Herrera Prada, Luis Omar

Working Paper

Ending the musical chairs game in higher education: How a software dashboard (SPADIES) improved information flow and educational outcomes in Colombia

IAI Discussion Papers, No. 251

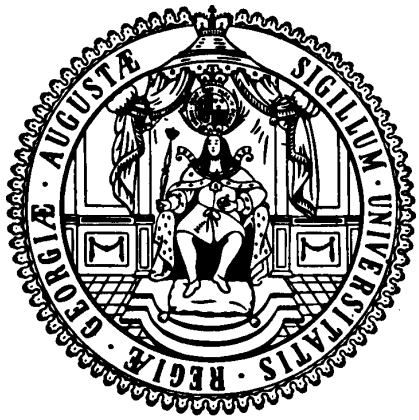
Provided in Cooperation with:
Ibero-America Institute for Economic Research, University of Goettingen

Suggested Citation: Herrera Prada, Luis Omar (2021) : Ending the musical chairs game in higher education: How a software dashboard (SPADIES) improved information flow and educational outcomes in Colombia, IAI Discussion Papers, No. 251, Georg-August-Universität Göttingen, Ibero-America Institute for Economic Research (IAI), Göttingen

This Version is available at:
https://hdl.handle.net/10419/233065

Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)



Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

Nr. 251

Ending the Musical Chairs Game in Higher Education: How a Software Dashboard (SPADIES) Improved Information Flow and Educational Outcomes in Colombia

Luis Omar Herrera-Prada

March 2021

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173
e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

Ending the Musical Chairs Game in Higher Education: *

How a Software Dashboard (SPADIES) Improved Information Flow and Educational Outcomes in Colombia

Author:

Luis Omar Herrera-Prada†

Ibero-America Institute for Economic Research

Universität Göttingen

International Monetary Fund

March 20, 2021

Keywords: Dropout, Higher Education, Education, Graduation, On-time graduation

JEL Classification Numbers: H40, H52, I23, I28, O33

*I am thankful to Inmaculada Martinez-Zarzoso, Julián Cristia, Alexandra Sastoque, Nadmy Daniela Valbuena González, Sarah Belford, Juan Esteban Saavedra, Darío Maldonado, Andrés Córdoba, Elkin Garavito, Diana Durán, Paula Andrea Triana Romero, Román David Zárate, Román Andrés Zárate, Mónica Ospina, Fabio Sánchez, Gustavo Canavire, SPADIES team 2005-2011 at Universidad de los Andes, and Ministry of Education. Comments received in the Development Economics Seminar at Universität Göttingen and Public Policy Seminar at Universidad de los Andes - Colombia were invaluable. An online appendix with descriptive analysis on covariables and outputs with a combination in the disaggregation of characteristics is available in `https://herreraprada.com` . All views and errors are mine.

†Corresponding Author: l.herreraprada@stud.uni-goettingen.de

Abstract

The new millennium brought with it a new challenge. Students reaching higher education were not only greater in number, they were also less prepared than their predecessors. This caused a sharp deterioration in the quality of the entire higher education system, which in turn affected enrollment and graduation rates. Colombia faced this challenge with severely scarce economic resources, but the Ministry of Education designed a plan to keep students in higher education and eventually increased the overall enrollment rate from 20% in 2002 to over 40% in 2010. One of the tools used to implement this plan was the System for the Prevention and Analysis of School Dropouts in Higher Education (SPADIES). This software dashboard facilitates the collection, analysis, and visualization of student data. This paper provides evidence that SPADIES was vital to achieving higher educational attainment outcomes across the country. Using a differences-in-differences approach, I find that SPADIES reduced the probability of dropping out by 1.2% (impact of 4.35%), increased the probability of graduating by 0.6% (impact of 1.3%), and increased the probably of graduating on-time by 0.6% (impact of 1.9%). I also find that the number of transitions (i.e. student status changes from absent or dropout to enrolled) decreased 3.4% (impact of 48.4%), and the average duration of a student's gap in enrollment decreased 0.2 semesters (impact of 14.4%).

1 Introduction

The dropout phenomenon is one of the most costly and difficult problems in higher education. The costs associated with not obtaining the degree - wasted resources (structural and financial)- the loss of efficiency in connecting graduates with the labor market, and the opportunity cost to dropouts - throws a stark spotlight on the need to reduce the dropout rate. Between 2000 and 2010, Colombia faced the "Crowding cohort"[1] phenomenon when the children of the "boomers" started to enter higher education, and this massive influx of less prepared students lead to a deterioration in the quality of higher education institutions in Colombia (Herrera-Prada, 2013; Ferreyra et al., 2017) .

The population that reached higher education between 2000 and 2010 was bigger than before because of demographic growth and the success of a policy to increase the secondary enrollment and graduation rate. However, the structure of the higher education system was not prepared and was essentially the same as it had been two decades ago Orozco Silva et al. (2006); Orozco Silva (2010). These new students had high levels of both academic and economic vulnerabilities. On one hand, these students faced academic vulnerabilities because the success of the policy that increased the enrollment rate in mid-90s was based on promoting all the students to the next grade level without any academic restrictions. On the other hand, the economic vulnerabilities stem from the fact that the children of the baby boomers were facing the opposite financial terms than their parents. In the 1970s, when the baby boomers were enrolling higher education, Colombia had an economic boom from the coffee sector. But in the early 2000s, when their children started to arrive to higher education, Colombia was suffering from the worst economic depression it had ever experienced (pre-COVID) and from a failed peace process dealing with the oldest guerrilla group on the continent, which created profound social-political turmoil (Lucio and Serrano, 1992; Bound and Turner, 2006; Herrera-Prada, 2013) and directed the use of scarce resources to military action.

Between 2002 and 2006, the Ministry of Education (MEN) designed a plan to improve educational outcomes in higher education and, specifically, to increase the enrollment rate by providing these cohorts with additional financial and academic support. The primary strategy to respond to the increase in demand was to use available resources from the public sector (mainly infrastructure), while at the same time creating a highly coordinated information system to track students

[1] "Crowding cohort" is a phenomenon when there is increased demand for fixed or reducing resources. This term was used for the first time in Bound and Turner (2006).

and improve the overall efficiency of the system(Orozco Silva, 2010). This is how the System for the Prevention and Analysis of School Dropout in Higher Education (SPADIES) was born. SPADIES was novel in its approach; it provided "real time" information about the enrollment, academic performance, quality of peers, dropout rates, and graduation rates to authorities in higher education institutions (HEIs), MEN, and the general public. MEN promoted the use of SPADIES to help each HEI predict which students were at risk of dropping out, get their students' complete academic profile, and target those students with different types of assistance to reduce the dropout rate. HEIs used three types of assistance: (1) academic assistance (e.g., tutoring or remedial classes), (2) financial assistance (i.e.. scholarships), and (3) all other assistance, which ranged from food vouchers and transportation to mental health services(Ministerio de Educación Nacional, 2008, 2009, 2010).

This paper uses SPADIES data from 1998-2017 and tracks the college life paths of 4.4 million students to evaluate the impact of SPADIES on educational and system efficiency outcomes. Using a differences-in-differences approach where the treated are those students enrolled in an HEI with SPADIES installed, I find that SPADIES not only reduced the probability of dropping out by 1.2% (impact of 4.35%), but also increased both the probability of getting the degree by 0.6% (impact of 1.3%) and of getting it on-time by 0.6% (impact of 1.9%). I also find that the number of transitions (student status changes from absent or dropout to enrolled) decreased 3.4% (impact of 48.4%), and the average duration of a gap in enrollment decreased 0.2 semesters (impact of 14.4%). While I do not find any effect of academic assistance on educational outcomes, I do find that public financial aid is particularly useful in reducing the dropout rate, and private financial aid and the "other" assistance category help reduce the probability of having a transition (stopping studying for some period and then returning to continue). I find SPADIES was most effective in reducing the dropout rate for low-income students. SPADIES increased the graduation rate, especially in certified HEIs, and the on-time graduation rate in Bachelor programs. Finally, results suggest that certified HEIs with SPADIES were better able to bring students back from absent status; and, the non-certified HEIs with SPADIES were better able to bring back students from the dropout status.

The following section presents a review of the literature. Section 3 details the SPADIES program design and the context in which it was created. Section 4 describes the dataset and variables. Section 5 presents the model specification. Section 6 discusses the results. Finally, Section 7 presents the conclusions and policy implications.

2 Literature Review

In this section, I review the literature that provides a foundation for the chapter. First, I will discuss the literature of the demand and supply sides for higher education, their implications on quality, and the dropout analysis evolution. Next, I will present some specific papers related to the Colombian context.

In 1999, the World Bank created the program "Education for All," in which one of the main relevant objectives was monitoring more closely the educational indicators access, enrollment, and quality. The idea was to implement some of the lessons learned from improvements in developed countries' Higher Education Systems in developing regions like Latin America and encourage the governments to track the results and indicators.

The evolutionary globalization conditions have identified that formal education and, more specifically, higher education plays a transversal role in the development and growth of human capital. This function consists of providing individuals with the capacities and skills to enter into productive sectors and boost social mobility (Becker, 1962). Given the importance of human capital formation, complex dynamics are presented from the demand and supply conditions that end up in an unbalanced market due to barriers of access, low enrollment rates, excess of demand and/or supply, and low quality (Epple et al., 2006). The discrepancy between supply and demand can be attributed to several reasons, including the lack of public resources due to an unexpected increase in applicants (Bound et al., 2009). Then, the growth of the demand raises problems such as resource allocation and efficiency, attrition, late graduation or a lag in studies, and a decrease in the quality of the education (Bound and Turner, 2006). Subsequently, institutional factors that may influence the individual's decision to enroll or continue or not continue in a HEI, such as the quality of the peers or the education received, and supports or services offered by the institution, are incorporated as a part of cost-benefit analysis (Tinto, 1975, 1982; Bank et al., 1990; MacLeod and Urquiola, 2015). The theoretical basis of the discussion around attrition is commonly classified into two strands: the first one is the student's integration or adaptation to the education system model (Tinto, 1975, 1982), and the second one explains attrition as a set of conditions linked to the individuals' socioeconomic factors, such as family conditions or academic performance during school (Bean, 1980, 1985).

