Higher Education Responses to Accountability

Anaely Machado*
Rafael Terra*
Maria Tannuri-Pianto*

July 24, 2022

Abstract

This paper estimates the impact of accountability scores on Brazilian higher education outcomes. We explore a natural experiment: the introduction of an accountability system for Brazilian undergraduate programs named SINAES that was implemented by the Ministry of Education in 2004. The design of the evaluation system enables us to implement a regression discontinuity strategy. We test whether program quality is sensitive to negative reinforcement, such as punishments imposed when a minimum threshold is not attained. We also test whether program administrators seek higher evaluation scores as a form of advertisement to attract prospective students. Our results show that program administrators respond to the threat of punishment by improving program quality in the next evaluation cycle, but we cannot determine whether administrators seek higher grades in order to advertise their programs and increase enrollments or to improve the quality and prestige of their programs.

JEL Classification: H75, I21, I23, I28
Keywords: Regression Discontinuity, Accountability, Higher Education, Impact Evaluation

*Faculdade de Economia, Administração e Contabilidade, Campus Darcy Ribeiro, Universidade de Brasília, Asa Norte, Brasília, DF 70910-900, Brazil

1 Introduction

Accountability systems aim to improve education quality. Such systems evaluate institutions on the basis of student performance on standardized tests and other instruments that reflect quality in terms of infrastructure and faculty profiles. Evaluation results can be used to inform people about institutional quality and to support regulatory initiatives. Nevertheless, accountability systems are most common in basic education, though such systems may certainly be introduced at any level or in any type of education, such as higher education, when governments want to promote quality and guarantee the rational use of public funds. Moreover, candidates for higher education programs can be better informed when making decisions based on publicly disclosed grades and choose programs in which students perform the best, setting positive incentives for undergraduate programs to always strive for improvement. On the other hand, negative incentives such as punishments imposed on the administrators of low-scoring programs can also encourage institutions to improve educational quality.

While plausible, the reaction of higher education institutions (HEIs) to the introduction of accountability incentives remains largely an empirically unexplored subject. This paper attempts to address this question by investigating the effects of negative or (weakly) positive incentives – introduced by an accountability system – on higher education outcomes[1].

To this end, we explore a natural experiment created by the Brazilian Ministry of Education, which enacted its current higher education accountability system in 2007. Thereafter, undergraduate programs are evaluated every three years based on the results of a standardized exam, the National Exam of Student Performance (ENADE)[2], faculty profiles and student feedback. The results in each of these dimensions are used to compose a continuous index that summarizes program performance, the Preliminary Program Grade (CPC)[3], which is used to classify programs into 5 possible levels based on sharp cutoffs[4]. The Ministry of Education publicly discloses the performance grades and uses them to regulate undergraduate expansion and activity, conditioning approval to renew programs on whether minimum achievement standards have been met. That is, the programs must obtain a minimum level of 3 to be recognized by the

[1]Previously, Rezende (2010) attempted to evaluate the impact of accountability on higher education by conducting a regression analysis of observational panel data on Brazilian undergraduate programs for the 1996-2003 period. The author concluded that scores on National Program Exam (ENC – Exame Nacional de Cursos in Portuguese) increased program slots and improved faculty profiles.

[2]Exame Nacional de Desempenho dos Estudantes in Portuguese.

[3]Conceito Preliminar de Curso in Portuguese.

[4]We refer to the continuous CPC score as CPC_{score}, while we use CPC_{level} when referring to the CPC levels.

federal authorities under penalty of suspension if they fail to reach that level. Since information about the quality of higher education institutions can also contribute to students' choice over undergraduate programs, institutions can advertise their good results to attract more students and expand their programs. Hence, discontinuities originating in the CPC level assignment rule create an opportunity to evaluate the short run effects of accountability incentives on undergraduate outcomes in the years following the evaluations.

Our research is closely related to the literature on the response to accountability by various agents (school administrators, teachers, families and others). For example, previous research has examined performance improvements in low-performing schools after the receipt of their evaluations in Brazil (Camargo et al. (2018)), Mexico (De Hoyos et al. (2017)), Portugal (Nunes et al. (2015)), South Korea (Woo et al. (2015)), and the United States, specifically Chicago, New York City, Florida and Wisconsin (Neal and Schanzenbach (2010); Rockoff and Turner (2010); Rouse et al. (2013); Chiang (2009); Chakrabarti (2014); Deming et al. (2016)). The literature has also identified the impacts of accountability ratings on resource allocation and administrator behavior (Figlio and Winicki (2005); Craig et al. (2013, 2015)). In addition, accountability evaluations are related to student and teacher flows from low- to high-performing schools (Feng et al. (2018)).

In particular, the interest in the impact of accountability on higher education has gained attention in the literature. Evidence on the impact of an evaluation system on offers, faculty profiles and program attractiveness are found in Rezende (2010) and Bowman and Bastedo (2009) for HEIs in Brazil and the United States. Less explored is the effect of higher education quality on labor market outcomes (see Canaan and Mouganie (2018)).

Public disclosure of the results of evaluations and college/school rankings contribute to institutional reputations which in turn influence student behavior when choosing an institution and undergraduate program (Bowman and Bastedo (2009); Rezende (2010)). This means that the expected effects of accountability depend, at least partially, on the publicity and transparency of the results obtained by institutions and programs (Hastings and Weinstein (2008); Deming and Figlio (2016)). In addition, the literature suggests that the impact of accountability on education is related to the incentives faced by different agents according to their performance. For example, Figlio and Rouse (2006) and Rouse et al. (2013) investigated the relevance of the voucher threat and state oversight on the impacts of accountability. While the former found that the impacts are mainly driven by grading stigma, the latter concluded that the changes in instructional policies and practices were a result of accountability pressure. In summary, the literature indicates that the different accountability systems affect educational outcomes provided that they are related to explicit rewards and sanctions.

