
- Number of years attending the same school

The examinees' grade levels were always matched exactly. Examinees were matched separately for each of the four content area tests. This means that a COC student may have been matched to different non-COC students for language than for math. For example, COC students of the same
grade level were matched between the two non-COC groups using their 2017 SimENEM language score, the number of years that they attended their current or last school, and household size.
The same process was then repeated for math scores, then the natural and human science test scores. Descriptive statistics for the matching variables can be found in Table C1 for the full
sample as well as by group.

CEM was the specific matching procedure used (See gking.harvard.edu/cem for more information on CEM). CEM is a monotonic imbalance bounding matching method where observable characteristics of the study groups of interest are made equivalent or balanced by the researcher, based on theory rather than discovered through the process of checking after the fact and repeated estimations. This method also adjusts the imbalance on one variable while having no effect on the imbalance of any other variable. That is, the researcher assigns cut-off scores to the group or 'coarsens' each matching variable separately into levels that are deemed meaningful or comparable, based on their experience with the distribution of these variables and potential influence on the outcome (Iacus et al., 2011).

Preprocessing data with CEM was conducted using the CEM package in Stata 14. SimENEM 2017 scores were processed using the default automated coarsening routine. The default procedure uses Scott's normal reference rule, a common method used to find cut-off points for histograms for normally distributed data. Grade level in 2018 was matched exactly. Years attending the same school was preprocessed using cut-off points of 0.5 and 3.5. These points have the effect of grouping the last school change as being in: 1. E1 or middle school grades, 2. During high school years, and 3. Post high school (PV). Household size (the number of people living at home) was also preprocessed using cut-off points, specifically 0.5 and 5.5 people. These points were chosen to distinguish between students living on their own, with their immediate family, or living in an extended family situation. The coarsening generally worked well except in two instances out of a total of 24. These are described in the results section.

The specific levels across the four factors define a matching profile for each subject, then the procedure exactly matches subjects on these profiles. When at least one COC and one comparison subject share the same profile, the standardized mean group difference for the subjects in that profile is calculated and added to a weighted sum for all matched profiles. Subjects with unmatched profiles are dropped from the analysis and not included in the weighted sum. The resulting weighted group mean differences are statistically tested against zero using robust standard errors (as described in the Analysis methods section). Note that the 2018 SimENEM scores are standardized before unmatched cases are dropped from the analyses.

Additionally, doubly-robust analyses were employed where the matching variables were used, not only to create closely matched groups, but also to further equate pre-existing group differences remaining after matching. The doubly-robust analyses conform to the recommendations of the What Works Clearinghouse (WWC) standards for baseline equivalence (What Works Clearinghouse Standards Handbook Version 4.0, p14). The balance between the COC students and each of the comparison groups after preprocessing data with CEM can be found in Tables C4 and C5. Since not all the standardized group mean differences for each matching variable used in each analysis was less than 0.05, the WWC would recommend the doubly-robust models to support the efficacy claims.

Pearson

Methods for multiple comparisons

The statistical tests performed for both analyses were adjusted using the Benjamini Hochberg (BH) procedure (Benjamini and Yekutieli, 2001) for controlling the family-wise false discovery rate (FDR) at the 5% level. The p-values for each test were compared to the BH adjusted critical p-values to determine statistical significance. The FDR is simply the proportion of erroneous significant tests among all tests found statistically significant for a family or grouping of tests. The BH adjustment controls the FDR at a specified level for each family of tests, limiting the probability of declaring comparisons statistically significant when some or all tests are truly not (making type I errors). For this research, the families of tests are controlled at a FWFDR = 0.05. This means that we expect to only make a type I error in 1 in 20 families of tests.

The BH procedure is recommended by the WWC as suitable for most cases encountered in educational research. The WWC Procedures Handbook V4 quotes Benjamini and Yekutieli to make their point: "a modification of the original BH procedure could be made, although it is very often not needed, and yields too conservative a procedure".

The statistical tests performed for the primary or matched groups analyses were grouped into two families, one for each of the two comparisons. Recall these comparisons are: 1. COC to public school examinees and 2. COC to other private school examinees. Within these two families, BH corrections are made across the four tests of SimENEM content area outcomes. In essence, these two families of analyses test orthogonal matched samples from different populations. The study was designed this way from the outset.

Lastly, the 10 statistical tests for the secondary analysis constitute a different situation. Here the test statistics for the total scores and the content area subtests are certainly dependent. That is to say, if we determine there is a statistical difference in the groups for the total score, we can expect to find significance for one or more of the content area tests. These tests are not orthogonal. It is not likely that this dependency is uniform across all tests. The Benjamini Yekutieli (BY) adjustment procedure, though more conservative than the BH, can control the FDR at q = 0.05 across all 10 tests. The BY is a refinement to the BH that can properly adjust for statistical tests with arbitrary dependence.

Pearson

Results

The wide availability of the 2018 SimENEM allowed the research team to collect a large sample (n = 12,608) of reliable achievement data from high school students across Brazil. This data was used to statistically compare the performance of current Sistema COC students to other private school examinees not using COC, as well as to public school examinees. The research team also collected educational information from examinees when registering for the 2018 SimENEM in an attempt to provide for more robust experimental comparisons.

To this end, two separate sets of analyses were conducted to compare and statistically test group mean differences. The primary set of analyses, used to support the efficacy statements, first created matched groups of similar COC and non-COC examinees, then compared the groups. The secondary analyses used all available examinees' scores from the 2018 SimENEM and included the additional examinee information in the statistical models to best equate study groups. The results of these two sets of analyses are described here.

Matched analyses results

Do students exposed to the Sistema COC for one school year demonstrate higher achievement in the national SimENEM exam in language arts, mathematics, human science and natural science compared to a matched group of (similar) students not exposed to COC?

The primary goal of the matching analysis was to assess the impact on student learning of being in a school that uses COC. Using CEM, the data was preprocessed to match COC students to similar non-COC students from both private and public schools. Examinees were matched separately for each of the four content area tests. Descriptive statistics for the model variables can be found in Table C1. Table C2 and Table C3 provide information on how well matched groups aligned on each of the matching variables. Although there are some differences between the groups before preprocessing the data, matching students did provide better balance between the groups.

The differences between the matched study groups are summarized in Table 2. These differences are typically less than 0.25 standard deviations, within the range specified by the WWC. This is always the case for 2017 SimENEM, where the difference is always less than 0.09 standard deviations. Number of people living at home is the only variable showing any group mean differences larger than 0.25 standard deviations after matching. Specifically, there are two instances of note: the difference of 0.45 standard deviations for the COC to public school comparison on language, and 0.28 standard deviations for the COC to non-COC private school comparison on math.