Economists have tried to integrate these two models with cost-benefit analysis and with the

interaction of variables containing social, economic, individual, family, and institutional factors (Cabrera et al., 1993). The analyses have not disregarded students' academic performance, or the level of education of their parents (Porto and Di Gresia, 2004); the individual characteristics of the student such as their gender or marital status, or their poor academic performance at the beginning of their academic life (Stratton et al., 2008); or their experiences after enrolling in higher education (St. John et al., 2016). The wide variety of studies also take into account how students' lives changed when they were accepted into the system, how their new social role changed, and how their interaction with their peers -and especially with their professors- lowered a student's likelihood of dropping out of school (Bank et al., 1990; Cabrera et al., 1996; Portes and Salas, 2007; Herrera-Prada, 2020). However, the enrollment, the dropout and graduation rates, and the quality of education depend on how students were sorted among HEIs; two main models developed by Epple et al., 2006 and MacLeod and Urquiola, 2015 illuminate how this process happened. But for this development, the information flow among agents was instrumental. Information provision, data access, and data mining have been at the center of the agenda to improve both education systems and their educational outcomes, especially in primary and secondary schools- over the past 25 years Campbell and Levin (2009); Asif et al. (2017); Allende et al. (2019); Kerr et al. (2020); Unesco (2020).

In Colombia, the first studies on attrition were developed in a very few HEI. For example, Universidad Nacional de Colombia (2007) study the lag, graduation rate, and dropout rate in the most important public university. The authors demonstrate the relevance of including contextual variables in the analysis. In general, they find that being a woman (especially 18 years old or younger) constitutes a significant characteristic that positively impacts the probability of obtaining a degree from any program. Financial aid or student loan programs decrease the dropout rate. In this sense, it is important to stress that studies agree on the importance of including in the public policy "Affirmative Action" programs, such as special admission mechanisms and alternative admission through pre-university courses. These programs aim at improving students' chances of staying in higher education and, as a consequence, improving access conditions and social equity (Sánchez et al., 2002). At the University of Antioquia, (Castaño et al., 2006) conducted a study on the students of the School of Engineering and the School of Economics in the second cohort of 1996. Using a duration model, including context variables, they found that being male, single,

and over 18 years old entails a greater risk of dropping out. They also found that living with their parents, having a better academic performance, not working, having parents with a high level of education, and being female decrease the risk of dropping out. The studies implemented in public universities have initiated and addressed the State's concern about the way resources are being used in the educational system; since they understand the dropout phenomenon as a waste of economic resources, human capital (professors and staff), and infrastructure (Cárdenas, 1996; Córtes et al., 2011; Facundo-Díaz, 2009). Finally, at a national level, ICFES (2002)found that household financial conditions were the primary determinant for becoming a dropout student. However, Ministerio de Educación Nacional (2008) showed that the dropout rate's main reasons were academic skills (measured in the secondary exit test score -ICFES test-), the mismatch in the career choice and skills, academic performance, and gender. Finally, Herrera-Prada (2013) showed a decrease in the graduation rate even though the dropout rate decreased in the whole country, explained by both the "Crowding cohort" and the simultaneous increase in the time needed to graduate.

3 Context and Program Design

In this section, I present some facts and statistics about the Colombian context and the history behind the creation of SPADIES.

3.1 Context

In the last 50 years, Colombia has experienced dramatic demographic changes, including massive migration from rural to urban areas that brought both cheap labor and young people hoping for opportunities of social mobility to the city. The new generations were eager to progress in a country that was not prepared to provide them with adequate educational resources(Lucio and Serrano, 1992). In the mid-1990s, the lack of capacity in high schools was evident, but the government preferred enrollment to quality and allowed students to advance through primary and secondary school without any academic restrictions. The promotion of students without restriction was successful, as coverage and completion rates in secondary schools increased, but the skills and quality of education significantly decreased (Herrera-Prada, 2013). The economic crisis of the late 1990s delayed the inevitable: few Colombians made it to higher education, and those who did were likely to drop

out early in their studies. Colombia started to analyze the phenomenon in the early 2000s and found that the main problem was the economic situation in students' households (ICFES, 2002). At that time, the nation's gross college enrollment rate was 20 percent, one of the lowest in the region(Ministerio de Educación Nacional, 2010).

In the middle of a very tough scenario, MEN created a plan to respond to the increasingly low qualified demand and the scarcity of resources. The plan was named "Educational Revolution". Amazing results came quickly as the enrollment rate jumped from 20% in 2002 to over 40% by the end of 2010 and then again to over 45% by 2016 (Orozco Silva et al., 2006; Ministerio de Educación Nacional, 2010, 2017). The process was mainly backed by HEIs in the public sector since the enrollment rate of private HEIs was basically zero or negative in the first half of the 2000s. For the first time since 1974, the public sector got more students than the private sector(Orozco Silva et al., 2006; Orozco Silva, 2010). However, the phenomenon was even more exceptional if the growth in the enrollment rate is analyzed. The pace of this growth was lower than the annual average of the last 70 years (8.74%) (Orozco Silva et al., 2006). Then, the question that arises is: how did this happen? The answer has two components: a steady growth rate of just over 6% per year during the first decade and a simultaneous reduction in the number of students leaving their schools. Basically, what seems to have happened in previous periods was a higher rate of enrollment but also a higher rate of dropouts. So, the solution was to reduce the dropout rate. To do so, MEN created SPADIES, thereby providing "real time" information about the enrollment, academic performance, quality of peers, dropout rates and graduation rates, and a list of students at risk of dropping out with their complete profile to be targeted in the correct aid they need, to authorities in the higher education institutions (HEIs), MEN, and some of the main characteristics of the system and HEIs performance indicators to the public.

3.2 Program Design

SPADIES is an information system conceived by MEN at the end of 2004. MEN commissioned the Universidad de los Andes to create a software tool that would collect information on all students enrolled in higher education and that would allow visualizations, statistical analysis, and the creation of lists of students at risk of dropping out. One module of the system was public and allowed students, authorities, and the general public to make visualizations and analyses comparing the

performance of HEIs or studying departmental or national aggregates.

When the first data were obtained, MEN's concern was focused on understanding the factors that determine dropout. Consequently, MEN and Universidad de Los Andes used the novel database to estimate Duration Models and also conducted a focal group analysis across the country. The results of these two analyses were later implemented in the software.

Once the database had enough information per HEI for the estimation of the models and the temporal analyses of the cohorts allowed finding consistent behavior patterns among the students who became dropouts, MEN and the Universidad de Los Andes visited each HEI to install the software, explain the results, and train the staff on the software. The visits lasted between 4 and 6 hours. The participation of the President in the visit was optional, but the participation of the academic vice-president, the head of the planning department, and the chief of the systems office was mandatory. In the first part of the visit, the results of the national and institutional estimates were presented. The second part consisted of the installation of the software in at least one of the offices of the directors required during the visit. After the installation, the final part of the meeting consisted of teaching HEI personnel how to use the tool to identify the population at risk of dropping out. Because SPADIES has information at the individual level, each HEI could visually analyze its population and generate a name list of students in different risk levels. The objective was to use this list to target the different assistance types to avoid dropout in students who were still enrolled. The assistance programs were sorted into three groups:

1. Financial aid (grants, scholarships, or any financial support that was not a financial instrument like loans or credits).

2. Academic aid (free tutorials or remedial courses).

3. "Other type of aid" (any other free aid given to the student different from those reported as financial or academic aid, such as mental health support, coaching, career selection, etc.)

At the national level and according to Ministerio de Educación Nacional (2008), the most vulnerable population used to be males with low household income and low academic skills, mainly in Associate programs or math-related programs.

Finally, between 12 and 18 months after installation, MEN and the Universidad de Los Andes made one follow-up visit ("Accompaniment") to check the new outcomes and the correct use of the

application. In 2011, all institutions had the software installed, and all of them had their follow-ups by the end of 2013. The Universidad de Los Andes was the consultant selected by MEN for the terms 2004-2011 and 2014-2017. Universidad EAFIT was in charge of the system between 2011 and 2013. After 2011, consultants only made follow-up visits, provided support to MEN and HEIs, improved the software, and developed research using the data for MEN. The consultants were required to have a team of people dedicated to answering doubts and analyzing the information from each institution. All the institutions had to have two delegates with permanent contact with MEN or the consultant on a monthly basis.

Starting in 2008, MEN organized several contests on strategies to reduce dropout using SPADIES as a mechanism to measure the dropout rate and to assign the prizes. Another important event to remark is that the collection of information at the individual level since 1998 was a historical milestone for many institutions–many of which were introduced to the digital world by SPADIES. By the end of 2005, few institutions had electronic records; the Ministry's information systems triggered an arduous migration process from paper to digital records. By 2017, about 60% of institutions were able to report their full data starting in 1998. SPADIES' installation and training were not randomly assigned because of the stark difference in the digital gap and quality of information of each HEI. Five rounds (2005-06, 2006-07, 2007, 2008-09, and 2010-11) were necessary to complete the collection of each of the HEIs' information and add it into the system (The list of HEIs in each round are shown in Figure 1). Each round corresponded to a new group of HEIs having their digital records in shape in order to trigger the process mentioned above. However, there is no difference in the assignation by the sex of the student, or the sector, quality, or region of the HEIs (Figure 2). The fact that some HEIs are biased toward high score and high income students pursuing bachelors degrees is evident; it is due these students' that HEIs were more biased to be enrolled in the first round because of their data quality. A time description of the composition of the HEIs (parallel trends in covariates) by quality, location, sector, and level can be found in the online appendix.

Also in 2008, MEN required that HEIs include the official dropout rate measured by SPADIES in the requirements for the quality certificate of programs. At the same time, a web portal was created for MEN's website so that the public could access to the main statistics aggregated per HEIs or the full higher education system as, including characteristics of the population, dropout rates, and graduation rates.

4 Data and Variables

In this section, I describe the SPADIES database and its characteristics. In the second part, I explain the variables that SPADIES uses, the program's definitions, and the new variables I created and included. In the final part of this section, I explain the final database used in the empirical analysis.