On the other hand, the literature also identifies critical issues for the effectiveness of accountability systems. For example, if educational assessments are tied to specific measures, then organizations seek to improve their performance related to those measures at the potential cost of other outcomes of interest due to their maximization behavior (Deming and Figlio (2016)). That is, the system sends a signal to society about what is most valued, and then, the administrators pursue that goal. Deciding which measures to include in an evaluation system for HEIs is even more difficult since different fields of study and organizations have different curricula and purposes (Deming and Figlio (2016)). Additionally, the long-run effectiveness of accountability systems may be limited by the strategic behavior of the agents, while institutional rankings tend to be effective only after the first publication (Deming and Figlio (2016); Bowman and Bastedo (2013))[5].

As is evident from the literature mentioned above, most previous research has focused on elementary to secondary education, most probably because of the absence of a structured accountability system for higher education in most countries or, when such a system exists, the lack of rules that would enable quasi-experimental evaluations of accountability systems for this level of education. To the best of our knowledge, this paper represents the first attempt to evaluate the impact of accountability on higher education using a regression discontinuity design (RDD) approach[6].

In particular, our empirical strategy is similar to that in previous works that have explored discontinuities in grade assignment rules to measure the impact of accountability, such as Chiang (2009), Rockoff and Turner (2010), Rouse et al. (2013), Chakrabarti (2014), Craig et al. (2015), Woo et al. (2015), Woo et al. (2015), Holbein and Ladd (2017), Canaan and Mouganie (2018), Feng et al. (2018).

Our main results suggest that undergraduate program administrators respond to negative incentives imposed by the federal authority – such as threats of closure, supervisory commission visits or punishment via the withdrawal of

[5]Bowman and Bastedo (2013) studied the impact of higher education rankings and found that the initial rankings influenced peer assessments of reputation in subsequent surveys but that second-year rankings were not related to changes in reputation in the third year, and these results may be associated with the anchoring theory.

[6]Rezende (2010) studied the effects of accountability on Brazilian higher education based on OLS estimations. In addition, the accountability system was replaced by the current system, investigated in this paper. Despite the use of an RDD approach, Canaan and Mouganie (2018) mainly explores the labor market returns to higher education accountability for low-skilled students.

recognition – by improving program accountability index values in comparison to other programs in the following evaluation cycle. Programs with evaluations that fall below the low-performance threshold, i.e., that have a CPC_{level} equal to 1 or 2 ($CPC_{score} < 1.945$), achieved better outcomes in the next evaluation cycle in terms of performance, faculty, infrastructure and quality overall. We also find evidence that programs just above the recognition threshold ($CPC_{score} \geq 1.945$) increase their program slots, receive more applications and admit more new students than programs just below the same threshold. We do not find clear patterns around the threshold that assigns $CPC_{level} = 5$ ($CPC_{score} \geq 3.945$), the maximum grade, which we expected programs could have used as an advertisement.

We also test whether accountability has heterogeneous effects on private and public HEIs. The Brazilian higher education system is composed of both private and public institutions. Public institutions are supported by public resources, students who attend public institutions do not pay fees, and faculty and staff enjoy job stability. These characteristics probably reduce the potential negative effects of a bad evaluation for public programs. On the other hand, in addition to competitive pressure from the private market, private institutions have more positive incentives to pursue quality in order to access public programs that offer scholarships and student loans.

The results suggest that even public institutions react to evaluation incentives. Since public institutions are not subject to the positive incentives of access to scholarships or student loans, we conclude that negative incentives – i.e., the threat of punishment – dominate their reaction. However, the magnitude of the reaction to low scores is greater among private institutions, which suggests that positive incentives may also affect administrator behavior, though it could also be the case that the more pronounced reaction among private institutions is the result of administrators having "skin in the game" and always trying to attract more students to keep their jobs, whereas public administrators enjoy job stability.

Finally, we conclude that the observed impacts are associated with clear punishment rules, while the achievement of higher grade levels does not significantly impact program effort nor candidate perceptions of future returns.

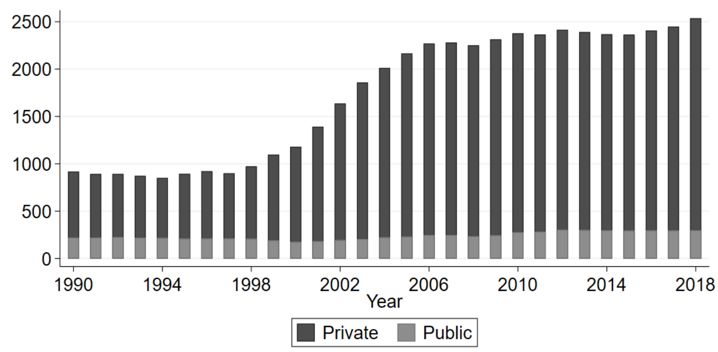
This paper is structured as follows. We describe the Brazilian higher education system and its accountability system in section 2. Section 3 describes the data and presents descriptive statistics. The empirical strategy is described in section 4. Section5 presents a discussion of the results and robustness tests, and section6 concludes.