Pearson

Table 2: Balance between matched study groups

Group	2018 SimENEM	2017 SimENEM	Years at the same school	Household size	Grade	Unmatched examinees COC/CP
		Mean diff. (Effect size)	Mean diff. (Effect size)	Mean diff. (Effect size)	Mean diff. (Effect size)	
COC private vs. non-COC public	Language	-3.96 (-0.04)	-0.05 (-0.01)	-0.45 (-0.41)	0	50/4
	Math	1.32 (0.02)	-0.26 (-0.07)	-0.06 (-0.06)	0	57/4
	Human science	-0.50 (-0.01)	-0.21 (-0.06)	-0.15 (-0.14)	0	53/2
	Natural science	0.26 (0.00)	0.03 (0.01)	-0.13 (-0.12)	0	55/5
COC private vs. non-COC private	Language	-8.00 (-0.09)	0.06 (0.02)	-0.12 (-0.11)	0	34/4
	Math	7.51 (0.09)	-0.07 (-0.02)	-0.28 (-0.26)	0	46/4
	Human science	3.43 (0.04)	-0.10 (-0.03)	-0.16 (-0.15)	0	39/6
	Natural science	1.95 (0.02)	-0.09 (-0.02)	-0.15 (-0.14)	0	31/5

Pearson

Table 3: Results for COC matched comparisons on 2018 SimENEM

Comparison	Test	Effect size[+] (%tile diff.)	Robust standard error	t	P-value	BH critical value	Sample size
Private COC vs. public	Language	0.552* (20.95)	0.192	2.88	0.005	0.0125	113/33
	Math	0.256 (10.10)	0.266	0.96	0.338	---	106/33
	Human science	0.198 (7.85)	0.232	0.85	0.396	---	110/35
	Natural science	0.406* (15.76)	0.167	2.44	0.016	0.025	109/32
Private COC vs. private non-COC	Language	-0.040 (-1.60)	0.147	-0.27	0.787	---	129/44
	Math	0.142 (5.65)	0.134	1.06	0.293	---	117/46
	Human science	-0.170 (-6.75)	0.155	-1.09	0.275	---	124/42
	Natural science	0.085 (3.39)	0.159	0.54	0.592	---	132/43

Note. The four matching variables include 2017 SimENEM, grade level, number of years attending current school and number of people living in the home.

[+]Positive effect sizes indicate COC group outperformed the non-COC group

Robust standard errors indicate the heteroscedasticity-consistent (HC) Huber–White 'sandwich' procedure was used to calculate standard errors for all statistical tests prior to the adjustments for multiple comparisons.

BH indicates the Benjamini Hochberg procedure for controlling the family-wise false discovery rate (FWFDR), here the FWFDR of q = 0.05 is controlled across SimENEM content domains within each of the two orthogonal COC comparisons.

* $p <$ BH adjusted critical p-value

Pearson

Table 4: Results for COC matched doubly-robust comparisons on 2018 SimENEM

Comparison	Test	Effect size[+] (%tile diff.)	Robust standard error	t	P-value	BH critical value	Sample size
Private COC vs. public	Language	0.558* (21.16)	0.228	2.45	0.016	0.025	113/33
	Math	0.279 (10.99)	0.181	1.54	0.126	---	106/33
	Human science	0.177 (7.02)	0.193	0.92	0.360	---	110/35
	Natural science	0.444* (17.15)	0.164	2.70	0.008	0.0125	109/32
Private COC vs. private non-COC	Language	0.009 (0.36)	0.118	0.07	0.942	---	129/44
	Math	0.086 (3.43)	0.087	0.98	0.326	---	117/46
	Human science	-0.210 (-8.32)	0.087	-2.40	0.018	0.0125	124/42
	Natural science	0.112 (4.46)	0.119	0.94	0.346	---	132/43

Note. The four matching variables include 2017 SimENEM, grade level, number of years attending current school, and number of people living in the home. The doubly-robust statistical models include all matching variables as covariates to further adjust the effect sizes for remaining differences in the matched groups.

[+]Positive effect sizes indicate COC group outperformed the non-COC group

Robust standard errors indicate the heteroscedasticity-consistent (HC) Huber-White 'sandwich' procedure was used to calculate standard errors for all statistical tests prior to the adjustments for multiple comparisons.

BH indicates the Benjamini Hochberg procedure for controlling the family-wise false discovery rate (FWFDR), here the FWFDR of q = 0.05 is controlled across SimENEM content domains within each of the two orthogonal COC comparisons.

* $p <$ BH adjusted critical p-value

Pearson

Table 3 provides information for the simple COC to non-COC group comparisons, while Table 4 provides information for the doubly-robust comparisons. Recall that the doubly-robust comparisons include the matching variables in the statistical models to further equate matched groups for any remaining differences. The doubly-robust results provide the basis for the efficacy statements.

We can see that after preprocessing students with CEM, COC students performed statistically significantly better than their non-COC public school counterparts in language (21 percentile points, or 0.56 standard deviations) and natural sciences (17 percentile points, or 0.44 standard deviations). COC students also had higher scores in math and human sciences but these differences were not statistically significant.

After again preprocessing and matching students to test the comparison between COC and non-COC private school students, there were no significant differences between the groups, see Table 3. The doubly-robust models also revealed non-significant differences, as shown in Table 4. The full results for the COC to non-COC private school comparisons can be found in Appendix E.

Full sample 2018 SimENEM analyses results

The full sample analyses compared 7,788 COC students to 1,928 non-COC private school students and 2,892 public school students. After analyzing the SimENEM total scores for all the examinees taking the 2018 SimENEM, COC students statistically outperformed (Type I error rate = 0.05) other non-COC private school attendees by a small amount, that is19 scaled score points, 3 percentile points, or 0.07 standard deviations. COC students, however, outperformed public school attendees by a wide margin (183 scaled score points, 25 percentile points, or 0.67 standard deviations). Furthermore, when we analyzed scores separately by subtest, COC students statistically outperformed non-COC private school attendees in language, math, and natural science (the difference in human science was not statistically significant), and outperformed public school attendees across all four subject areas including human science.

It is important to note that these analyses included additional examinee information in the statistical models to equate the group differences (explained in the Analysis Methods section) for the effects of these factors on 2018 SimENEM score. Tables 5 and 6 give the statistical significance tests and effect sizes for the equated group mean differences. The group differences without the addition of this information has the non-COC private school examinees performing the best, with the COC examinees second (-0.18 standard deviations), and the public school examinees last (-0.45 standard deviations versus COC). This trend is, of course, changed to COC performing the best when group mean differences are made conditional on the additional available examinee information.

Pearson

Table 5: Results for COC vs. non-COC public model adjusted comparisons on 2018 SimENEM

2018 SimENEM	Group mean difference	Robust standard error	t	P-value	Effect size (%tile diff.)
Total score	182.930	5.802	-31.528	0.000	0.669* (24.822)
Language	57.990	2.027	-28.612	0.000	0.637* (23.802)
Math	37.537	1.521	-24.671	0.000	0.541* (20.568)
Human sciences	44.291	1.908	-23.212	0.000	0.514* (19.623)
Natural sciences	43.113	1.489	-28.959	0.000	0.641* (23.930)

Robust standard errors indicate the heteroscedasticity-consistent (HC) Huber–White 'sandwich' procedure was used to calculate standard errors for all statistical tests prior to the adjustments for multiple comparisons.

BY indicates the Benjamini Yekutieli procedure for controlling the family-wise false discovery rate (FWFDR), here the FWFDR of q = 0.05 is controlled across the 10 SimENEM total score and content domains across the two orthogonal COC comparisons.

* $p < 0.0017$, the first BY adjusted critical p-value

Pearson

Table 6: Results for COC vs. non-COC private model adjusted comparisons on 2018 SimENEM

2018 SimENEM	Group mean difference	Robust standard error	t	P-value	BY critical value	Effect size (%tile diff.)
Total score	18.838	6.543	-2.879	0.004	0.014	0.067* (2.660)
Language	6.996	2.195	-3.188	0.001	0.012	0.076* (3.042)
Math	4.283	1.701	-2.518	0.012	0.015	0.059* (2.371)
Human sciences	0.564	2.126	-0.265	0.791	---	0.006 (0.252)
Natural sciences	6.995	1.701	-4.112	0.000	0.010	0.101* (4.004)

Robust standard errors indicate the heteroscedasticity-consistent (HC) Huber-White 'sandwich' procedure was used to calculate standard errors for all statistical tests prior to the adjustments for multiple comparisons.

BY indicates the Benjamini Yekutieli procedure for controlling the family-wise false discovery rate (FWFDR), here the FWFDR of q = 0.05 is controlled across the 10 SimENEM total score and content domains across the two orthogonal COC comparisons.

* p < the BY adjusted critical p-value

See Tables B1–B4 for:

Descriptive statistics comparing COC and non-COC private school group 2018 SimENEM scores
A breakdown of the total score statistical model effects Statistical significance tests Model parameters

Tables B5–B8 give the same information for the statistical model, but also compare COC to public school examinees.