4.1 SPADIES Database

The SPADIES database is consists of the merger of three sources: MEN's National Information System for Higher Education (SNIES) database, the Colombian Institute for the Promotion of Higher Education (ICFES) database, and the HEIs semestral report.

The SNIES data is time-invariant and includes all the characteristics of the HEIs and their programs. The ICFES data is time-invariant as it is collected during the Saber 11 exam at the end of secondary school. The HEIs' report updates the information about its students every semester.

SPADIES uses the life history approach to collect its data. It means that SPADIES only tracks students who began college in 1998H1 or later. This paper uses SPADIES data from 1998 until 2017, about 8 million students. The dataset is an unbalanced panel per individual semester.

4.2 Variables

This subsection explains the variables that SPADIES has, how they are measured, and the new variables I created.

4.2.1 SPADIES Time-Invariant Variables

All the ICFES' variables collected during the Saber 11 exam include the Saber 11 exam score, gender, year of birth, number of siblings, the position among siblings if they live in their own home when presented the exam, household income, and the parents' education level. The exam result is an essential requirement to enroll in higher education, so all the students who enrolled in higher education have a score; all the other questions are optional for the students. This information is used for the students' characterization and the empirical analysis in this paper. SPADIES uses a standardized version of the Saber 11 test score. As the ICFES used a different score range over

time in the Saber 11 test score, MEN required SPADIES' developers to standardized the score by assigning each student their percentile on the Saber 11 exam in the period the student took the exam. The variable contains values from 1 to 100 depending on the student's percentile. However, I also created an additional dummy that takes the value of 1 if a student has a high score (above 66) and 0 otherwise. The MEN considers a high score when the value is over 90.

Gender is reported by the ICFES database and by the Freshmen report; SPADIES uses the data from the ICFES, but in the case of missing values, it completes it with the data from the Freshmen report. If both reports have no-data on gender, SPADIES runs an algorithm with a dictionary of names to assign it.

The year of birth is also reported in the ICFES database and the Freshmen's report. SPADIES uses the data from the ICFES database and, in the case of missing values, completes it with the Freshmen report's data. In case that there are missing values in both databases, SPADIES looks for the year in the identification number; if it is not found, the variable will be a missing value.

I use the number of siblings and the position among siblings as they appear in the SPADIES database. I have created a new set of dummy variables that take the value of 1 if the student's family lives in their own house, the mother has a college degree or higher, and the household income is above 3 minimum monthly wages, and 0 otherwise. SPADIES uses the data from the characteristics of the HEIs and their programs reported by SNIES. The HEIs characteristics include the sector (public or private), the category (Universities or Community Colleges), and their location. The program's characteristics include the level (bachelor or associate) and the field of knowledge.

Finally, SPADIES has a variable that qualifies the quality of the data reported per HEI. The grade depends on the number of semesters reported and the details of the data reported per semester. The grade could be A, B, or C; the highest grade is A. I created a dummy that takes the value of 1 in the HEI is qualified with A in the report as the cut of the database in 2017.

4.2.2 SPADIES Time-Variant Variables

SPADIES received three main reports per semester from each HEI: Freshmen, Graduates, and Enrolled. Each HEI also provides the students' id, their academic performance, information about financial or academic aid received, and the fields of study.

SPADIES looks for the students reported as Freshmen in the Enrolled report during every

semester until they are reported as Graduates; if the student is not found in the Enrolled report or Graduated report, their status changes.

After multiple meetings between presidents of HEIs, government authorities, and other scholars, the following statuses and definitions were agreed upon:

1. **Graduated:** SPADIES defines this status as a person who finished the coursework and earned a degree from a higher education program. I divided this category into two groups:

   (a) Graduated on time (those who graduated within one year of the expected time of graduation).

   (b) Graduated late (those who graduated more than 1 year after their expected time of graduation).

   Graduated late and Graduated on time encompass all individuals that graduated from higher education. The expected graduation year was estimated as five years for professional programs and three years for associate programs. So, $Graduated = Graduated\,on\,time + Graduated\,late$.

2. **Dropout:** SPADIES defined dropout as a student that has not been reported in the system or Graduated after 2 or more consecutive semesters as of 2017.

3. **Absent** students are those who missed only one semester and are not reported as Graduated.

4. **Active** is any student taking classes as of 2017.

5. The **student's cohort** is the year and semester (YYYY-SS) in which the student enrolled as a freshman.

6. **Freshmen students** are those enrolled in their first semester in an HEI-program.

7. The **time in the system** is the account of semesters that a student is reported as enrolled; it is not the number of semesters since the start of its cohort (e.g., the student could have been absent one semester). SPADIES records a timestamp per semester in each student's record when it can find the student in the semestral enrollment report.

Therefore, every semester, the student's time in the system, academic performance, and the report of aid received are updated. The Status is assigned at the moment of the cut of the database; in the

case of this document, it is 2017. However, I created three new variables to measure the transitions. I compare the time since they first enrolled in the college and the numbers of semesters reported in SPADIES. If there is a difference, it means that they left the school, so if their current status is graduated or active, it means that they have a transition. There are two types of transition:

1. **Transition from absent,** means that in a period "T" the student was "Absent" and in "T+1" he became "Active" or "Graduated". The dummy variable following these transitions takes the value of 1 in "T+1" and 0 otherwise.

2. **Transition from dropout**, means that in a period "T" the student was a "Dropout" and in "T>=3" he became "Active" or "Graduated". The dummy variable following these transitions takes the value of 1 when the transition ends and 0 otherwise.

Finally, I created the variable **time gap of transition** that counts the amount of periods during each transition; in the case of transitions from "Absent", it is always 1, but for the transitions from dropout, it is always bigger than 2.

4.3 Final Data

SPADIES merges the time-invariant information from the ICFES database and the SNIES with the reported data of Freshmen, Enrolled, and Graduates every semester. As in the definition of Freshmen students, the individual is defined by a student who is enrolled in a program in an HEI. Because SPADIES uses the life history approach, and I want to have periods with students in all stages of their program (from freshmen to seniors), I only use data for students reported as "Active" since 2002. Remember that to be an "Active" in 2002 means that the student was reported as "Freshmen" in any year before (and including) 2002, and they should appear in the report of Enrollment in 2002. After this change, the final dataset has information for 6.4 million students.

I created a dummy that identified the time when the student was enrolled in an HEI that had SPADIES installed. As a participant in the Universidad de Los Andes' team, I kept the record on when each institution received the visit from MEN and Universidad de Los Andes (described in the previous section); I use the semester of this visit as the start of the "treatment" of SPADIES on an HEI.

I included the unemployment rate estimated by the National Department of Statistics (DANE) by HEI's department and year.

The final database is an unbalanced panel individual-program-HEI and time, including the students' gender, year of birth, number of siblings, the position among siblings, a dummy if their family owns a home when they presented the Saber 11 exam, a dummy that takes the value of 1 if their mother has an University degree or higher and it is 0 otherwise, a dummy that takes the value of 1 if their household's income is higher than 3 minimum monthly wages and it is 0 otherwise, a dummy that takes the value of 1 if the program they are attending conducts to a bachelor level degree and it is 0 otherwise, a set of dummies that take the value of 1 depending on the area of knowledge of the program they are attending, a dummy that takes the value of 1 if their institution is public and it is 0 otherwise, a dummy that takes the value of 1 if their HEI is certified and it is 0 otherwise, a dummy that takes the value of 1 if their HEI is a main campus and it is 0 otherwise, a dummy that takes the value of 1 if the data from their HEI is A and it is 0 otherwise, a set of dummies that take the value of 1 depending on the HEI's region (Bogotá, Valle del Cauca, Antioquía, and Atlántico), a set of dummies that take the value of 1 depending on the round of implementation their HEI was included in the program, a dummy that takes the value of 1 if the period of Enrollment is later than the starting time of the program in their HEI, the unemployment rate by their HEI's department and the year they were enrolled, the status in the system, the transition from absent, the transition from dropout, the time gap of transition, a dummy that takes the value of 1 if the student is a dropout student and 0 otherwise, a dummy that takes the value of 1 if the student is graduated and 0 otherwise, and a variable that accounts the number of transitions that a student had during their program.

[Table 1 about here.]

[Table 2 about here.]

5 Model specification

To measure the impact of SPADIES in different outcomes, I use the following equation:

$$Eq:1 \qquad y_{it} = \beta_0 + \gamma SPADIES_{it} + \beta X_{it} + \epsilon_{it}$$

where y_{it} is a dummy that, depending on the model, measures the dropout rate, graduation rate, and transitions rate. In the case of Time gap, y_{it} is a continuous variable. The variable of interest is $SPADIES_{it}$, and it takes the value of one (1) if the individual is enrolled in a HEI in a period t when SPADIES was already installed and zero (0) otherwise. The vector of controls X_{it} includes time variant variables including academic performance, if the student received any type of assistance (financial, academic o "other type"), the time that the student has been enrolled in the HEI, and the departamental unemployment rate. This vector also includes time invariant variables as a dummy for females, the year of birth, the number of siblings, the position among siblings, a dummy if their family owns a home when they presented the Saber 11 exam, a dummy if their mother has an University degree or higher, a dummy 1 if their household's income is higher than 3 minimum monthly wages, a dummy if the program they are attending conducts to a bachelor level degree, a set of dummies depending on the area of knowledge of the program they are attending, a dummy if their HEI is public, a dummy if their HEI is certified , a dummy if their HEI is a main campus, a dummy if the data from their HEI is graded as A by the MEN, and a set of dummies depending on the HEI's region. $\alpha_{r(i)}$ comprends the fixed effects by HEI.

Finally, for the graph analysis, I use the same model presented in the Equation 1, but restricting the sample to the individuals that belong to the groups of the time invariant dummies. Therefore, I present the results of the variable of interest $SPADIES_{it}$ for the sub-samples (in example: for males and for females).

6 Results

The results section contains two parts: The first part presents a table with the general results of the impact of SPADIES including the controls of the different types of assistance. The second part shows the results for the coefficient of SPADIES from Equation 1 by different samples of student characteristics in charts per output.