2 The Brazilian higher education system

2.1 The recent expansion of the Brazilian higher education system

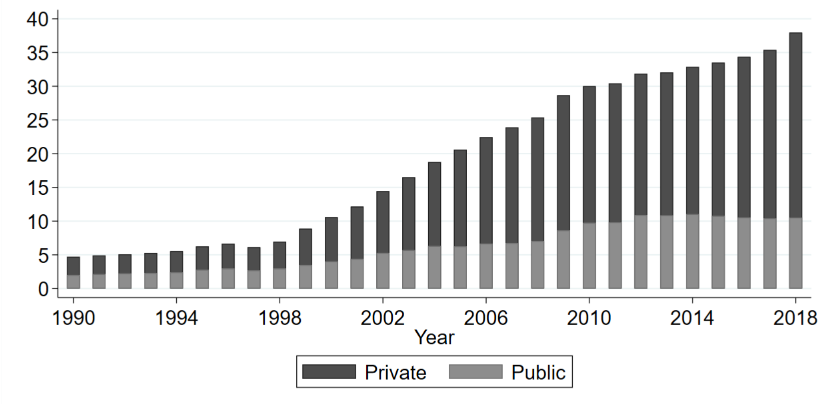
The Brazilian higher education system consists of public and private institutions. Public institutions may be fully supported by the federal government, as is the case for federal universities and federal institutes; by state governments, as is the case for state universities; and, in some cases, by municipal governments. Public institutions cannot charge tuition or fees, and faculty and staff enjoy legal job security after a three-year probationary period. In contrast, private HEIs charge tuition and fees from their students. There are various types of private institutions: publicly traded companies, private limited companies, Christian colleges and universities, think tanks, and foundations. Employees typically do not enjoy job security – although there are a few cases of institutions granting tenure to some professors.

The Brazilian higher education system has expanded significantly since the 1990s. Figure 1 shows that the number of HEIs tripled during that decade, which was led by the expansion in the number of private institutions. The number of undergraduate programs has evolved similarly, with the private sector representing more than 70% of the increase in undergraduate programs – see figure 2. Enrollments increased from 1.5 million in 1990 to almost 8.5 millions in 2018; see figure 3.



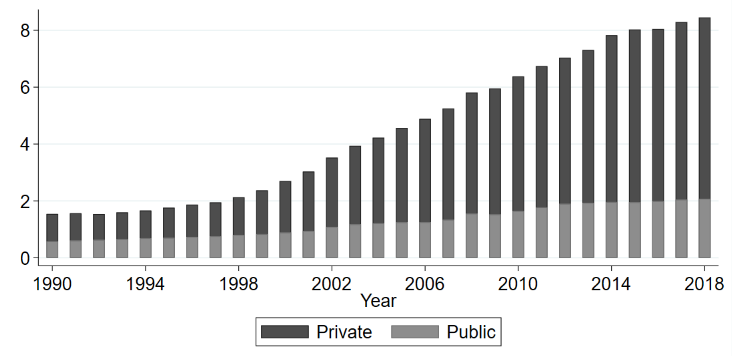
Source: Anísio Teixeira National Institute for Educational Studies and Research (INEP).

Figure 1: Number of higher education institutions in Brazil



Source: INEP.

Figure 2: Number of undergraduate programs in Brazil (in thousands)



Source: INEP.

Figure 3: Number of enrollments in undergraduate programs in Brazil (in millions)

The expansion of Brazilian higher education can be attributed to a few critical factors: the increasing number of students completing a high school level, changes in regulations that facilitated the entry of new institutions into the higher education market, and public policies that promoted higher education (Rezende (2010); OECD (2018)).

According to the School Census[7] reported by the Anísio Teixeira National Institute for Educational Studies and Research (INEP), [8] an agency within the Ministry of Education, the number of students graduating from high school increased from 960 thousand in 1995 to more than 2 million in 2018. During the same period, the Higher Education Census [9], also conducted by INEP, shows that the number of applications to undergraduate programs increased from 2.7

[7]Censo Escolar in Portuguese. Available at `http://portal.inep.gov.br/web/guest/censo-escolar`.

[8]Instituto Nacional de Estudos e Pesquisas Educacionais "Anísio Teixeira" in Portuguese.

[9]Censo da Educação Superior in Portuguese. Available at `http://portal.inep.gov.br/web/guest/censo-da-educacao-superior`.

to 12.4 million[10]. Although enrollments in high schools have been decreasing in recent years in Brazil (because of demographics and improvements in school progress), there is still high demand for undergraduate programs as evidenced by the increase in applications.

A second explanation for the higher education expansion in Brazil relates to changes in the regulations established in the 1990s that facilitated the market entry of new institutions and the creation of new undergraduate programs as long as such institutions and programs underwent periodic assessment for accreditation and recognition of diplomas (Rezende (2010); OECD (2018)).

Lastly, the Brazilian federal government sought the expansion of higher education by financing tuition in private institutions and expanding the number of slots in public universities. The federal authorities created the Student Loan Fund (Fies), an alternative source of credit for students to obtain educational loans and pay for their studies in private undergraduate programs, and the University for All Program (ProUni), which provides full scholarships for students in private HEIs[11]. Together, Fies and ProUni account for 22% of private enrollments and are regarded as important contributors to the increase in higher education enrollments among private institutions (Corbucci et al. (2016)).

In 2014, the federal government enacted the National Plan for Education (PNE)[12] for the period between 2014 and 2024. That plan set specific goals for increasing enrollment in public higher education institutions, as well as for improving the quality of education and access to higher education among socioeconomically disadvantaged students (OECD (2018)), thus reinforcing the role of the state in setting the conditions for the development of higher education.

2.2 The accountability system for Brazilian undergraduate programs

In the last two decades, the Brazilian government and Brazilian society have discussed the relevance of an accountability system for assessing, monitoring and assuring the quality of HEIs in face of the intended expansion of undergraduate programs and enrollments (Inep (2009); OECD (2018)). In 2004, the National System of Higher Education Evaluation (SINAES)[13] was established[14]. This system guides the Ministry of Education in its decisions about the accredita-

[10] Numbers for on-site undergraduate programs.

[11] See Law No 10260 from July 12, 2001, and Law No 11096 from January 13, 2005.

[12] See Law No 13005 from June 26, 2014.

[13] Sistema Nacional de Avaliação do Ensino Superior in Portuguese.