Efficacy statements

In the context of this study, we can make two comparative statements about the efficacy of Sistema COC. Comparative statements are based on research that includes strict controls, but not strict enough to be deemed causal evidence. The research team equated comparison groups by matching examinees older than 18, on the previous year's (2017) SimENEM scores, grade level, number of people living in the household and years attending the same school[2].

- COC students statistically outperformed other similar non-COC public school attendees in both Language and Natural Science by a large margin (21 and 17 percentile points or 0.56 and 0.44 standard deviations, respectively). There were no significant differences in Math and Human Science.
- There were no significant differences in performance between COC students and non-COC private school students in Language, Math, Human Science and Natural Science.

2. The efficacy statements that come from the results of this study are considered comparative rather than causal for four main reasons:
   - Differences in Brazilian private and public schools cannot be equated through matching.
   - Half the COC examinees attended COC schools before 2017, and thus were exposed to the study treatment prior to baseline testing, precluding the matching or equating of study groups prior to treatment.
   - One of the matching variables, the number of people living at home, had group mean differences larger than 0.25 standard deviations after matching for the COC to public school comparison on language.
   - Examinees who reported they used COC but were public school students were omitted from the study, because all COC schools are private schools, which makes these attendees' reports inconsistent.

Pearson

Further robustness checks

A potential concern with matching analysis is a potential for a decrease in sample size and statistical power when matching on multiple covariates, significant relationships may not be detected. One potential robustness check is to reduce the number of matching variables to the most important ones based on theory. A set of robustness checks was conducted that replicated the main analysis using CEM and doubly-robust methods where students were matched on only the two most important variables: prior SimENEM scores and grade. Matching on two variables resulted in a larger sample size but similar results to the main analysis with four matching variables.

On average, COC private school students performed significantly better than their non-COC public school counterparts in language (β = 0.863, SE = .150, p = 0.000) and natural sciences (β = 0.347, SE = 0.174, p = 0.047) in the bivariate regressions and again in language (β = 0.897, SE = 0.240, p = 0.000) and natural sciences (β = 0.375, SE = 0.159, p = 0.019) when using doubly-robust methods. Also, when comparing COC private school students with non-COC private school students, there were no significant differences in performance in the bivariate or doubly-robust methods. The results can be found in Tables D1 and D2.

Given the similarities in results between matching on two variables (found in Tables D1 and D2) and on four variables in the main analysis (found in Tables C2 and C3), it can be assumed that the main analysis presents a more conservative estimate of the effect of COC on students.

Another potential concern lies with the use of CEM, given the large breadth of various types of statistical matching techniques that one can employ to make causal claims. It is possible that one matching technique will yield significant relationships while a similar one would not. To account for a potential inconsistency between various matching techniques, inverse-propensity-weighted regression adjustment estimators (IPWRA) were used.

IPWRA calculates the propensity that a student would receive COC from the matching variables and generates weights accordingly. IPWRA was selected due to its capabilities for performing a similar doubly-robust process found in CEM.

The robustness check using IPWRA uses three of the four matching variables used in CEM: prior achievement, grade, and years in the same school. Household size was not included due to overlap violations. The replication analyses using IPWRA yielded very similar results to those in the main analyses. COC private school students were found to have performed significantly better than their non-COC public school counterparts in language (β = 0.717, SE = .152, p = 0.000) and natural sciences (β = 0.456, SE = 0.141, p = 0.001). Also, when comparing between COC private school students and non-COC private school students, there were no significant differences in performance in the bivariate or doubly-robust methods. The results can be found in Table D3. Summary statistics of the differences between the groups before and after weighting with inverse propensity scores can be found in Tables D4–D6.

Given the similar results between those found in the main analyses and the robustness checks found in Appendix D (CEM with fewer matching covariates and replication analysis with IPWRA), the main matching analyses illustrate a set of conservative and robust results that replicate across different model specifications and matching techniques. Based on the findings across the different models, COC had a consistently positive impact on language and natural science test scores when comparing COC private school students and non-COC public school students.

Pearson

Discussion

Very little research has been conducted on the merits of the Sistema COC used by private schools in Brazil. In fact, no recent scientific research could be found that evaluates COC's impact on student achievement. High schools in Brazil are monitored by the Brazilian Ministry of Education and their effectiveness judged by student performance on the ENEM. The ENEM is an optional Brazilian college entrance exam, and students with the highest scores win a place in the best universities in the country. To combat the misuse of ENEM scores, the Ministry of Education announced in 2017 that they would no longer report ENEM results that could be linked to schools. This development would seem to make it difficult to research the impact of COC.

Fortunately, the widely available 2018 SimENEM allowed the research team to collect a large sample (n = 13,653) of reliable achievement data from high school students across Brazil. This data was used to statistically compare the performance of current Sistema COC students to other private school examinees not using COC, as well as public school examinees. The research team also collected educational information from examinees registering for the 2018 SimENEM in an attempt to provide more robust comparisons.

Two separate sets of analyses were conducted. The first set, or primary analyses, provides the results that form the basis of the efficacy statements. This set of analyses first matched individual 18-year-old students in the same grade on 2017 SimENEM test scores, on the number of people living at home and on the number of years attending the same school. These matching variables were also included in the statistical model to further equate pre-existing group differences remaining after matching. The secondary analyses included all available examinees' scores in the 2018 SimENEM and statistically adjusted group mean differences with available information from their registration forms. As this analytic sample included students under the age of 18, it was not possible to include 2017 SimENEM scores. In place of these, a simple self-reported proxy measure was used (see Measures section for details).

The estimated effect sizes from the two sets of analyses largely agreed. They only substantially differed in that the effect comparing COC to non-COC private students for human sciences was negative for the doubly-robust matched sample, while the full sample effect size was nearly zero. There were two main differences in the primary versus secondary analysis, which may explain the difference in results between the matched and the full samples. In particular, the matched sample was statistically equated for pre-existing differences in achievement of the two groups by matching on 2017 scores. We know from the test-retest analysis that 2017 scores were highly salient predictors of 2018 scores. It could be that only by controlling for these prior achievement differences can we observe the true contribution of the COC program to 2018 scores. In addition, the matched sample only included 18-year-old repeat test-takers, whereas the full sample included first-time test-takers and students in other grades.

Conclusion

The doubly-robust match group comparisons suggest that COC students, after an additional school year, can be expected to perform better than similar public school students in language and natural science. The results further suggest that COC students would be expected to perform at least as well as other private school students in the SimENEM. By 'similar', we mean students from the same grade that demonstrated very close previous year achievement in each subject area in the SimENEM, the very measure being compared.

Limitations of the study and generalizability of the findings

For four main reasons, the efficacy statements that come from the results of this study are considered comparative rather than causal, Comparative statements include strict controls, but not strict enough, according to the What Works Clearinghouse, to be deemed causal evidence. Recall that the primary or matched groups analyses form the basis for the efficacy statements. The sample of students used in these analyses included those examinees over 18 who took the 2017 SimENEM before taking the 2018 SimENEM.

The four reasons are:

1. COC schools are private schools, and as such, there are no students from COC-using public schools that use COC to match to the students in the COC private school treatment group. This means that the substantial differences in Brazilian private and public schools cannot be equated through matching as prior SimENEM performance was. Thus, we cannot know how much of the comparative difference in the COC and public school groups was due specifically to school-related factors, such as the quality of instruction or availability of technology, and not simply the use of COC.

The research team chose the best option available to them, which was comparing COC to other private and public school students separately. In this way, the fairest possible comparisons could be made.