The main results (Table 4) show that SPADIES had a significant impact in all the analyzed variables; in addition, the signs, including the controls, are the expected ones. SPADIES reduced the dropout rate by 1.2 percent, which means an impact of 4.3 percent; increased the graduation rate by 0.6 percent, an impact of 1.3 percent; and increased the on-time graduation rate by 0.6

percent, an impact of 1.9 percent. The transitions collapsed due to SPADIES. The probability of having a transition dropped 3.4%, an impact of 48 percent, while absentee transitions dropped 1.3%, an impact of 35 percent. However, the most significant drop was in transitions from dropout, which declined 2.1% with an impact of 63 percent. The latter implies that HEIs retained more students thanks to SPADIES and that the HEIs attracted students who were dropouts back to school. Finally, the time gap for each transition decreased by 0.18 semesters, an impact of 14 percent.

On the other hand, as expected, the greater the number of semesters completed, the lower the dropout rate, the higher the on-time graduation, and the greater the probability of having transitions. The greater the ratio of failed classes on the total of classes took in a semester, the greater the probability of dropping out and the lower the probability of graduating or on-time graduation. Finally, the Transitions increase if the ratio of missed classes increases.

Regarding the types of assistance provided to students during their time in school, I find that academic assistance had no impact. However, financial aid from HEIs and the other type of aid reduced the transitions and the transitions' time gap. However, public financial aid shows fascinating results. They are highly significant, and their impact is similar to the results of SPADIES, only except by the SPADIES result in the dropout rate.

Analyzing the impacts according to the students' characteristics or their schools, SPADIES had better results regarding the vulnerable population's dropout rate (the main focus on the training) -Figure 4-. SPADIES reduced the dropout rate, mainly for males with low household income and low academic skills. The household income is the only variable with a statistical difference between counterparts (i.e., High income, low income). Although the dropout rate has historically been lower in public institutions, and these institutions received the largest number of people after the expansion of the system at the beginning of the century, and thanks to SPADIES, they reduced the dropout rate more than HEIs from the private sector. There is no significant difference in the impacts according to the institutions' quality, the program's level, or the area of knowledge. However, SPADIES was much more effective in HEIs outside of Bogotá. In the case of the dropout rate, all the variables are statistically different from zero. The most robust and most significant effects in terms of dropout occurred in rounds 3 and 4.

As with the dropout rate results, the population that benefited the most in terms of graduation

from SPADIES (Figure 5) was males with a low-income level. SPADIES had better results in increasing the graduation rate for high-income individuals and high-quality HEIs (the only statistical difference between the characteristics) in Bachelor programs and STEM programs. There is no significant difference between the regions or the sector of the institution. Unlike the results in dropout, the graduation rate results are not significantly different from zero in all cases. On the other hand, on-time graduation (Figure 6) increased, mainly in low-income males in certified HEIs, Bachelor programs, and STEM programs. SPADIES does not seem to impact on-time graduation for Associate programs. HEIs from rounds 1 and 3 were the most effective in increasing the graduation rate and on-time graduation thanks to SPADIES.

Transitions decreased mainly in institutions outside of Bogotá; public, non-certified HEIs; and in all rounds except round 5 -perhaps without much time in the database to evidence improvement- (Figure 7). This last result is particularly relevant, as transitions from dropout at non-certified institutions were reduced significantly more than at certified institutions. Interestingly, the results indicate that high-qulity institutions avoided students becoming dropouts by bringing them back from the absent status (Figure 8). At the same time, non-certified HEIs recovered them from the dropout status (Figure 9). Finally, thanks to SPADIES, the gap time during transitions decreased mainly in the non-certified HEIs in Associate programs (consistent with what I found in Figure 9) and in people with low academic skills outside of Bogotá. Round 4 was the most effective in reducing the number of transitions and its time gap. (Figure 10).

[Table 3 about here.]

7 Conclusions

The new millennium brought with it an interesting challenge - the children of the "boomers" arrived to higher education. These new students were less prepared and had lower skills than their predecessors, increasing the risk that they become dropouts(Ferreyra et al., 2017).

Colombia developed a strategy that would use dynamic system tools to respond to the challenge; they called this strategy the "Educational Revolution"(Orozco Silva, 2010). One of MEN's tools was SPADIES, a software application to collect higher education students' data with a user-friendly interface that would help institutions target at-risk students for anti-dropout assistance(Ministerio

de Educación Nacional, 2008). In fact, the requirement to participate in this "revolution" modernized the protocols and records from the HEIs that, when used with SPADIES's tools, provided information to the different agents in this market: authorities in MEN and HEIs, students, and the public(Ministerio de Educación Nacional, 2010). All the agents were able to know "real time" statistics of the most important outcomes for the higher education system.

The flow of information was crucial; not only were HEIs receiving training about how to operate the dashboard, but they were also informed about strategies and protocols that other HEIs were using successfully to reduce the dropout rate. The HEIs were able to compare themselves to their peers in the main outcomes. Students were aware of what was happening as the dropout rates were more promoted in media, and these rates were also used in the promotion of the programs and the HEIs. The students were also the main target of a variety of developed and targeted programs using SPADIES. Thus, no matter the equilibrium model used the maximization of income or quality from Epple et al. (2006) or the effect of the peers from MacLeod and Urquiola (2015), the information provided by SPADIES impacted the outcomes rates in my analysis.

On the one hand, in Epple et al. (2006)–where a set of HEIs maximize quality and other profits, and the students attend them depending on their skills and financial capabilities–the students choose where to study depending on the institution's quality, the cost of the tuition, and the admission policies. HEIs with high quality will have more selection, where selection can be by prices, admission policies, or both. Reputation matters, so for high-quality HEIs, SPADIES will impact the selectivity; HEIs will accept only those students with a lower risk of becoming a dropout student. For the low-quality HEIs, SPADIES information becomes a benchmark for promotion to capturing more students. Finally, students will attend HEIs where their chances of getting the degree are high enough for their risk of becoming dropouts and the prices they paid.

On the other hand, the peer's effect from MacLeod and Urquiola (2015) triggered the competition among HEIs, so they were more selective, increased the number of programs to provide aid to students in risk, targeted their resources better, and traced their students better. One impressive result is the reduction of transitions, and it is no-doubt due to the competition among HEIs peers. Students also responded that they migrated to similar HEIs, where their peers had more efficient programs to reduce dropout rates or increase graduation rates. The competition helped to reduce the requirements to graduate, particularly in low-quality institutions, but the data I have do not

allow me to track those students. Future research exploring the migration between HEIs would be a substantial contribution to the literature. Analyzing how the increase in selection recomposed the population within the institutions and how it affected students' future income is an interesting topic to address for future research.

My research shows that SPADIES not only resulted in higher student retention, but also higher rates of on-time graduation and graduation overall. SPADIES achieved its primary purpose of reducing the dropout rate for all types of students, and in particular for the population most at-risk of dropping out - males with low household income and low academic skills (see Figure 4).

SPADIES reduced the probability of becoming a dropout student by 1.2%, increased the probability of graduation by 0.6%, and graduation on-time by 0.6%. A massive improvement in transitions supports these relevant results. SPADIES reduced the transitions by 3.4%, where the major improvement was the reduction of the transitions from dropouts (2.1%) and the time students were absent in 0.2 semesters (which means an impact of 14.4 percent). The transitions decreased mainly in public non-certified HEIs outside of Bogotá (Figure 7). However, results suggest that certified HEI bring back students from the absent status, while the non-certified HEIs recovered them from the dropout status. SPADIES reduced the gap time during transitions in the non-certified HEIs for Associate programs (Figure 9) and in students with low academic skills from outside Bogotá (Figure 10).

Interestingly, I find that academic assistance did not demonstrate impact on any of the outcomes, while private financial aid and the "other type of assistance" did result in a reduction in transitions and in the duration of the transition time gap (see Table 4).

Finally, I find that SPADIES improved the higher education system's efficiency because it helped decongest a bulging student population. SPADIES helped more students to graduate and to do so on-time. As a result, the quality of the system improved. By reducing the time it takes students to graduate, SPADIES was gradually reducing the burden of overpopulation on HEIs, which helped institutions better serve their students and serve more students. SPADIES served as a bridge to bring information to students and HEIs about their students to mitigate the vulnerabilities of the new population that were arriving into the system.

SPADIES was instrumental in breaking down the musical chairs game that was the higher education system in Colombia. By providing information to agents in the market, SPADIES improved

the system's efficiency by changing the future of many students who would have otherwise dropped out.

References

Adams, Jack E, "Affirmative Action in Higher Education: A Sourcebook," *The Journal of Higher Education*, 1984, *55* (1), 113.

Allende, Claudia, Francisco Gallego, and Christopher Neilson, "Approximating the Equilibrium Effects of Informed School Choice," *Working Paper*, 2019.

Asif, Raheela, Agathe Merceron, Syed Abbas Ali, and Najmi Ghani Haider, "Analyzing undergraduate students' performance using educational data mining," *Computers and Education*, 2017.

Auguste, Byron, Paul Kihn, and Matt Miller, "Closing the talent gap: Attracting and retaining top-third graduates to careers in teaching," Technical Report, Mckinsey & Company 2010.

Bank, Barbara J., Ricky L. Slavings, and Bruce J. Biddle, "Effects of Peer, Faculty, and Parental Influences on Students' Persistence," *Sociology of Education*, 1990.

Barro, Robert J., "Human capital: Growth, history, and policy - A session to honor Stanley Engerman: Human capital and growth," *American Economic Review*, 2001, *91* (2), 12–17.

Bassi, Marina, Matias Busso, Sergio Urzúa, and Jaime Vargas, "Desconectados Habilidades, educación y empleo en América Latina," 2012, p. 36.

Bean, John P., "Dropouts and turnover: The synthesis and test of a causal model of student attrition," *Research in Higher Education*, 1980.

Bean, John P, "Interaction Effects Based on Class Level in an Explanatory Model of College Student Dropout Syndrome," *American Educational Research Journal*, 1985.

Becker, Gary, "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 1962, *70* (S5), 9.

Bound, John and Sarah Turner, "Cohort Crowding: How Resources Affect Collegiate Attainment.," 2006, *12424*.

_ , **Brad Hershbein, and Bridget Terry Long**, "Playing the admissions game: Student reactions to increasing college competition," in "Journal of Economic Perspectives" 2009.