[14] Previous efforts to evaluate higher education include the Institutional Evaluation Program for Brazilian Universities (Paiub – Programa de Avaliação Institucional das Universidades Brasileiras in Portuguese), a voluntary evaluation for universities introduced in 1993, and the ENC, a standardized exam for undergraduate students in effect between 1996 and 2003. Graduate programs, in turn, have been evaluated by the General Coordination for the Improvement of Higher Education Personnel (CAPES) since 1976.

tion of institutions and the authorization for and recognition of undergraduate programs[15]. SINAES evaluates all private and federal public institutions, accounting for 91% of total enrollments in Brazilian undergraduate programs[16], and is administered by INEP.

SINAES sets rules and procedures for monitoring and evaluating undergraduate programs in order to act on results indicating low-performance programs. Accordingly, every three years, INEP calculates a CPC value for each undergraduate program.[17]. The CPC reflects the overall "quality of the program"; it is a composite index that summarizes (a) student performance, (b) teaching staff profiles and (c) feedback from students about the program. The CPC formula is given in equation 1.

$$\begin{aligned} CPC_{score} = 0.2 \cdot ENADE_c + 0.35 \cdot IDD_c + 0.075 \cdot NM_c + 0.15 \cdot ND_c \\ +0.075 \cdot NR_c + 0.075 \cdot NO_c + 0.05 \cdot NF_c + 0.025 \cdot NA_c \end{aligned} \tag{1}$$

The ENADE index evaluates learning quality and reflects student results on the ENADE, a standardized exam taken by students in their senior year covering the core disciplines of each program. ENADE results feed into the ENADE Index, which consists of the mean grade achieved by students in each discipline. ENADE results also feed into the Index for the Difference between Observed and Expected Performance (IDD),[18] which measures the value added by the higher education programs by comparing the grades achieved by students on the ENADE with their grades from the National High School Exam (ENEM)[19]. Together, the ENADE and IDD indexes account for more than half of the total weight in the CPC_{score} .

The quality of the faculty is evaluated by the proportion of its members with a master's degree (NM), the proportion with a PhD (ND), and the proportion of full or part-time faculty (NR).

Lastly, students complete questionnaires before taking the ENADE test wherein they provide feedback about the undergraduate program they attended. The student responses are used to produce indexes for teaching and learning (NO), infrastructure (NF) and academic and professional opportunities (NA).

[15]See Law 10861 from April 14, 2004.

[16]State- or municipality-controlled institutions can voluntarily participate in SINAES, as they are subject to local legislation and regulations regarding education.

[17]See Regulatory Ordinance no 560 from July 9th, 2019. See also Technical note n.58 from 2020 for the CPC methodology.

[18]Indicador da Diferença entre os Desempenhos Observado e Esperado in Portuguese.

[19]Exame Nacional do Ensino Médio in Portuguese. ENEM is a national exam that evaluates the quality of high school education. Its results are also used as an entrance exam for the main universities – public or private – and as a criterion for receiving scholarships and loans.

Each of the indexes that make up the CPC are standardized and rescaled to range between 0 and 5. Weights sum to one and are distributed according to equation 1. These index values are calculated for each program every three years, following the ENADE cycle, which determines the fields evaluated each year[20].

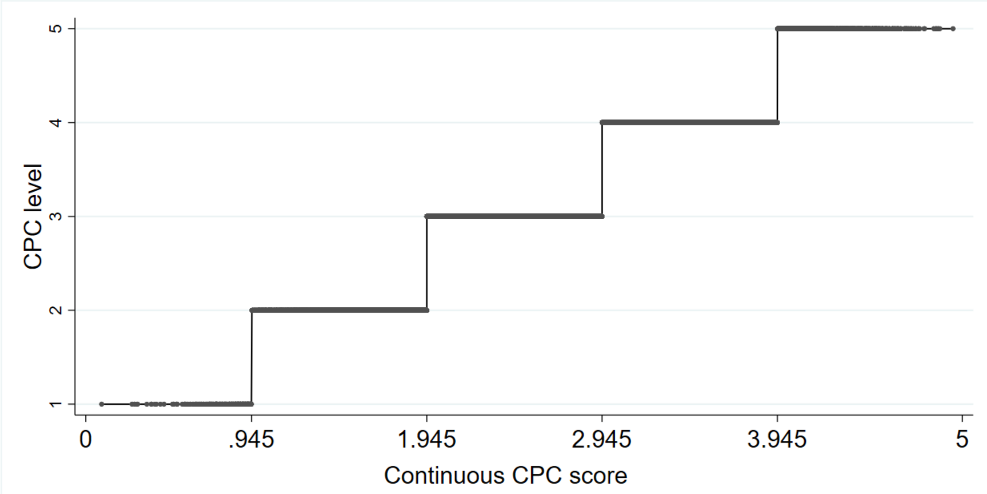
The CPC also influences the HEI general quality index, named the General Index of Programs (IGC)[21]. Coupled with graduate program scores[22], CPC scores are used to calculate the IGC, which is calculated as the mean of the graduate and CPC program scores weighted by the number of students enrolled in each program and degree level. INEP updates the IGC every year with the results of the current evaluation cycle.

Based on the CPC score – which is continuous and ranges from 0 to 5 –, the programs are classified into quality levels (i.e., CPC levels) – which are discrete and range from 1 to 5. We use the notation CPC_{score} and CPC_{level} to refer to the continuous score and the level, respectively. Programs with a CPC score below the threshold of $CPC_{score} < 0.945$ are classified as having a CPC_{level} of 1. A CPC_{score} equal to or above 0.945 but less than 1.945 results in a CPC_{level} of 2. The CPC_{level} is equal to 3 when the CPC_{score} is equal to or above 1.945 but less than 2.945. Level 4 is attained whenever CPC_{score} is equal to or above 2.945 but less than 3.945. Finally, programs with a CPC_{level} of 5 are assigned to the "excellence programs" category, i.e., the category of those programs whose CPC_{score} is equal to or greater than 3.945. Figure 4 shows the empirical relation between CPC_{score} and CPC_{level} (as determined by the rules described above), which makes evident the existence of sharp discontinuities in program level designations.