2. Probably half the COC examinees attended COC schools before 2017, and thus were exposed to the study treatment prior to baseline testing. This precludes the matching or equating of study groups prior to treatment. The absence of a true baseline, before the start of the treatment, means the full effect of COC cannot be estimated. The efficacy statements suggest the comparative effect of COC from a single school year of study, at the end of the students' high school years, and after already taking the SimENEM at least once.

3. One of the matching variables, the number of people living at home or household size, had group mean differences larger than 0.25 standard deviations after matching. Specifically, there are two instances of note: the difference of 0.45 standard deviations for the COC to public school comparison on language, and 0.28 standard deviations for the COC to non-COC private school comparison on math. These are the only two instances out of a total of 24. All matching variables were included in the statistical model to further equate pre-existing group differences remaining after matching. These doubly-robust results formed the basis of the efficacy statements.

4. Lastly, examinees who reported they used COC but were public school students were omitted from the study. COC partner schools are exclusively private schools, so reporting both COC and public is inconsistent. This is likely to be due to public high school graduates in their post-graduation ENEM study using COC. Nearly all these examinees reported they were in their pre-vestibular study (63/67 = 94%). This group scored almost exactly between the COC private and non-COC public groups in the 2017 SimENEM: 0.35 standard deviations above the non-COC public group and 0.34 standard deviations below the COC private group. This evidence would suggest these examinees could be a fourth group for study.

   If these examinees should be included with either the COC private or non-COC public groups, their exclusion may bias the results of the COC private to non-COC public school comparisons. The research team sees this issue as a trade-off between the potential bias in the results from excluding these students we are uncertain about, and the bias that may be introduced by including them, as they do not appear to fit into our current groups.

Another limitation of the current study is a concern over its generalizability. This is specifically the generalizability of the sample used to generate the efficacy statements to the wider population of Brazilian high school students. The sample of students available for matching (and thus supporting the efficacy statements) was a tiny fraction of all students taking the 2018 SimENEM (248/13,653 = 1.8%). The COC group made up the majority of these 248 examinees and the two comparison groups were relatively small (COC = 163, non-COC private = 48, public school = 37), making generalizability even more of a concern (i.e., both less than 1%).

Recall that these students needed to be both 18 at the time of their 2018 SimENEM registration to provide consent for the research, and to have scores for the 2017 SimENEM. It should be noted that these 248 examinees come from 118 different schools. This issue is a classic trade-off between research rigor and protecting the subjects versus wider generalizability.

As with any research study that uses a matched groups design, the goal is to create artificially similar groups prior to treatment, preparing them for fair comparison after treatment is applied. These matched groupings can have the effect of limiting the generalizability of the research findings. In this case, top performing public school students were selected for one matched group. These public school students are unlikely to represent the entire population of Brazilian public school students. There exists a very large achievement gap between public and private high school students, in the order of 30 percentile points (as seen in the 2015 ENEM). Shockingly, this means that the average private school student is expected to score at the 80th percentile of the public school students' distribution.

The fact that students will have taken the SimENEM at least once before may also limit generalizability to the wider public school population. This is not much of a concern for the private school population, as nearly all private school students practice the ENEM in some fashion during their high school years. On the other hand, taking the SimENEM is less common among public school students. In general, public school students are about 80% of the high school population and the same number took the ENEM in 2015, but only 30% of the 2018 SimENEM examinees were public school students. Therefore, public school students who have taken the SimENEM multiple times are likely to be among the top-performers in their schools, as this suggests they are serious about attending college. That does not mean that the sample of public school SimENEM-takers score as highly as private school students. In fact, they do not. The average private school student scores at the 66th percentile of the public school students' 2018 SimENEM distribution.

To further understand the type of students in the analytic sample, we contrast them to both the larger samples of all 2018 SimENEM examinees (n = 13,653) and the sample of people that registered for the SimENEM exam (n = 33,126). Table E3 shows the percentages side-by-side for each sample, within each study group, across the levels of seven variables. The seven variables come from fields on the registration form that all registrants were asked to provide.

The variables were:

- Self-assessment math (orange dots in Figure 1)
- Self-assessment Portuguese (yellow dots)
- Number of people living at home (green dots)
- When did you start school? (pink dots)
- Have you failed to graduate? (blue dots)
- Have you left school during the school year and not returned that year? (red dots)
- What will you do after finishing high school? (large black dots)

Two comparisons were made for three groups at 26 levels across the seven variables for 156 total comparisons. In Table E3, cells are highlighted red if the percentage observed for the larger samples are outside the exact binomial 95% confidence interval around the percentage observed for the analytic sample. It should be noted that a confidence interval could not be calculated for 18 of the 78 analytic sample percentages. For this reason, we also calculated and plotted the differences in percentages.

In addition to Table E3, we provide Figure 1. The figure shows the differences for the comparisons between the largest sample (people registering for the 2018 SimENEM) and the analytic sample by the observed percentages in the analytic sample. We plot the differences against the observed percentages because, as the value range for percentages is bounded, differences in the middle of the range are expected to be larger and will have larger standard error. We can see from Figure 1 that most differences fall within plus or minus 10%. The exception being the obvious large black dots, which we discuss next.

The one area in which the analytic sample stands out as substantively different is in what the students expect to do after finishing high school. Here, more members of the analytic sample expect to be able to continue their education without needing to work. The analytic sample differs from the other two samples in this expectation for all three groups, including the public school students. It would seem that these students tend to perceive themselves at a higher socioeconomic level or expect to receive scholarships. The difference in percentages for the comparisons on this variable are shown in Figure 1 as the large black dots. Also, the comparisons that do not fall within the 95% confidence interval are highlighted red in Table E3.

The only other instance of a comparison falling outside the 95% confidence interval was within the COC group on the variable *'H*ave you failed to graduate?'. The COC analytic sample was a little more likely to have failed a grade than the students in the registration sample.

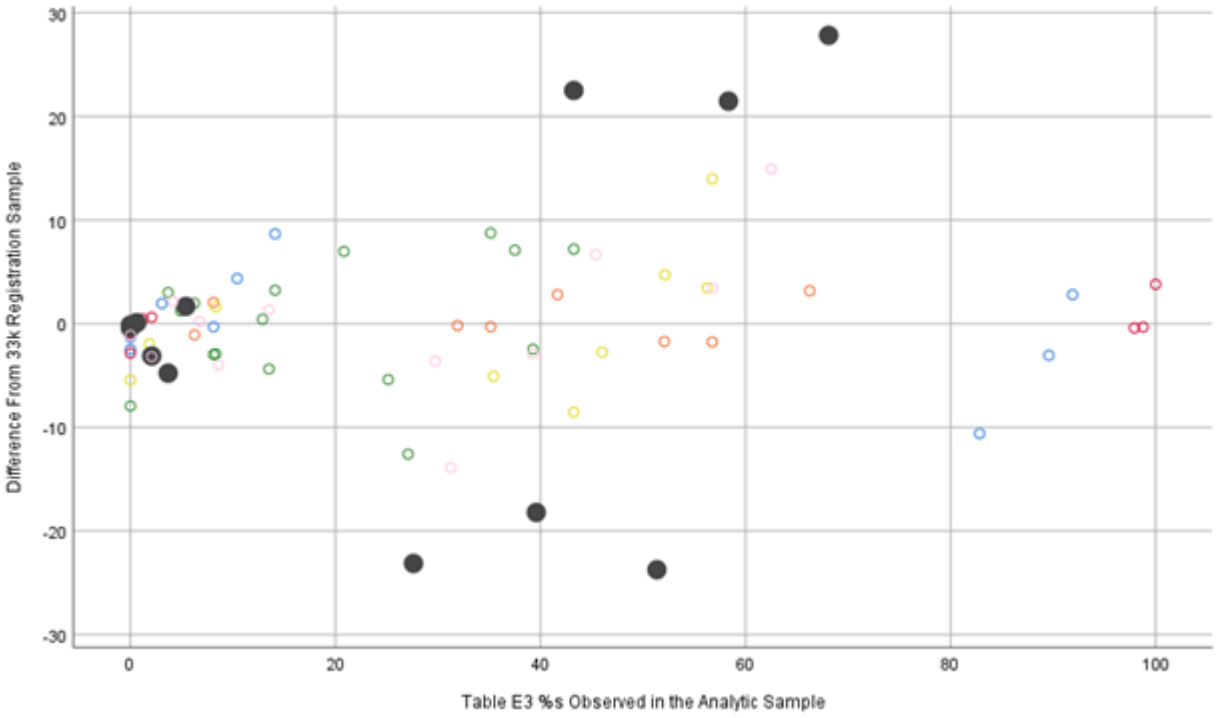

Figure 1: Percentage differences in analytic and 2018 SimENEM registration sample by percentages observed for characteristics in the analytic sample