_ , **M Lovenheim, and Sarah Turner**, "Understanding the decrease in college completion rates and the increased time to the baccalaureate degree," *Ann Arbor, Mich.: Population Studies Center, University of Michigan Institute for Social Research*, 2007, pp. 1–91.

Braxton, John M, "Student Success in College: Creating Conditions that Matter," *Journal of College Student Development*, 2007.

Cabrera, Alberto, Linda Serra Hagedorn, Amaury Nora, and Ernest Pascarella, "Differential impacts of academic and social experiences on college-related behavioral outcomes across different ethnic and gender groups at four-year institutions," *Research in Higher Education*, 1996.

_ , **María Castañeda, and Amaury Nora**, "College Persistence: Structural Equiations Modeling Test of Integrated Model of Student Retention," *The Journal of Higher Education*, 1993, *64* (2), 123–139.

Campbell, Carol and Ben Levin, "Using data to support educational improvement," *Educational Assessment, Evaluation and Accountability*, 2009.

Cárdenas, E, "Estudio de la deserción estudiantil en programas de ingeniería en la Universidad Nacional de Colombia." Tesis de maestria en dirección universitaria, Universidad de los Andes 1996.

Castaño, Elkin, Santiago Gallón, Karoll Gómez, and Johanna Vásquez, "Análisis de los factores asociados a la deserción y graduación estudiantil universitaria," *Lecturas de Economia*, 2006, *65*, 9–36.

Córtes, Hernán, Luis Gallego, and Gerardo Rodríguez, "The engineering faculty today: an approach towards consolidating academic indicators," *Ingeniería e Investigación*, 2011, *31* (1).

Cruces, Guillermo, "Inequality in education: evidence for Latin America," *Falling Inequality in Latin ...*, 2014.

de Roux, Nicolás and Evan Riehl, "Disrupted academic careers: the returns to time off after high school," 2019.

Epple, Dennis, Richard Romano, and Holger Sieg, "Admission, tuition, and financial aid policies in the market for higher education," 2006.

Facundo-Díaz, Ángel, "Análisis sobre la deserción en la educación superior a distancia y virtual: el caso de la UNAD - COLOMBIA," *Revista de investigaciones UNAD*, 2009, *8* (2).

Ferreyra, María Marta, Ciro Avitabile, Javier Botero Álvarez, Francisco Haimovich Paz, and Sergio Urzúa, *Momento decisivo: La educación superior en América Latina y el Caribe* 2017.

Gentsch, Kerstin, "How admission policy shapes college access: Evidence from two sectors." PhD dissertation 2016.

Glewwe, Paul W., Eric Hanushek, Sarah D Humpage, and Renato Ravina, "School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010," 2011.

Hanushek, Eric and Ludger Woessmann, "Does Early Tracking Affect Educational Inequality and Performance? Differences-in-Differences Evidence across Countries," *Economic Journal*, 2006, *116* (115), C63–C76.

_ **and** _ , "Schooling, educational achievement, and the Latin American growth puzzle," *Journal of Development Economics*, 2012, *99* (2), 497–512.

Hastings, Justine S and Jeffrey M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Herrera-Prada, Luis Omar, "Determinantes de la tasa de graduación y graduación a tiempo en la educación superior de Colombia 1998-2010," *Coyuntura Económica*, 2013, *XLIII* (1), 143.177.

_ , "The Economic Consequences of Setting Foot in a College in Colombia," 2020.

ICFES, "Estudio de la deserción estudiantil en la educación superior en Colombia.," Technical Report, Universidad Nacional - ICFES, Bogotá 2002.

Kerr, Sari Pekkala, Tuomas Pekkarinen, Matti Sarvimäki, and Roope Uusitalo, "Post-secondary education and information on labor market prospects: A randomized field experiment," *Labour Economics*, 2020.

Lucio, Ricardo and Mariana Serrano, *La Educación Superior: Tendencias y Políticas Estatales.* 1992.

MacLeod, W. Bentley and Miguel Urquiola, "Reputation and school competition," 2015.

Ministerio de Educación Nacional, "Análisis de determinantes de la deserción en la educación superior colombiana con base en el SPADIES.," Technical Report, Ministerio de Educación Nacional - Universidad de los Andes, Bogotá 2008.

_ , "Deserción estudiantil en la educación superior colombiana. Metodología de seguimiento, diagnóstico y elementos para su prevención.," Technical Report, Bogotá 2009.

_ , "La Revolución Educativa 2002 - 2010. Informe de gestión.," Technical Report, Bogotá 2010.

_ , "Boletín Educación Superior," Technical Report, Bogotá 2017.

Mizala, Alejandra and Miguel Urquiola, "School markets: The impact of information approximating schools' effectiveness," *Journal of Development Economics*, 2013.

ONU, "Objetivos de Desarrollo del Milenio," *Naciones Unidas*, 2013, p. 64.

Orozco Silva, Luis Enrique, *La Política de Cobertura: eje de la revolución educativa, 2002-2008.*, Bogotá: Ediciones Uniandes, 2010.

_ **, Alberto Roa Valero, and Luis Carlos Castillo Gómez**, "La Educación Superior en Colombia," 2011, pp. 1–83.

_ **, Javier Medina Vásquez, María Pérez Piñeros, and Alberto Roa Valero**, "Informe Colombia," in "Proyecto Informe Sobre la Educación Superior en Iberoamérica" 2006.

Park, Kang H and Peter M Kerr, "Determinants of academic performace: A multinomial Logit Approach," *Journal of Economic Educations*, 1990, *21*, 101–111.

Portes, Pedro R. and Spencer Salas, "The dream deferred: Why multicultural education fails to close the achievement gap. A cultural historical analysis," *Cultura y Educación*, 2007.

Porto, Alberto and Luciano Di Gresia, "Rendimiento de estudiantes universitarios y sus determinantes," *Revista de Economía y Estadística*, 2004.

Saavedra, Juan Esteban and Carlos Medina, "Formación para el Trabajo en Colombia *," 2012.

Sánchez, Fabio, Margarita Quirós, Carlos Reverón, and Alberto Rodríguez, "Equidad Social En El Acceso Y Permanencia En La Universidad Pública Determinantes Y Factores Asociados," 2002, *7191*, 1–48.

SPADIES, "Reporte Modelo SPADIES al MEN," Technical Report, Universidad de los Andes, Bogotá 2008.

St. John, Edward P., Alberto F. Cabrera, Amaury Nora, and Eric H. Asker, "Economic Influences on Persistence Reconsidered," in "Reworking the Student Departure Puzzle," Vanderbilt University Press, oct 2016, pp. 29–47.

Stratton, Leslie S., Dennis M. O'Toole, and James N. Wetzel, "A multinomial logit model of college stopout and dropout behavior," *Economics of Education Review*, 2008, *27* (3), 319–331.

Tinto, Vincent, "Dropout from Higher Education: A Theoretical Synthesis of Recent Research," *Review of Educational Research*, 1975.

_ , "Limits of Theory and Practice in Student Attrition," *The Journal of Higher Education*, 1982.

Trow, Martin, "Problems in the Transition from Elite to Mass Higher Education," *International Review of Education*, 1974, *18*, 61–82.

Unesco, "Using data to improve the quality of education," 2020.

Universidad Nacional de Colombia, *Cuestión de Supervivencia. Graduación Deserción y Rezago*, Bogotá: Beta Impresores Ltda, 2007.

Table 1: Students description

Variable	Obs.	Mean	Standard Deviation
Female	6,411,602	.506	.5
Year of birth	5,641,802	1988	7.96
Secondary test score	6,411,602	61.5	28.5
No. of siblings	4,201,498	2.05	1.49
Position among siblings	4,194,907	1.61	1.12
Live in their own home	4,194,907	.791	.406
Mother with University degree or higher	6,411,602	.175	.38
Household with high income	6,411,602	.384	.486
Bachelor level	6,411,602	.66	.474
Agriculture and Related	6,411,602	.018	.131
Art, Architecture and Related	6,411,602	.039	.193
Education	6,411,602	.087	.282
Health Related	6,411,602	.07	.255
Social Sciences	6,411,602	.176	.381
Business, Economics and Related	6,411,602	.298	.457
Theology	6,411,602	0	.003
Engineering	6,411,602	.289	.453
Natural Sciences and Maths	6,411,602	.02	.14

Source: ICFES-HEIs. STEM=Engineering+Natual Sciences and Math. Field of studies variables from HEIs' report.

Table 2: Higher Education System data description

Variable	Obs.	Mean	Standard Deviation
Public institution	6,411,602	.421	.494
High-quality certified institution	6,411,602	.289	.453
Main campus	6,411,602	.654	.476
Good data report	6,411,602	.959	.199
Institution located in Bogota	6,411,602	.419	.493
Institution located in Valle del Cauca	6,411,602	.067	.25
Institution located in Antioquia	6,411,602	.139	.346
Institution located in Atlantico	6,411,602	.058	.235
HEI from 1 round of implementation including sub-locations	6,411,602	.261	.439
HEI from 2 round of implementation including sub-locations	6,411,602	.181	.385
HEI from 3 round of implementation including sub-locations	6,411,602	.229	.42
HEI from 4 round of implementation including sub-locations	6,411,602	.323	.468
HEI from 5 round of implementation including sub-locations	6,411,602	.006	.076
Unemployment rate by department year	6,411,602	10.9	2.66