[20] See Regulatory Ordinance no 40 from December 12, 2007.

[21] Índice Geral de Cursos in Portuguese.

[22] Every four years, master and doctorate programs in different fields are evaluated by CAPES. See Regulatory Ordinance n. 59 from March 21, 2017.



Source: SINAES Tables (INEP). Authors' elaboration.

Figure 4: Empirical relation between CPC_{level} and CPC_{score}

Note: The CPC_{score} is used to classify programs into 5 possible levels. The figure illustrates the rule that determines the CPC_{level}, which makes evident the existence of sharp discontinuities in program level designations with multiple thresholds.

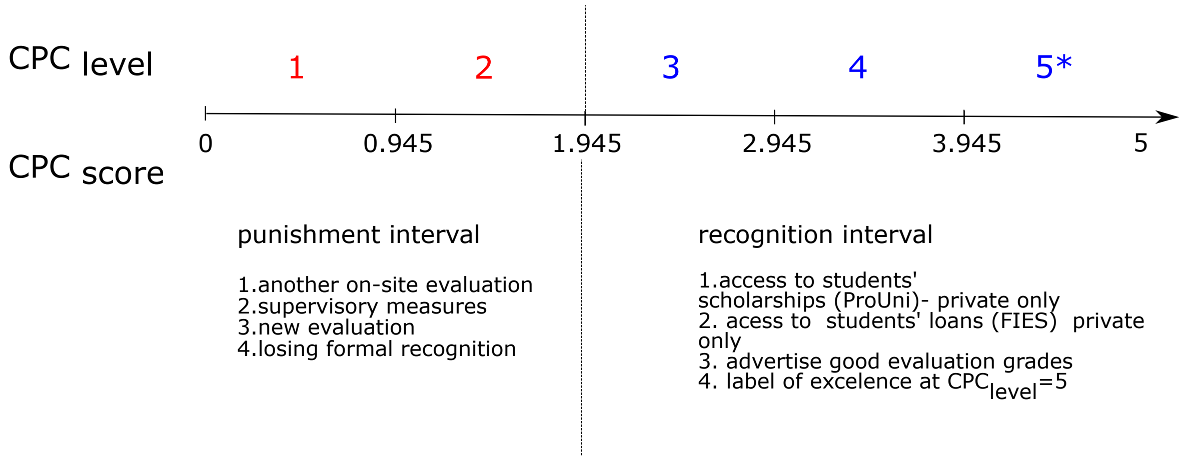
Table 1 shows how programs transition from the CPC level achieved in year t0 to the level achieved in the next evaluation period, t+3, i.e., in the following evaluation. Between 2007 and 2018, 18.5% of programs transitioned to a lower score in their next evaluation and 26.8% climbed to a higher level, while 54.7% remained at the same level in their subsequent evaluation.

Table 1: Rating transition in CPC_{level}

CPC_{level} in t	CPC_{level} in t+3					
	1	2	3	4	5	Total
1	0.02%	0.12%	0.14%	0.04%	0.00%	0.32%
2	0.15%	3.53%	9.26%	2.14%	0.08%	15.14%
3	0.09%	5.58%	34.84%	13.73%	0.42%	54.66%
4	0.02%	0.58%	10.25%	15.94%	0.90%	27.70%
5	0.00%	0.01%	0.29%	1.48%	0.40%	2.18%
Total	0.27%	9.83%	54.78%	33.32%	1.80%	100.00%

Source: SINAES Tables (INEP). Authors' elaboration.

The SINAES results determine the accreditation process for HEIs and their undergraduate programs as well as their access to publicly funded scholarships and student loans. Figure 5 summarizes the potential bonuses and penalties associated with each quality level.



Source: The authors.

Figure 5: Quality levels and corresponding bonuses and penalties

Programs that receive a CPC_{level} of 3 or above have their recognition renewed automatically. Programs with an unsatisfactory CPC_{level}, i.e., those that are classified as level 1 or 2, are subject to an additional on-site evaluation by an external reviewing commission[23]. This second evaluation involves a questionnaire – completed by the external commission – about the faculty (30% weight), the infrastructure (30%) and teaching and learning policies and practices (40%), resulting in a new index named the Program Index (CC)[24], with a CC of 3 or above being the criterion for the renewal of program recognition (OECD (2018)). In the event those programs still fail to achieve a satisfactory assessment (3 or above), the number of program slots must be reduced and the institution must sign a compromise protocol with the federal government in order to establish goals for improving quality. If the program still does not improve its evaluation scores, its formal recognition may be suspended or canceled and any diplomas issued will not be valid.

An unsatisfactory CPC also limits the participation of the institutions and programs in publicly funded programs for higher education. For example, the current legislation excludes programs evaluated at CPC levels 1 or 2 from accessing Fies, a federal government fund that provides student loans, or ProUni, a federal program that grants scholarships for disadvantaged and minority students.[25]

These accountability results are informative for society, people interested in ap-

[23]Neglect to fulfil that obligation may result in penalties such as the temporary suspension of new enrollments, the revocation of the HEI's authorization to operate, suspension of program recognition and, for public institutions, warnings or the suspension of the person in charge of the evaluation process within the institution.

[24]Conceito de Curso in Portuguese.

[25]See Laws n. 10.260 from July 2001 and n. 13.530 from December 2017 regarding the FIES regulations. See Law n. 11.096 from January 2005 and Normative Ordinance n. 22 from November 2012 regarding the ProUni regulations.

plying to higher education and undergraduate students. They also enable HEIs to seek to improve their programs and conform them to the quality standards needed to continue functioning.