Pearson

Further limitations that could be noted are that the collection of variables used to match on was relatively small. There are undoubtedly other variables (both observed and unobserved) that we were not able to control for and which could confound the results. These include student motivation and study habits. Also, there is no information available on how well the SimENEM predicts the actual ENEM. The research team does not have the capacity to follow up on students taking the SimENEM and collect their subsequent ENEM scores. This, of course, is a necessary next step for validating the SimENEM exam. At this point, the research team would caution against comparing SimENEM and ENEM scores directly.

Implications of findings for product implementation and further research

Although this study provides some sense of how student achievement compares across COC schools, non-COC private schools and public schools, it tells us nothing about why observed differences exist. In order to explain how COC students perform better in language and natural science than public school students and worse in human sciences than other private school students, we would need to investigate more closely how the program operates within the schools. We have essentially ignored implementation fidelity of the COC program in this study. It seems reasonable to hypothesize that schools implementing the COC program with greater fidelity would have better results than those that do not. This could be an area for future research.

Pearson

References

Aikenhead, G. S. (2003, August 19-23). *Review of research on humanistic perspectives in science curricula*. Paper presented at the European Science Education Research Association 2003 Conference, Noordwijkerhout, The Netherlands. Retrieved from https://education.usask.ca/documents/profiles/aikenhead/ESERA_2.pdf.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics, 29*(4), 1165-1188.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review, 24*(2), 205-249.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88*(40), 715-730.

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2014). *Academic Tenacity: Mindsets and Skills that Promote Long-Term Learning*. Seattle, WA: Bill & Melinda Gates Foundation.

Iacus, S. M., King, G., Porro, G. (2011) Multivariate matching methods that are monotonic imbalance bounding. J*ournal of the American Statistical Association*, *106*(493), 345-361.

Jacobs, H. H. (Ed.) (1989). *Interdisciplinary curriculum: Design and implementation*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mansilla, V. B. (2010). Learning to synthesize: The development of interdisciplinary understanding. In R. Frodeman, J. Thompson Klein, & C. Mitcham (Eds.), *The Oxford handbook of interdisciplinarity* (1st ed.). Oxford, UK: Oxford University Press.

Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15-52). Hillsdale, NJ: Lawrence Erlbaum Associates.

Serra, M. J., & Metcalfe, J. (2009). Effective implementation of metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition* (pp. 278-298). New York, NY: Routledge.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*, 3-14.

What Works Clearinghouse Standards Handbook Version 4.0

Pearson

Appendix A. Data cleaning process

Global Schools research team received 2018 SimENEM (SimENEM) registration data from vendor (Dataeduc).

- Of the N = 33,126 potential examinees registered, 13,653 had 2018 SimENEM scores

Research team received 2017 SimENEM scores from vendor for 9,678 examinees.

Research team matched (by name) examinees with 2017 SimENEM scores to people in 2018 registration datafile.

- Found and merged N = 2,570 examinees with 2017 SimENEM scores also registered for 2018 SimENEM
  - 50 of these cases had repeated taxpayer IDs and these were omitted
- Examinee registration data included examinees' taxpayer ID and this unique ID was used to merge 2018 SimENEM scores

Examinee consent and personal identifying information (PII) procedures.

- 2018 SimENEN score file was scrubbed of PII by Brazil research lead before being sent to research team
  - These scores were used to compare COC and non-COC students
- Second file of 2018 SimENEN scores for those examinees 18 years or older (age of consent in Brazil), not scrubbed of PII was sent to research team
  - These scores were used to make matched comparison groups using SimENEM 2017 scores
  - PII was necessary to match individual examinees' scores

Number of examinees that had SimENEM scores in 2017 and 2018 and were 18at the time of registration

- COC n = 163,
- Non-COC public n = 37,
- Non-COC private n = 48)

Pearson

Appendix B. COC vs. non-COC private and public results tables

Table B1. COC vs. non-COC private 2018 SimENEM descriptive statistics

Study group	Count	Language	Human science	Natural science	Math	Total
		Mean (Standard deviation.)	Mean (Standard deviation)	Mean (Standard deviation)	Mean (Standard deviation)	Mean (Standard deviation)
Non-COC private	1928	532.94 (88.87)	493.67 (87.21)	408.12 (70.54)	396.18 (73.17)	1830.91 (283.76)
COC	7788	520.86 (92.20)	473.45 (89.26)	400.86 (69.27)	383.80 (71.52)	1778.96 (280.99)
Total	9716	523.25 (91.67)	477.46 (89.22)	402.30 (69.58)	386.25 (72.01)	1789.27 (282.29)

Pearson

Table B2. COC vs. non-COC private 2018 SimENEM total score statistical model fixed effects

Model effect	Level	N
Grade	9th grade	859
	1st year high school	1632
	2nd year high school	2161
	3rd year high school	3954
	Pre-vestibular	1110
Have you taken a SimENEM exam before?	No	5236
	Yes	4480
Self-assessment math	Best of my class	588
	Good student	3479
	Not good or bad student	5649
Self-assessment Portuguese	Best of my class	450
	Good student	4758
	Not good or bad student	4508
Number of people living at home	1	56
	2	943
	3	2960
	4	4186
	5	1232
	6	339

Pearson

Table B3. COC vs. non-COC private 2018 SimENEM total score statistical significance tests

Model effect	Type III sum of squares	df	Mean square	F	Sig
Corrected model	258488024.743[a]	16	16155501.546	303.864	0.000
Intercept	2637850586.555	1	2637850586.555	49614.603	0.000
COC vs. non-COC private	461367.202	1	461367.202	8.678	0.003
Grade	148814362.335	4	37203590.584	699.75	0.000
Have you taken a SimENEM exam before?	340353.736	1	340353.736	6.402	0.011
Self-assessment math	71263021.150	2	35631510.575	670.183	0.000
Self-assessment Portuguese	11498442.151	2	5749221.076	108.136	0.000
Number of people living at home	717452.621	5	143490.524	2.699	0.019
Number of years attending same school (covariate)	1186898.589	1	1186898.589	22.324	0.000
Error	515664975.596	9699	53166.819		
Total	31879786808.134	9716			
Corrected total	774153000.339	9715			

a. R squared = .334 (adjusted R squared = .333)