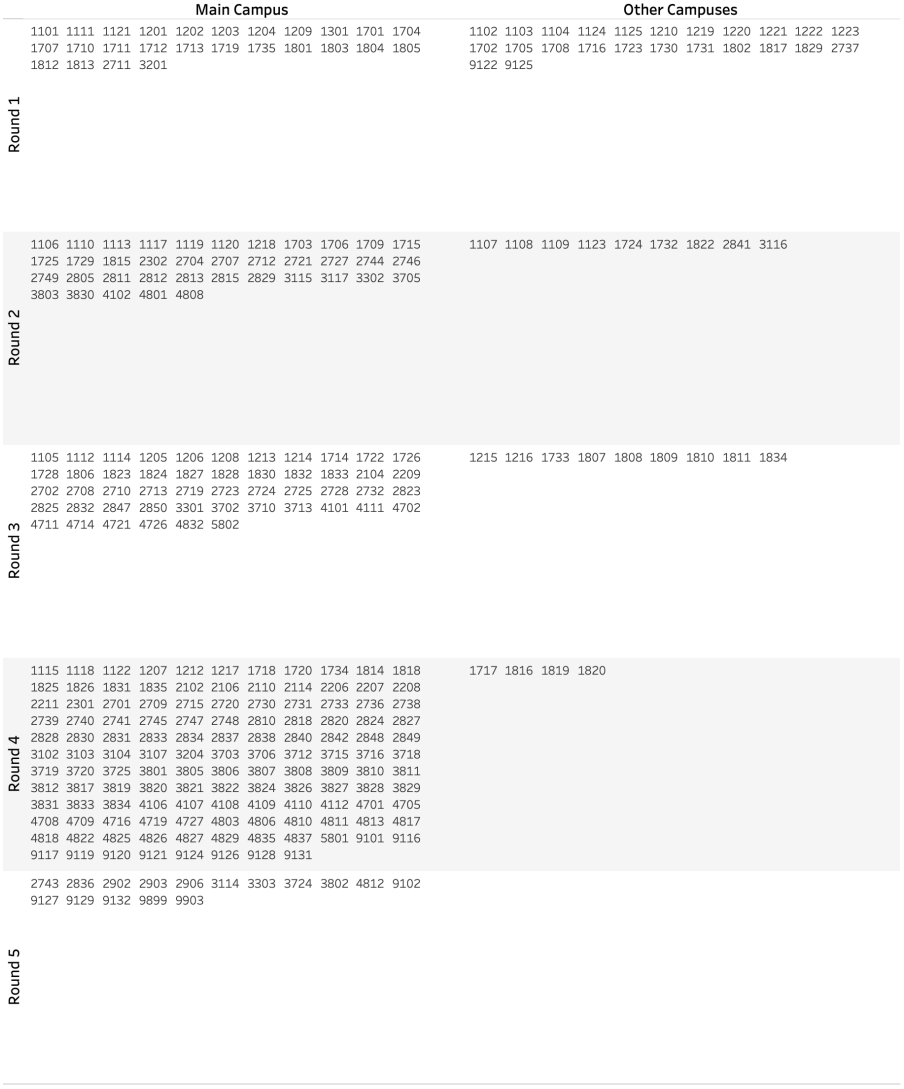
Notes: A good data report was an index that the MEN and Universidad de Los Andes created to track the development of the indicators in the HEIs. The measure was done by grades A, B, or C. SPADIES' installation was performed when HEI reached "A" grade; the last wave included the remaining HEIs. The grade corresponds to an evaluation of the full data submission (the report for first-year students, enrolled, and graduated students since 1998). A low grade means that the HEI reported a short run frame (e.g., since 2007) or did not report at least one of the tables required (first-year students, enrolled, or graduates). The Unemployment rate from DANE (National statistical office).

Table 3: SPADIES impact on educational outputs (Main results)

	Drop-out rate	Graduation rate	On-time graduation rate	Number of transitions	Transitions from absent	Transitions from drop-out	Time gap in the transition
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SPADIES	-1.230***	.559**	.591***	-3.425***	-1.280***	-2.144***	-.181***
	(.170)	(.227)	(.127)	(.459)	(.402)	(.320)	(.038)
Semester	-.048**	.180***	.090***	.366***	.089***	.277***	-.031***
	(.023)	(.049)	(.023)	(.033)	(.019)	(.023)	(.003)
Share of failed classes	.910***	-.686***	-.614***	1.264***	.471***	.793***	-.100***
	(.147)	(.124)	(.111)	(.286)	(.151)	(.179)	(.013)
Academic aid	-1.798	2.932	1.457	-.957	-.583	-.374	.036
	(1.365)	(2.379)	(1.182)	(.624)	(.416)	(.605)	(.071)
Private financial aid	-.108	.187	.120	-1.140***	-.578**	-.563***	-.031*
	(.132)	(.125)	(.083)	(.350)	(.279)	(.157)	(.017)
Other aid	-1.299	1.334	.740	-2.475***	-1.308**	-1.167**	-.060
	(.863)	(1.311)	(.678)	(.776)	(.518)	(.481)	(.061)
Public financial aid	-.275***	.546***	.434***	-2.114***	-1.073***	-1.041***	-.020
	(.105)	(.109)	(.089)	(.257)	(.187)	(.149)	(.014)
Unemployment rate	.351***	.186**	.072*	.776***	.187***	.588***	.043***
	(.048)	(.091)	(.043)	(.106)	(.058)	(.073)	(.009)
Academic aid x SPADIES	2.157	-2.317	-1.232	.854	.502	.353	.012
	(1.386)	(1.928)	(1.045)	(.623)	(.429)	(.609)	(.068)
Private financial aid x SPADIES	-.181	-.171	-.091	.285	.124	.160	.049**
	(.185)	(.180)	(.141)	(.463)	(.398)	(.192)	(.022)
Other aid x SPADIES	1.328	-.837	-.606	2.232***	1.242**	.990**	.132**
	(.890)	(.939)	(.569)	(.755)	(.500)	(.501)	(.065)
Public financial aid x SPADIES	.455***	.398**	.380***	2.387***	1.031***	1.356***	.120***
	(.120)	(.155)	(.110)	(.295)	(.235)	(.177)	(.018)
Constant	23.907***	41.807***	3.621***	-2.696**	1.455**	-4.150***	.753***
	(.548)	(1.147)	(.535)	(1.167)	(.714)	(.770)	(.083)
Observations	30,962,767	30,962,767	30,962,767	30,962,767	30,962,767	30,962,767	30,962,767
R^2	.006	.005	.002	.004	.001	.005	.039
Number of id	6,411,602	6,411,602	6,411,602	6,411,602	6,411,602	6,411,602	6,411,602
Dependent variable mean	28.3	43.7	31.1	7.07	3.68	3.4	1.26

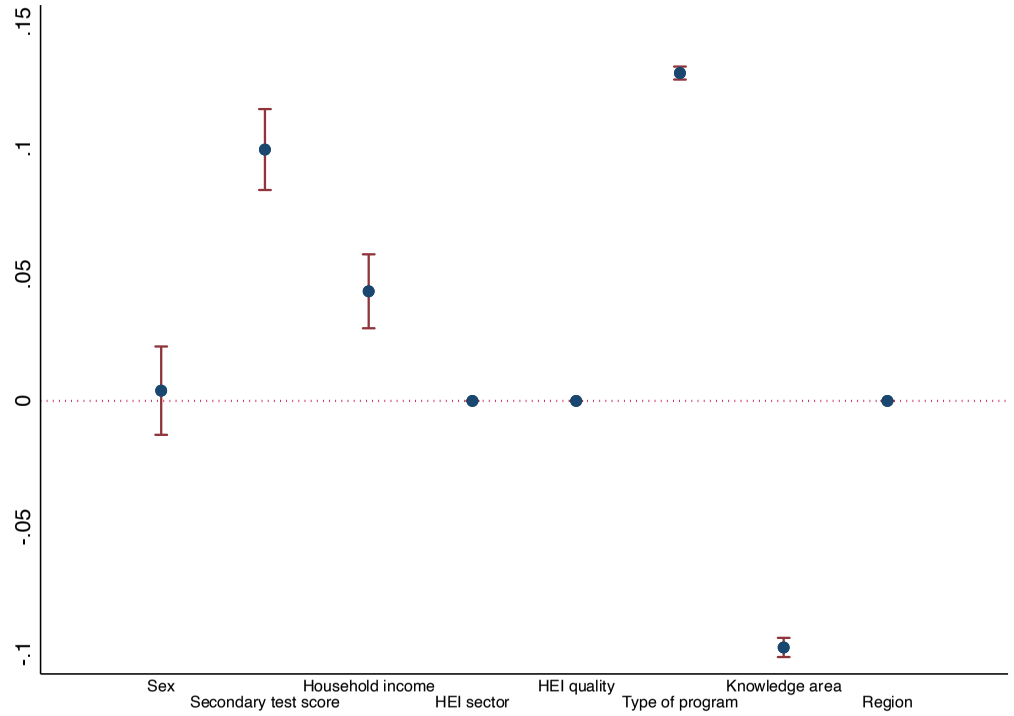
Notes: SPADIES would take the value of 1 if SPADIES was installed in the HEI when the student was enrolled and 0 otherwise. The semester is an account of semesters that a student has been enrolled. Share of failed classes is the ratio of non-approved classes divided by the total of classes that the student took in a period. Academic aid is a tutoring or remedial course that the HEI provided without cost to the students. Private financial aids are internships or grants given by the HEI to the students; private financial loans were not considered private financial aid. Public financial aid is the student's loan received provided by the government's bureau (ICETEX). Other aids are all the other programs that HEIs provided to students (e.g., professional orientation or psychological help). The unemployment rate is at the department level reported by the Colombian statistical office (DANE). Robust standard errors clustered at the HEI level. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: HEIs distribution by round



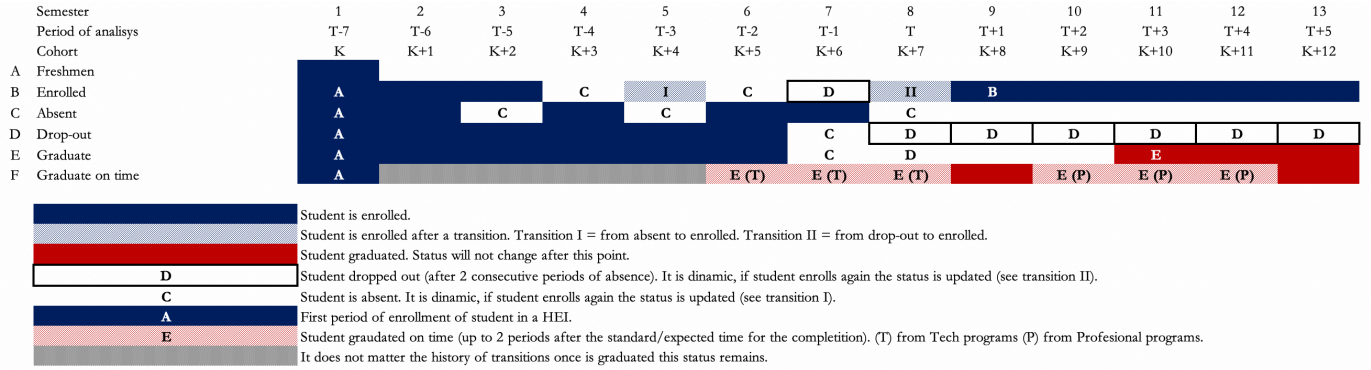
Notes: Five rounds -Round 1(2005-06), Round 2 (2006-07), Round 3 (2007), Round 4 (2008-09), and Round 5 (2010-11)- were necessary to complete all the HEIs' information into the system. The MEN and the Universidad de Los Andes visited and installed the dashboard in the main HEI; MEN expected that the main HEI shared SPADIES with its other campuses."Other campuses" were included in the same round that its parent HEIs; they are an extension of one of the main HEIs in other regions (e,g. Universidad Nacional de Colombia code 1101 is the main public national university, located in Bogotá, and it is the parent of 1102 that is the campus located in Medellín. In some cases, the "Other campuses" administration is autonomous, and in other cases, it administration depends directly on the main campus. There is not a rule about this).