In addition, HEIs can also advertise their evaluation results to attract more students[26]. If the expected economic return of obtaining a CPC_{level} of 5 is sufficiently high, i.e., the revenue increase surpasses the costs (including opportunity costs), institutions will invest in the pursuit of that evaluation level – and will not invest if the costs exceed expected revenues.

2.3 Potential effects of the Brazilian higher education accountability system

As described above, the Brazilian higher education accountability system assigns quality levels to each undergraduate program based on an assignment rule that generates discontinuities. Based on this rule, we test the impacts of falling just above each cutoff relative to falling just below the same cutoff. Since each cutoff is associated with different mechanisms that would affect agents' behavior (see figure 5 in the previous section), we also expect to find different impacts depending on the cutoff analyzed.

First, since the cutoff that assigns programs to $CPC_{level} = 2$ does not imply any incentives or penalties that differ from those imposed on programs that receive a $CPC_{level} = 1$, falling just above or just below this cutoff may have no impact on administrator behavior. Similarly, because both of these levels are associated with the same risk of having diplomas invalidates, students and families would potentially not prefer programs with a $CPC_{level} = 2$ over those with a $CPC_{level} = 1$. This means that we do not expect the accountability system to have strong effects around the cutoff $CPC_{score} = 0.945$.

Second, the cutoff that determines whether a program is assigned to $CPC_{level} = 3$ is strongly associated with the sanctions and benefits of having a recognized program, which means that this cutoff potentially affects the behavior of both members of society and program administrators. For those programs to the left of the cutoff ($CPC_{score} < 1.945$), we expect administrators to react to their low performance by investing in the resources needed to obtain a better result in the next evaluation. This may be achieved by improving the infrastructure of the institution, hiring more professors or changing pedagogical strategies, for example. On the other hand, when applying to HEIs, students and families may prefer programs that score at least the minimum level needed to have their diplomas formally recognized by the federal government. That is, we

[26]Figure 11 in Appendix A illustrates how institutions use their results for advertisement.

would expect an increase in offers and applications for programs that are to the right of this cutoff ($CPC_{score} \geqslant 1.945$) relative to programs that are to the left.

Finally, achieving a higher level of quality ($CPC_{level} = 4$ or $CPC_{level} = 5$) does not imply any additional bonuses nor does it guarantee more resources for the institution. However, achieving a higher quality grade ($CPC_{level} = 5$) can be a signal to society of the high performance of the programs and thus can be used as a form of positive marketing to attract more (and better) students[27]. Based on this line of reasoning, we predict that the administrators of those programs that receive a $CPC_{score} < 3.945$ increase effort in order to seek the highest quality rating.

In summary, we expect accountability grades to have stronger effects on programs around the cutoff that determines the minimum quality level needed to be recognized ($CPC_{level} = 3$) and for programs around the cutoff for $CPC_{level} = 5$, which is a signal of the high performance of such programs.

3 The data

We obtain our data from INEP – the main federal authority for education evaluation in Brazil. The first dataset consists of the quality index files, which contain annual assessment results for undergraduate programs from 2007 to 2018. The aforementioned files list the CPC scores and levels for each field of study and institution[28]. Because assessment results are aggregated by field of study and institution, we use microdata from the ENADE to identify each undergraduate program within these fields of study and institutions. Therefore, we organize the data so that our unit of observation is the program.

The accountability result tables also present the undergraduate programs' performance in each component of the CPC: faculty characteristics, mean student performance and student feedback about the program. Regarding the faculty profiles[29], we observe the percentage of faculty with a PhD, the percentage with a master's degree, and the percentage with full-time appointments (dedicating 40 hours or more per week to the program with which they are associated). Combining data on faculty and enrollments from the Higher Education Census (described below), we also estimate the ratio of students to faculty members[30].

[27] We would also expect a potential increase in tuitions for programs that received the highest quality rating. Unfortunately, there is no publicly available data on program tuition to test this hypothesis.

[28] Until 2015, evaluations were conducted at the field and institution level, so if an institution had two or more programs in the same field, all programs within that field received the same CPC score. From 2015 onwards, evaluations have been conducted for each program separately.

[29] Public data from the Higher Education Census do not contain information about faculty profiles for each program. This information is only available in files with the SINAES results.

[30] The number of faculty members in each undergraduate program was not published for the years 2012, 2013

Student performance is captured by two indexes, one of which is the mean score achieved by students on the ENADE and the other consists of the mean value added by the program (obtained by comparing the ENADE and ENEM results), named the IDD. Feedback from students is summarized in three indexes, which also range from 0 to 5. One index is for infrastructure, another is for learning and teaching, and the final index is about perceptions of professional and academic opportunities[31].

We also analyze the institution's overall evaluation grade, given by the IGC (the general grade assigned to each HEI) to take into account the quality of the institutions and the effect of their reputation on their programs' reactions to the disclosure of their evaluation grades.

The second dataset is from the Higher Education Census for the years 2007 to 2018. The microdata on programs and students in this dataset provide information on the number of students enrolled, the number of slots and applications in the selection processes, the number of new students and the dropout rate[32]. The data also include a variable that indicates the status of the program (i.e., whether the program is still open). All data is organized at the program level.

We discard observations from online undergraduate programs[33], as the accountability system for Brazilian higher education was developed primarily for evaluating on-site programs and does not take into account the specificities of online education OECD (2018).

We paired information on the SINAES evaluation from year t with information from year t+3 in the SINAES tables for each undergraduate program. In addition, we analyzed variables from the Census for t+1, t+2, and t+3.