Pearson

Table B4. COC vs. non-COC private 2018 SimENEM statistical model parameters

Model effect	B	Robust standard error[a]	t	Sig	95% Confidence interval	
					Lower bound	Upper bound
Intercept	1936.667	15.477	125.134	0.000	1906.330	1967.005
Non-COC private	-18.838	6.543	-2.879	0.004	-31.664	-6.011
COC	0[b]					
Grade = 9	-555.052	11.584	-47.917	0.000	-577.758	-532.346
1st year high school	-433.449	10.426	-41.574	0.000	-453.886	-413.012
2nd year high school	-323.946	9.999	-32.399	0.000	-343.546	-304.347
3rd year high school	-226.779	9.291	-24.408	0.000	-244.992	-208.567
Grade = pre-vestibular	0[b]					
Have you taken a SimENEM exam before = no	-13.085	5.209	-2.512	0.012	-23.296	-2.874
Have you taken a SimENEM exam before = yes	0[b]					
Self-assessment math = best of my class	262.672	10.812	24.294	0.000	241.478	28.866
Self-assessment math = good student	158.088	5.101	30.994	0.000	148.090	168.087
Self-assessment math = not good or bad student	0[b]					
Self-assessment Portuguese = best of my class	121.401	12.212	9.941	0.000	97.463	145.339
Self-assessment Portuguese = good student	61.034	4.811	12.685	0.000	51.603	70.466

Self-assessment Portuguese = not good or bad student	0[b]					
Number of people living at home = 1	38.129	35.111	1.086	0.278	-30.696	106.954
Number of people living at home = 2	10.929	15.292	0.715	0.475	-19.047	40.905
Number of people living at home = 3	27.649	13.653	2.025	0.043	0.886	54.412
Number of people living at home = 4	27.649	13.459	2.054	0.040	1.266	54.031
Number of people living at home = 5	9.169	14.547	0.630	0.529	-19.345	37.683
Number of people living at home = 6	0[b]					
Years attending same school (covariate)	3.171	0.681	4.656	0.000	1.836	4.505

a. Heteroskedastic-consistent HC3 method

b. This parameter is set to zero because it is redundant.

Pearson

Table B5. COC vs. Non-COC public 2018 SimENEM descriptive statistics

Study group	Count	Language	Human science	Natural science	Math	Total
		Mean (Standard deviation)	Mean (Standard deviation)	Mean (Standard deviation)	Mean (Standard deviation)	Mean (Standard deviation)
Non-COC public	2892	479.95 (80.44)	446.84 (74.11)	369.12 (55.14)	359.99 (60.06)	1655.90 (229.24)
COC	7788	520.86 (92.20)	473.45 (89.26)	400.86 (69.27)	383.80 (71.52)	1778.96 (280.99)
Total	10680	509.78 (91.00)	466.25 (89.22)	392.26 (67.24)	377.35 (69.41)	1745.64 (273.48)

Pearson

Table B6. COC vs. non-COC public 2018 SimENEM total score statistical model fixed effects

Model effect	Level	N
Grade	9th grade	846
	1st year high school	1616
	2nd year high school	2368
	3rd year high school	4718
	Pre-vestibular	1132
Have you taken a SimENEM exam before?	No	6331
	Yes	4349
Self-assessment math	Best of my class	647
	Good student	3813
	Not good or bad student	6220
Self-assessment Portuguese	Best of my class	485
	Good student	5277
	Not good or bad student	4918
Number of people living at home	1	58
	2	1028
	3	3158
	4	4534
	5	1441
	6	461

Pearson

Table B7. COC vs. non-COC public 2018 SimENEM total score statistical significance tests

Model effect	Type III sum of squares	df	Mean square	F	Sig
Corrected model	258356392.259[a]	16	16147274.516	318.652	0.000
Intercept	2615200812.741	1	2615200812.741	51608.638	0.000
COC vs. non-COC public	49577802.413	1	49577802.413	978.373	0.000
Grade	124664790.888	4	31166197.722	615.037	0.000
Have you taken a SimENEM exam before?	736472.012	1	736472.012	14.534	0.000
Self-assessment math	64251048.167	2	32125524.084	633.968	0.000
Self-assessment Portuguese	11149539.080	2	5574769.540	110.013	0.000
Number of people living at home	1088004.906	5	217600.981	4.294	0.001
Number of years attending same school (covariate)	2197.000	1	2197.000	0.043	0.835
Error	540333695.831	10663	50673.703		
Total	33343348592.098	10680			
Corrected total	798690088.090	10679			

a. R squared = .323 (adjusted R squared = .322)

Pearson

Table B8. COC vs. non-COC public 2018 SimENEM statistical model parameters

Model effect	B	Robust standard errorr[a]	t	Sig	95% Confidence interval	
					Lower bound	Upper bound
Intercept	1922.460	13.733	139.989	0.000	1895.541	1949.379
Non-COC public	-182.930	5.802	-31.528	0.000	-194.303	-171.557
COC	0[b]					
Grade = 9	-500.652	11.451	-43.722	0.000	-523.098	-478.207
1st year high school	-386.999	10.271	-37.678	0.000	-407.132	-366.866
2nd year high school	-279.377	9.580	-29.163	0.000	-298.155	-260.599
3rd year high school	-191.303	8.907	-21.477	0.000	-208.763	-173.843
Grade = pre-vestibular	0[b]					
Have you taken a SimENEM exam before = no	-19.695	5.253	-3.749	0.000	-29.992	-9.399
Have you taken a SimENEM exam before = yes	0[b]					
Self-assessment math = best of my class	239.991	10.166	23.606	0.000	220.063	259.919
Self-assessment math = good student	142.189	4.737	30.015	0.000	132.903	151.475
Self-assessment math = not good or bad student	0[b]					
Self-assessment Portuguese = best of my class	119.429	11.027	10.830	0.000	97.814	141.045
Self-assessment Portuguese = good student	56.034	4.479	12.510	0.000	47.254	64.814

Self-assessment Portuguese = not good or bad student	0[b]					
Number of people living at home = 1	43.787	34.477	1.270	0.204	-23.795	111.369
Number of people living at home = 2	19.069	13.219	1.443	0.149	-6.843	44.982
Number of people living at home = 3	33.218	11.646	2.852	0.004	10.389	56.046
Number of people living at home = 4	37.328	11.422	3.260	0.001	14.849	59.627
Number of people living at home = 5	16.046	12.383	1.296	0.195	-8.227	40.319
Number of people living at home = 6	0[b]					
Number of years attending same school (covariate)	0.135	0.663	0.204	0.838	-1.165	1.436

a. Heteroskedastic-consistent HC3 method
b. This parameter is set to zero because it is redundant

Pearson

Appendix C. Primary analyses: matched group balance tables

Table C1. COC primary analyses: SimENEM summary statistics before matching

	Variable	Full sample	COC private	Non-COC public	Non-COC private
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Dependent	2018 Language	562.85 (87.22)	574.71 (87.35)	504.49 (76.89)	567.58 (77.24)
	2018 Math	436.67 (79.32)	445.69 (78.96)	392.91 (71.08)	439.75 (76.84)
	2018 Human sciences	526.23 (98.06)	530.95 (102.60)	489.55 (78.35)	538.46 (90.86)
	2018 Natural sciences	450.49 (82.09)	458.72 (84.84)	412.27 (60.22)	451.97 (80.25)
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Matching variables	2017 Language	517.19 (93.09)	521.10 (92.21)	471.59 (97.36)	534.86 (85.13)
	2017 Math	461.28 (87.51)	469.09 (85.97)	413.11 (80.73)	471.87 (87.49)
	2017 Human sciences	488.83 (88.91)	494.19 (89.30)	450.42 (80.55)	496.44 (88.21)
	2017 Natural sciences	515.15 (89.98)	521.35 (86.98)	482.93 (82.55)	518.93 (101.62)
	Years at same school	2.57 (3.74)	2.67 (3.82)	2.57 (3.58)	2.25 (3.64)
	Household size	3.63 (1.09)	3.58 (1.14)	3.62 (0.83)	3.79 (1.07)
	3rd year high school	25%	24%	35%	23%
	Pre-vestibular	75%	76%	65%	77%
	Observations	248	163	37	48

Pearson

Table C2. Matching variables balance summary statistics for COC vs. public students