Figure 2: SPADIES assignation balance



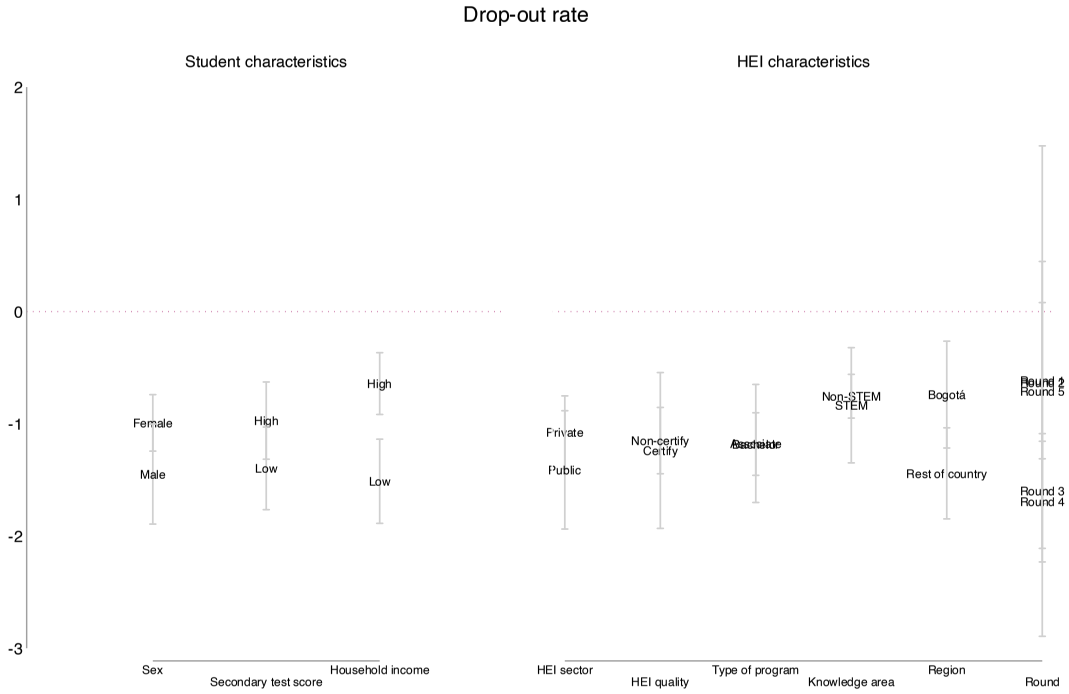
Notes: The figure shows the regression coefficients explaining the treatment variable using the students' main characteristics and HEIs. Whiskers show the 95% confidence interval. Sex is a dummy that is 1 if the student is female. The secondary test score is a dummy that is 1 if the student is in the top 10

Figure 3: SPADIES status analysis



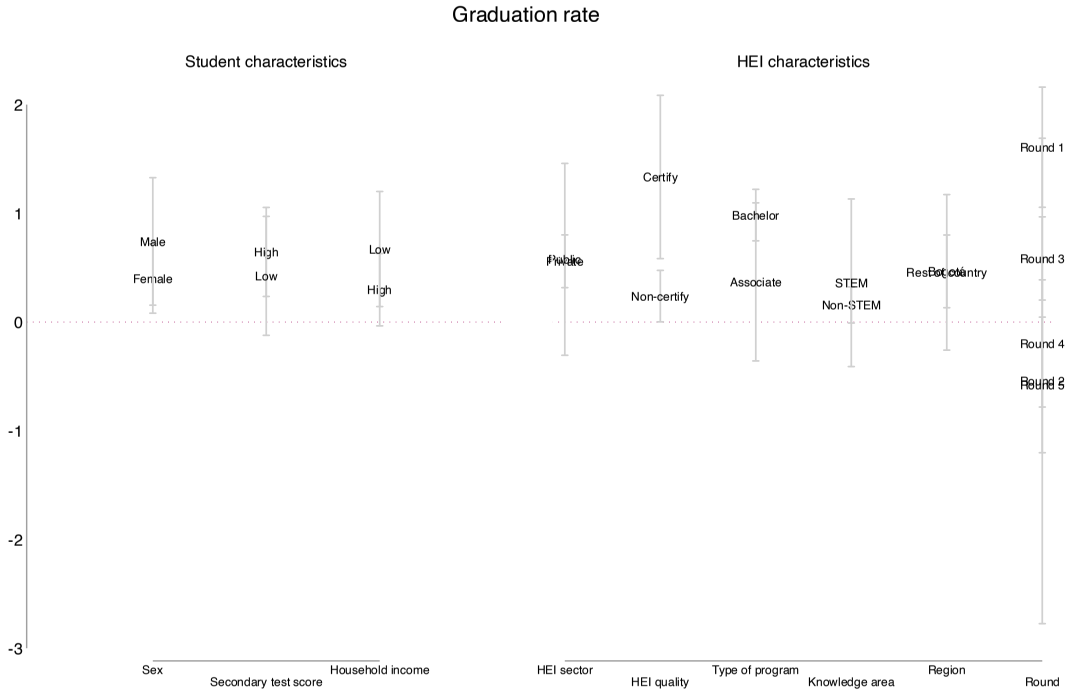
Notes: All the statuses in SPADIES are dynamic in time except by Graduated. Transition I shows an enrolled status once the student was absent per 1 semester and returned to the system. Transition II shows that after being considered a dropout student the status was updated to "enrolled" once the student returned to the system. Letters from "A" to "F" show the status of the student. Each row shows an example of the status in T. "E(T)" if graduated from a tech and "E(P)" from a professional program. Letters in cells are reference to examples of dynamic statuses in time. Time is recorded as the times an student is enrolled; i.e. for cohort K, if our analysis is in period "T", an student who has been enrolled all the semesters will be in 8th semester, but none of the examples account this. In the example "B Enrolled" the student will account 5th, as the student has been enrolled in 5 semesters and considered absent in 2 semesters and dropout in 1. In the example "E Graduated" the student will account 6 semesters (1 as absent and 1 as dropout).

Figure 4: SPADIES impact coefficient in dropout rate by characteristic



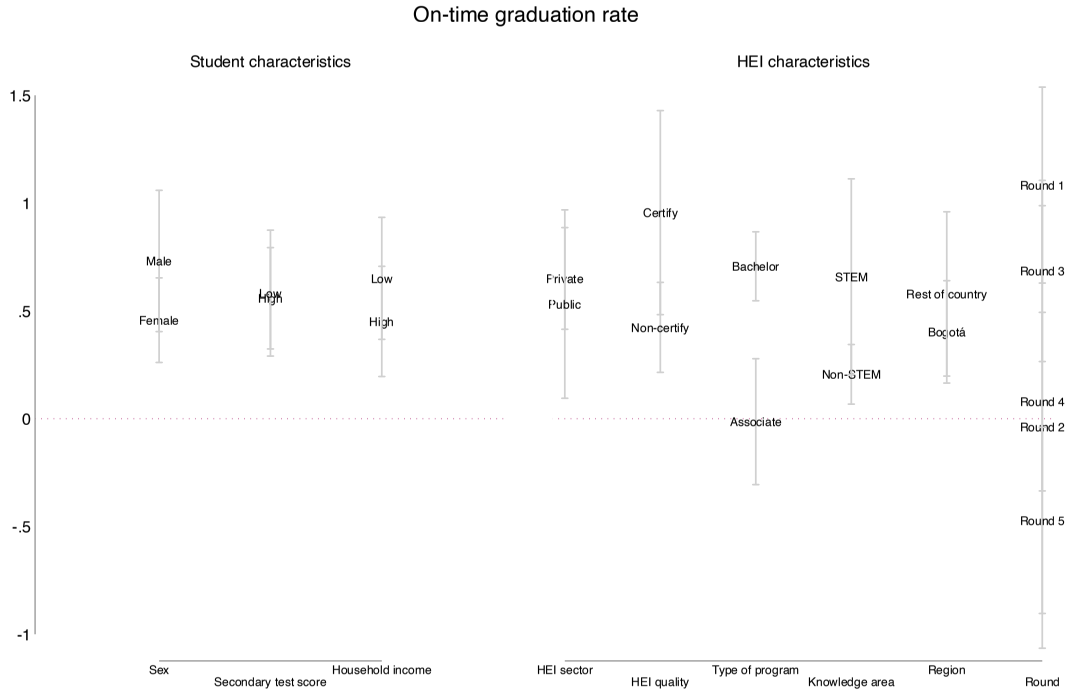
Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.