Table 2 presents a summary of the characteristics of undergraduate programs by CPC_{level} between 2007 and 2018. We see that the number of programs classified as CPC level 1 or 2 has decreased since 2009, whereas the number of programs with a CPC level of 3, 4 or 5 has increased significantly over the same period.

Although public HEIs are fewer in number, they tend to be relatively more likely to be classified at having a CPC level of 4 or 5. In particular, within the highest level, public programs are almost as common as private programs.

---

and 2018. Therefore, those years are excluded in regressions over variables that depend on this information (specifically, the number of faculty and the number of students per faculty member).

[31] Available from 2013 onwards.

[32] We calculate the dropout rate as the percentage of students whose situation is characterized as "inactive enrollment", "canceled enrollment" or "transferred to another program in the same HEI". We assume that each of these situations represents a temporary or permanent interruption of the program, which negatively affects the total number students who complete the program.

[33] Educação à Distância (EaD) in Portuguese.

Universities are responsible for most of the programs classified into the higher levels. Table 2 also shows that the distribution of areas of study is similar over different CPC levels. Finally, as expected, the best-performing programs are concentrated among HEIs with the highest IGC scores.

Table 2: Characteristics of undergraduate programs by CPC_{level}

	CPC_{level}				
	1	2	3	4	5
Number of programs					
2007-2009	120	3,720	6,355	2,091	243
2010-2012	67	2,254	8,197	4,360	377
2013-2015	43	1,989	10,210	5,418	307
2016-2018	85	1,973	11,547	7,787	443
Total	230	7,963	24,762	11,869	927
Distribution by type of administration (%)					
Private	74.92	82.20	77.09	61.72	50.88
Public	25.08	17.80	22.91	38.28	49.12
Distribution by type of academic organization (%)					
University	29.70	33.41	44.72	62.15	69.13
University Center	7.92	13.06	15.25	14.68	10.63
College	61.72	52.66	37.71	20.80	18.84
Federal Institute	0.66	0.87	2.33	2.37	1.39
Distribution by field of study (%)					
Agriculture and veterinary	3.96	2.19	2.17	3.76	4.92
Social sciences, business and law	29.04	38.76	37.14	30.82	26.71
Natural sciences, mathematics and ICTs	10.56	10.67	9.55	10.85	12.47
Education	20.79	18.70	21.06	23.35	22.67
Engineering, manufaturing and construction	13.53	11.39	11.11	10.76	12.55
Humanities and arts	6.93	2.42	2.72	3.25	5.87
Health and welfare	12.54	13.39	14.31	15.43	13.06
Services	2.64	2.49	1.93	1.78	1.76
Distribution by HEI quality (%)					
IGC=1	11.36	0.06	0.00	0.00	0.00
IGC=2	46.52	31.78	3.74	0.17	0.00
IGC=3	37.73	62.07	78.52	39.13	15.06
IGC=4	4.03	5.62	16.52	53.57	63.62
IGC=5	0.37	0.48	1.21	7.14	21.33

Source: SINAES Tables and Higher Education Census (INEP). Authors' elaboration.

Table 3 presents a summary of the response variables used in the following analysis by CPC_{level}. The quality indexes in t+3 increase with CPC_{level}. The same behavior is noticed among faculty attributes (faculty size, faculty with an MA, faculty with a PhD, and full-time faculty) and offer variables (slots, applications, new students and enrollments). On the other hand, the higher the quality level of the program, the fewer students per faculty member. The dropout rate does not vary by quality level. Finally, the probability of a program closing in the years following an evaluation is higher among the programs that performed the worst.

Table 3: Program response variables by CPC_{level}

	CPC_{level}				
	1	2	3	4	5
Program quality in t+3					
ENADE	1.78	1.98	2.30	2.93	3.60
Infrastructure	2.99	3.28	3.23	3.20	3.38
Teaching and learning	2.81	2.96	3.04	2.99	3.00
Opportunity	2.68	2.81	2.97	3.15	3.41
IDD	1.95	2.23	2.41	2.72	3.10
CPC	2.04	2.34	2.61	3.06	3.50
Program faculty profile in t+3					
Students/Faculty	10.40	10.51	10.27	8.18	6.04
Faculty	28.68	31.15	35.65	50.89	68.21
MA	19.09	23.27	29.78	46.57	63.39
PhD	10.96	11.97	16.82	33.04	47.27
Full-time	21.28	23.84	29.47	47.24	65.47
Program status and flow indicators in t+1, t+2 and t+3					
Slots in t+1	100.21	145.77	172.85	161.59	120.38
Slots in t+2	91.01	140.02	178.82	166.52	126.16
Slots in t+3	89.52	137.91	186.56	171.84	127.40
Applications in t+1	168.00	257.26	398.87	585.18	519.93
Applications in t+2	163.15	247.60	420.20	590.53	526.48
Applications in t+3	179.69	253.99	428.73	607.16	586.73
New students in t+1	50.29	73.13	87.42	84.47	71.54
New students in t+2	42.06	61.92	85.20	83.80	72.00
New students in t+3	37.89	61.22	80.92	79.78	71.73
Total enrollment in t+1	159.35	231.56	284.47	266.85	225.23
Total enrollment in t+2	145.69	217.22	276.61	266.11	228.35
Total enrollment in t+3	135.56	203.25	266.21	261.53	232.58
Dropout in t+1	37.73%	55.29%	48.58%	49.78%	39.05%
Dropout in t+2	37.86%	57.20%	59.32%	57.12%	49.99%
Dropout in t+3	50.59%	60.91%	59.37%	52.63%	40.21%
Activity status in t+1	89.47%	95.26%	97.58%	97.48%	96.86%
Activity status in t+2	81.18%	91.95%	95.81%	96.19%	96.59%
Activity status in t+3	74.53%	88.96%	94.04%	94.67%	95.03%

Source: SINAES Tables and Higher Education Census (INEP). Authors' elaboration.