Group	2018 SimENEM	2017 SimENEM	Years at same school	Household size	Grade	Observations
		Mean (SD)	Mean (SD)	Mean (SD)	Percent	Count
COC private	Language	502.28 (83.59)	2.00 (3.30)	3.52 (1.02)	27%	113
	Math	448.27 (72.40)	1.91 (3.45)	3.42 (1.05)	27%	106
	Human science	499.31 (85.36)	1.52 (3.14)	3.43 (1.04)	18%	110
	Natural science	515.53 (84.68)	1.34 (2.67)	3.48 (1.04)	23%	109
Non-COC public	Language	506.24 (86.98)	2.05 (3.50)	3.97 (0.91)	27%	33
	Math	446.95 (68.83)	2.17 (3.91)	3.48 (0.79)	27%	33
	Human science	499.81 (89.03)	1.73 (3.64)	3.58 (0.76)	18%	35
	Natural science	515.27 (78.95)	1.31 (2.66)	3.61 (0.93)	23%	32
COC private vs. Non-COC public		Mean diff. (Effect size)	Mean diff. (Effect size)	Mean diff. (Effect size)	Mean diff. (Effect size)	Unmatched examinees
	Language	-3.96 (-0.04)	-0.05 (-0.01)	-0.45 (-0.41)	0	54
	Math	1.32 (0.02)	-0.26 (-0.07)	-0.06 (-0.06)	0	61
	Human science	-0.50 (-0.01)	-0.21 (-0.06)	-0.15 (-0.14)	0	55
	Natural science	0.26 (0.00)	0.03 (0.01)	-0.13 (-0.12)	0	59

Effect size is calculated using the full sample standard deviations in Table C1

Pearson

Table C3. Matching variables balance summary statistics for COC vs. non-COC private students

Group	2018 SimENEM	2017 SimENEM	Years at same school	Household size	Grade	Observations
		Mean (SD)	Mean (SD)	Mean (SD)	Percent	Count
COC private	Language	532.87 (85.93)	2.47 (3.70)	3.47 (1.05)	19%	129
	Math	481.10 (65.85)	2.19 (3.61)	3.48 (1.08)	13%	117
	Human science	513.66 (75.09)	2.19 (3.54)	3.48 (1.10)	14%	124
	Natural science	527.47 (79.85)	2.14 (3.55)	3.55 (1.06)	17%	132
Non-COC private	Language	540.87 (87.58)	2.41 (3.69)	3.59 (0.97)	19%	44
	Math	473.59 (68.81)	2.26 (3.83)	3.76 (0.98)	13%	46
	Human science	510.23 (74.50)	2.29 (3.73)	3.64 (0.99)	14%	42
	Natural science	525.52 (81.02)	2.23 (3.81)	3.70 (0.92)	17%	43
		Mean diff. (Effect size)	Mean diff. (Effect size)	Mean diff. (Effect size)	Mean diff. (Effect size)	Unmatched examinees
COC Private vs. Non-COC Private	Language	-8.00 (-0.09)	0.06 (0.02)	-0.12 (-0.11)	0	38
	Math	7.51 (0.09)	-0.07 (-0.02)	-0.28 (-0.26)	0	48
	Human science	3.43 (0.04)	-0.10 (-0.03)	-0.16 (-0.15)	0	45
	Natural science	1.95 (0.02)	-0.09 (-0.02)	-0.15 (-0.14)	0	36

Years = number of years in current school, and household = number of people living in the home

Effect size is calculated using the full sample standard deviations in Table C1

Pearson

Appendix D. Primary matched analyses robustness checks

Table D1. Results for COC matched comparisons on 2018 SimENEM using CEM

Comparison group	Test	Effect size[2]	Robust standard error	t	P-value
Public	Language	0.863*	0.150	5.76	0.000
	Math	0.326	0.253	1.29	0.198
	Human science	0.431+	0.238	1.81	0.072
	Natural science	0.347*	0.174	2.00	0.047
Private	Language	0.043	0.149	0.29	0.774
	Math	0.233	0.206	1.13	0.260
	Human science	-0.035	0.170	-0.20	0.839
	Natural science	0.145	0.162	0.89	0.375

The two matching variables include 2017 SimENEM and grade level.

Statistical significance tests in this table are not controlled for a family-wise false discovery rate
$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Pearson

Table D2. Results for COC matched comparisons on 2018 SimENEM Using CEM and doubly-robust methods

Comparison group	Test	Effect size[2]	Robust standard error	t	P-value
Public	Language	0.897***	0.240	3.74	0.000
	Math	0.318	0.228	1.40	0.165
	Human science	0.424+	0.226	1.87	0.063
	Natural science	0.375*	0.159	2.36	0.019
Private	Language	0.082	0.111	0.73	0.463
	Math	0.231	0.145	1.59	0.113
	Human science	-0.056	0.092	-0.61	0.542
	Natural science	0.202	0.133	1.52	0.131

The two matching variables include 2017 SimENEM and grade level. The statistical models include all matching variables as covariates to further adjust the effect sizes for remaining differences in the matched groups. Statistical significance tests in this table are not controlled for a family-wise false discovery rate

+p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Pearson

Table D3. Results for COC matched comparisons on 2018 SimENEM
Using inverse-probability-weighted regression adjustment estimator (IPWRA)

Comparison group	Test	Effect size	Robust standard error	t	P-value
Public	Language	0.717***	0.152	4.71	0.000
	Math	0.468+	0.268	1.74	0.081
	Human science	0.284+	0.164	1.73	0.084
	Natural science	0.456**	0.141	3.23	0.001
Private	Language	0.163	0.113	1.44	0.151
	Math	0.126	0.105	1.20	0.231
	Human science	-0.060	0.102	-0.59	0.558
	Natural science	0.123	0.125	0.99	0.323

The matching variables include 2017 SimENEM, grade level, and years at the same school. The statistical models include all matching variables as covariates to further adjust the effect sizes for remaining differences in the matched groups. Statistical significance tests in this table are not controlled for a family-wise false discovery rate

+p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Pearson

Table D4. Covariate balance summary statistics COC private (treatment) vs. non-COC public (control)

	Standardized differences					
	2017 SimENEM		Grade		Years at same school	
	Raw	Weighted	Raw	Weighted	Raw	Weighted
Language	.5221354	-.0132434	-.366931	-.0947088	-.0657062	.083004
Math	.6713488	.0908227	-.2644569	-.1233959	.0337446	.1160324
Human science	.5147825	.0745458	-.366931	-.083863	-.0657062	.0820308
Natural science	.4530016	.0745569	-.2644569	-.1009876	.0337446	.0507106
	Variance ratio					
Language	.8969871	.7371358	.7287148	.9159825	1.065244	1.17309
Math	1.134181	1.193359	.7319872	.8750538	1.072642	1.203016
Human science	1.229183	1.057478	.7287148	.9236615	1.065244	1.18195
Natural science	1.110442	1.173961	.7319872	.8915081	1.072642	.975139

Pearson

Table D5. Covariate balance summary statistics COC private (treatment) vs. non-COC private (control)

	Standardized differences					
	2017 SimENEM		Grade		Years at same school	
	Raw	Weighted	Raw	Weighted	Raw	Weighted
Language	-.1551229	-.0030919	.0046427	.025926	.1044015	.0211459
Math	-.0320345	.0050037	.0052099	-.0019722	.0968343	-.0083904
Human science	-.0252686	-.0147929	.0161235	-.0027654	.1175332	-.0007493
Natural science	.0255884	-.0128354	.0052099	-.0044216	.0968343	-.0102711
	Variance ratio					
Language	1.173192	1.102587	.9906115	1.034115	1.094696	.9704252
Math	.9655385	.9729012	.9913835	.9971378	1.048295	.8870438
Human science	1.024842	1.029427	1.005744	.9965319	1.109211	.9223966
Natural science	.7327621	.7561429	.9913835	.9936079	1.048295	.8837551

Pearson

Table D6. Covariate balance summary statistics observations COC private (treatment) vs. non-COC private (control)