Figure 5: SPADIES impact coefficient in graduation rate by characteristic



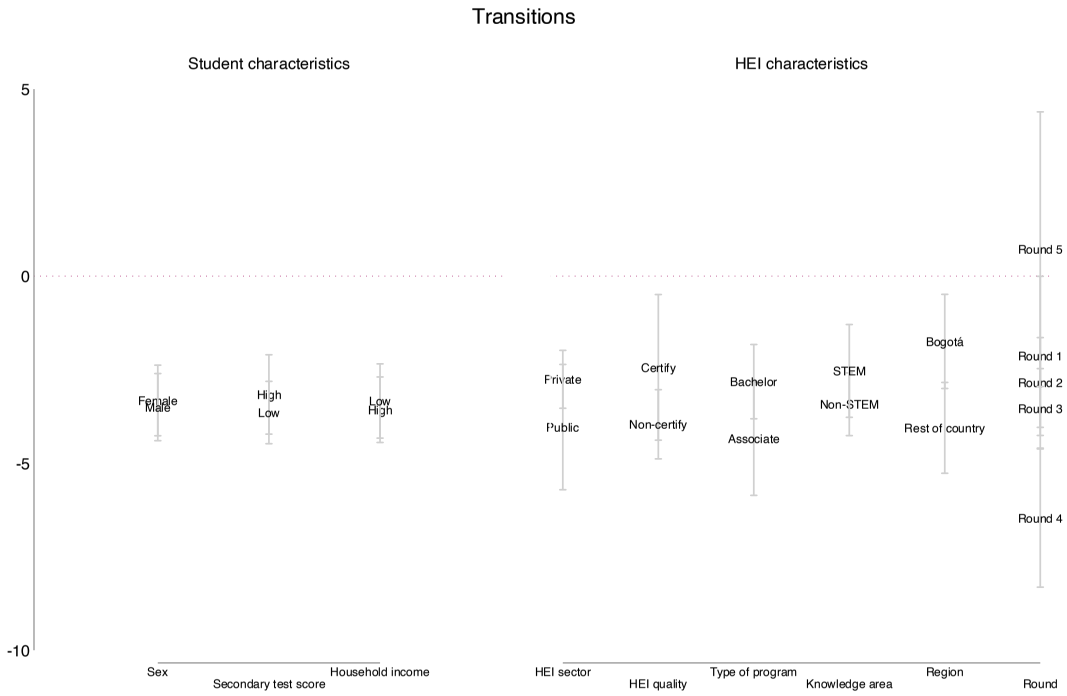
Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.

Figure 6: SPADIES impact coefficient in on-time graduation rate by characteristic



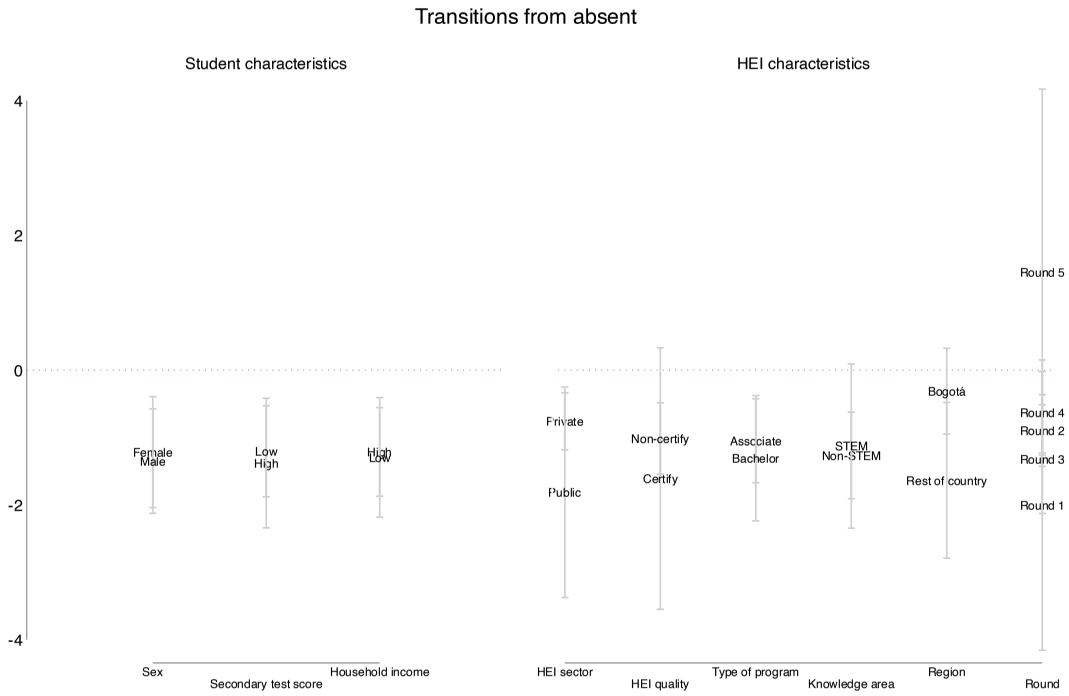
Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.

Figure 7: SPADIES impact coefficient in total of transitions by characteristic



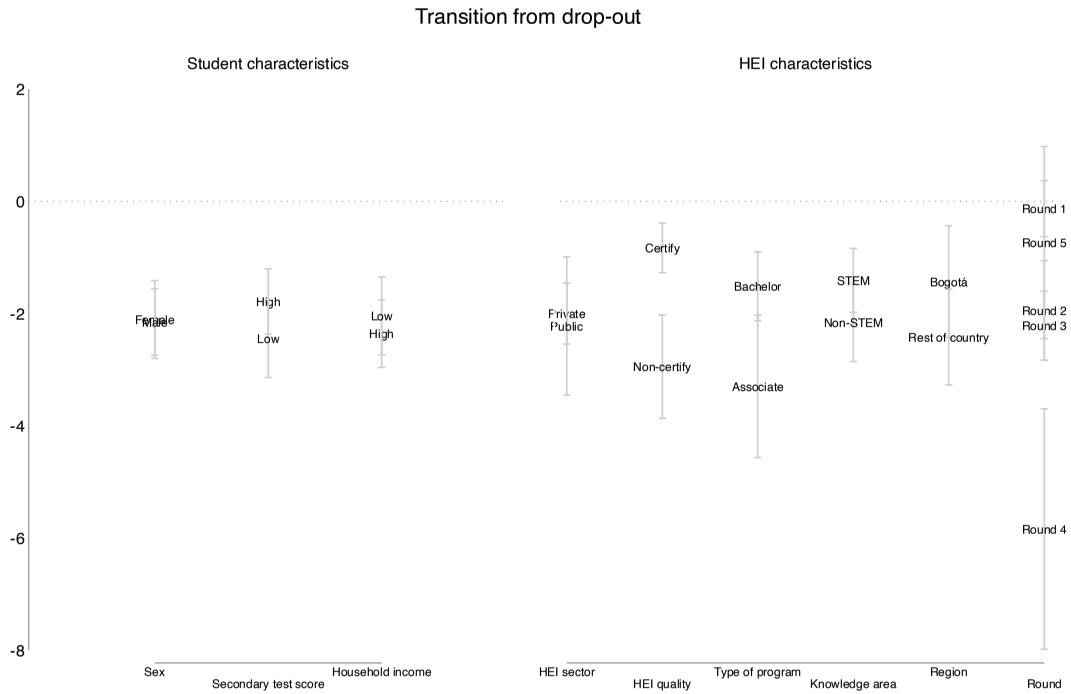
Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.

Figure 8: SPADIES impact coefficient in transitions from absent by characteristic



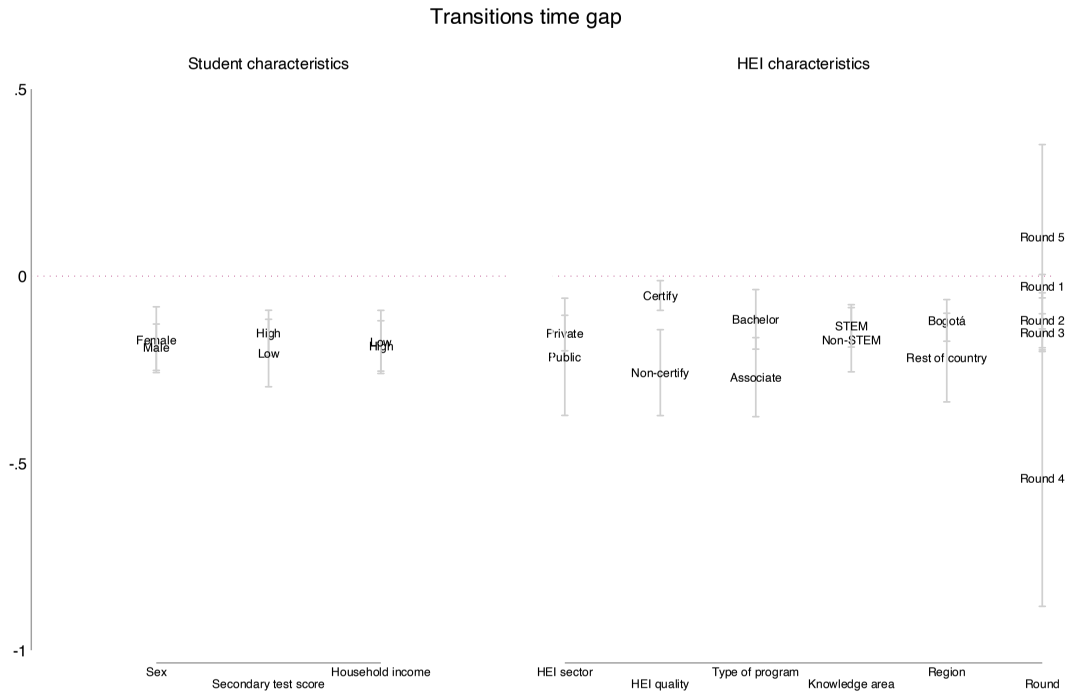
Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.

Figure 9: SPADIES impact coefficient in transitions from dropout status by characteristic



Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.

Figure 10: SPADIES impact coefficient in time gap during transitions by characteristic



Notes: Chart shows the coefficient for the Equation 1 using different disaggregations for the student characteristics (Gender, Secondary test score, Household income), or the HEI-program characteristics. Whiskers show the 95% confidence interval. Secondary test score is the standarized version of the score in the range 0-100. An score over 90 points (including) is considered a High score in the secondary test. High income households have monthly income higher than 3 times the monthly minimum wage. Technology programs (Associate) are 2-year programs, University/professional (Bachelor) programs are 4-year or 5-year programs. STEM refers to science and engineering knowledge areas.