	COC private (treatment) vs. non-COC public (control)							
	Language		Math		Human science		Natural science	
	Raw	Weighted	Raw	Weighted	Raw	Weighted	Raw	Weighted
Total	193	193.0	184	184.0	193	193.0	184	184.0
COC private	161	97.8	150	94.4	161	98.5	150	93.5
Non-COC public	32	95.2	34	89.6	32	94.5	34	90.5
	COC private (treatment) vs. non-COC private (control)							
	Raw	Weighted	Raw	Weighted	Raw	Weighted	Raw	Weighted
Total	208	208.0	194	194.0	209	209.0	194	194.0
COC private	161	104.0	150	97.0	161	104.5	150	97.0
Non-COC private	47	104.0	44	97.0	48	104.5	44	97.0

Pearson

Appendix E. Additional tables

Table E1. Model parameters for the results presented in Tables 3 and 4 – simple and doubly-robust comparisons for COC private vs. non-COC public

	Language		Math		Human science		Natural science	
	Simple	Doubly robust	Simple	Doubly robust	Simple	Doubly robust	Simple	Doubly robust
COC vs. public	0.552** (0.192)	0.558* (0.228)	0.256 (0.266)	0.279 (0.181)	0.198 (0.232)	0.177 (0.193)	0.406* (0.167)	0.444** (0.164)
Covariates								
2017 SimENEM		0.563*** (0.095)		0.715*** (0.092)		0.581*** (0.074)		0.649*** (0.092)
Grade (2018)		-0.311 (0.262)		0.020 (0.264)		-0.457 (0.309)		0.158 (0.254)
Years at same school		0.003 (0.024)		-0.053* (0.026)		0.002 (0.036)		-0.035 (0.025)
Household size		-0.073 (0.076)		-0.088+ (0.053)		-0.137* (0.059)		0.007 (0.059)
Constant	-0.596*** (0.168)	-0.170 (0.257)	-0.343 (0.249)	0.233 (0.270)	-0.187 (0.211)	0.316 (0.283)	-0.374** (0.131)	-0.317 (0.267)
R^2	0.059	0.447	0.012	0.558	0.008	0.472	0.029	0.421
Observations	146	141	139	129	145	140	141	131

Robust standard errors in parentheses
Statistical significance tests in this table are not controlled for a family-wise false discovery rate
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Pearson

Table E2. Model parameters for the results presented in Tables 3 and 4 – simple and doubly-robust comparisons for COC private vs. non-COC private

	Language		Math		Human science		Natural science	
	Simple	Doubly robust	Simple	Doubly robust	Simple	Doubly robust	Simple	Doubly robust
COC vs. private	-0.040 (0.147)	0.009 (0.118)	0.142 (0.134)	0.086 (0.087)	-0.170 (0.155)	-0.210* (0.087)	0.085 (0.159)	0.112 (0.119)
Covariates								
2017 SimENEM		0.600*** (0.072)		0.813*** (0.062)		0.713*** (0.062)		0.647*** (0.072)
Grade (2018)		-0.394* (0.199)		-0.417+ (0.220)		-0.607** (0.193)		-0.156 (0.175)
Years at same school		0.012 (0.014)		-0.009 (0.013)		0.019 (0.013)		-0.004 (0.013)
Household size		-0.059 (0.048)		-0.045 (0.039)		-0.085* (0.040)		0.033 (0.051)
Constant	0.262* (0.120)	0.369+ (0.209)	0.126 (0.104)	0.256 (0.159)	0.355** (0.130)	0.533** (0.170)	0.103 (0.133)	-0.047 (0.212)
R^2	0.000	0.522	0.005	0.658	0.007	0.636	0.002	0.427
Observations	173	171	163	152	166	166	175	164

Robust standard errors in parentheses
Statistical significance tests in this table are not controlled for a family-wise false discovery rate
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Pearson

Table E3. Comparison of primary analytic sample to larger samples on 2018 SimENEM registration information

Variable and levels	COC			Non-COC private			Non-COC public		
	33K registered 2018	13K 2018 tested	18-year-olds tested 2017 and 2018	33K registered 2018	13K 2018 tested	18-year-olds tested 2017 and 2018	33K registered 2018	13K 2018 tested	18-year-olds tested 2017 and 2018
Self-assessment math (percent)									
Best of my class	4.83	5.35	1.84	7.34	8.87	6.25	6.06	7.95	8.11
Good student	32.08	34.31	31.90	38.86	41.86	41.67	35.43	39.45	35.14
Not good or bad student	63.09	60.34	66.26	53.80	49.27	52.08	58.51	52.59	56.76
Self-assessment Portuguese (percent)									
Best of my class	3.80	4.10	1.84	6.70	6.79	8.33	5.43	5.74	0.00
Good student	48.76	48.48	46.01	52.81	50.93	56.25	51.78	51.90	43.24
Not good or bad student	47.44	47.42	52.15	40.49	42.27	35.42	42.79	42.36	56.76
Number of people living at home (percent)									
2	10.88	9.82	14.11	11.24	9.23	8.33	11.08	9.09	8.11
3	30.55	30.30	25.15	30.41	31.12	37.50	26.39	27.59	35.14
4	41.73	43.50	39.26	39.67	41.39	27.08	36.06	39.63	43.24
5	12.47	12.37	12.88	13.87	13.95	20.83	17.88	16.53	13.51
6+	3.68	3.42	4.91	4.27	3.79	6.25	7.95	6.74	0.00
Live alone	0.69	0.59	3.68	0.54	0.52	0.00	0.63	0.41	0.00

When did you start school? (percent)									
Nursery (0 to 3 years old)	42.15	42.17	39.26	45.14	44.92	31.25	33.36	32.54	29.73
Kindergarten (4 to 5 years old)	38.74	39.28	45.40	47.58	47.77	62.50	53.34	55.33	56.76
1st grade Elementary 1 (6 to 7 years old)	6.53	6.12	6.75	5.27	4.82	2.08	12.17	11.24	13.51
After 1st grade	12.58	12.43	8.59	2.02	2.49	4.17	1.13	0.90	0.00
Have you failed to graduate? (percent)									
No	93.42	94.77	82.82	92.65	94.24	89.58	89.09	94.05	91.89
Yes, once	5.45	4.30	14.11	6.06	5.29	10.42	8.42	4.94	8.11
Yes, two times or more	1.13	0.92	3.07	1.30	0.47	0.00	2.49	1.00	0.00
Have you left school during the school year and not returned that year? (percent)									
No	99.11	99.41	98.77	98.33	99.01	97.92	96.21	98.58	100.00
Yes, once	0.78	0.53	1.23	1.46	0.73	2.08	2.85	1.14	0.00
Yes, two times or more	0.11	0.06	0.00	0.20	0.26	0.00	0.94	0.28	0.00
What will you do after finishing high school? (percent)									
I don't know yet	8.45	8.51	3.68	5.20	4.93	2.08	3.74	4.01	5.41
Keep studying and working	50.74	49.31	27.61	57.79	55.13	39.58	75.08	70.99	51.35
Only keep studying	40.29	41.81	68.10	36.87	39.83	58.33	20.75	24.90	43.24
Only work	0.52	0.37	0.61	0.13	0.10	0.00	0.43	0.10	0.00
Total frequency									
Total	14,148	7,788	163	5,942	1,928	48	10,514	2,892	37

The analytic sample consists of those examinees that were 18 at the time they registered for the 2018 SimENEM, and had scores for both the 2017 and 2018 SimENEM.

The yellow cells are instances where an exact 95% confidence interval cannot be calculated because the frequency (0 or 1) for the sample of the 18-year-olds does not permit it.

The red cells are instances where the 33k or 13k samples are outside the 95% confidence interval around the 18-year-old sample percentage.