Notes: Quality indexes are continuous variables and range from 0 to 5. Students per faculty member is the ratio of students to faculty members in the program. Faculty, MA, PhD and Full-time refer to the number of faculty members, the percentage of faculty with a master degree, the percentage of faculty with a PhD, and the percentage of faculty with full-time appointments (dedicating 40 hours or more per week to the program with which they are associated), respectively. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether the programs are still open in the following years.

4 Empirical strategy

The empirical analysis used in this paper is similar to that used in previous studies that have applied an RDD to identify the impact of accountability grades on education issues (Chiang (2009); Rockoff and Turner (2010); Rouse et al. (2013); Chakrabarti (2014); Craig et al. (2015); Woo et al. (2015); Holbein and Ladd (2017); Feng et al. (2018)). In particular, our estimation strategy is quite similar to that of Rockoff and Turner (2010), which explored the heterogeneous effects of different performance levels on school outcomes. In the same way, we explore the discontinuities in CPC levels arising from the continuous grades used to determine the levels in order to compare the performance in subsequent years of undergraduate programs that received different grades. The main assumption behind this strategy is that when comparing programs that fall on either side of the grade cutoff, the assignment of a high or a low level to each program is as good as randomly determined.

We examine the impact of the grades received by each program evaluated in the period 2007-2015 on program performance in the following three years. Because the probability of treatment (i.e., being above a specific level) changes from 0 to 1 at each cutoff, we have a sharp RDD. We estimate the reduced-form regression specification described by equation 2[34].

$$Y_{jt+3} = \alpha + \lambda_L \mathbf{CPC}^L_{jt} + \beta \boldsymbol{f}(\mathbf{P}_{jt}) + \gamma \mathbf{D}_{jt} + \varepsilon_{jt}, \tag{2}$$

where Y_{jt+3} is the variable of interest for program j in year t, $\mathbf{CPC}^L_{jt}$ is a vector of dummies indicating whether a program is above CPC_{level} (L) based on the CPC_{score} that it achieved in $t = 0$ relative to the cutoffs described previously (CPC_{level}=2 when CPC_{score}≥0.945, CPC_{level}=3 when CPC_{score}≥1.945, CPC_{level}=4 when CPC_{score}≥2.945, and $CPC_{level} = 5$ when CPC_{score}≥3.945), $\mathbf{P}_{jt}$ is a vector of continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics), $\mathbf{D}_{jt}$ is a vector of program characteristics control variables (the number of programs within the same field of study[35] in the same institution and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution), and ε_{jt} is an idiosyncratic error term. We add a quartic polynomial in $\mathbf{P}_{jt}$.

[34]Because the assignment of general higher education quality ratings (the IGC_{level}) follows rules similar to those for CPC_{level} assignments, we adapt equation 2 to evaluate the impact of the IGC on the aggregated outcomes of HEIs. These results are presented in Appendix E.

[35]Until 2015, programs were evaluated in groups within the same area and institution, and so these programs could behave differently from programs that are evaluated individually.

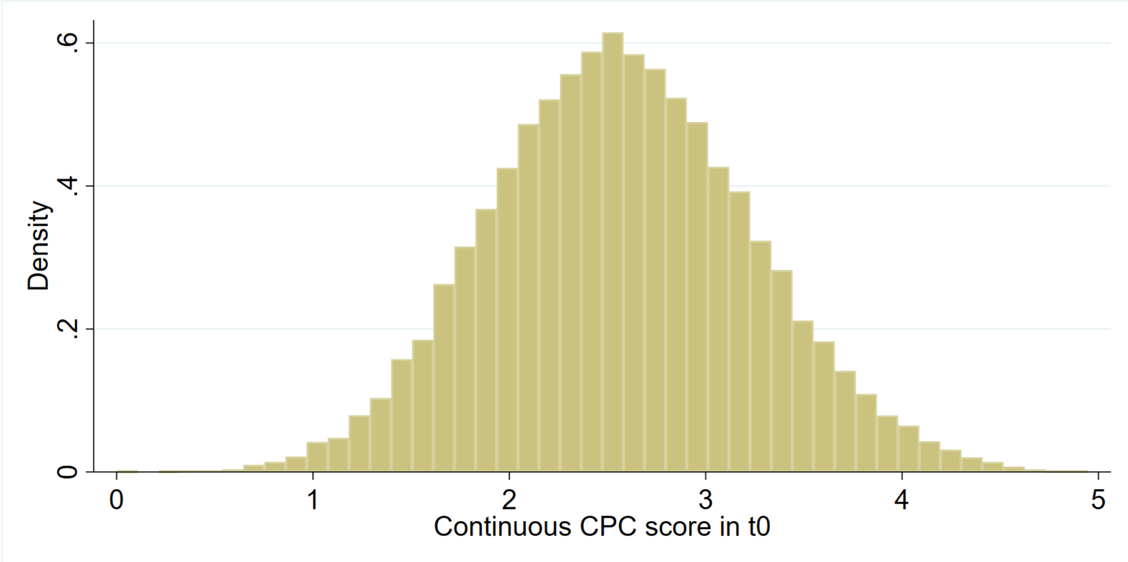
We categorize the variables of interest into three major groups: quality indexes, offer variables, and faculty characteristics{The SINAES measures of student performance (the ENADE and IDD) are not comparable over time, because they are not based on Item Response Theory or similar methodologies that make comparisons over time more credible (OECD (2018)). As our estimates are not based on variation over time, this fact does not affect our results. For the quality indexes, we weight the regressions with the number of graduating students taking the ENADE because those indexes are based on answers given by this group of students on the feedback questionnaire. For the offer variables and faculty characteristics, the regressions are weighted by the total number of students enrolled in the year of evaluation since these variables are based on data for all students in the program (not only final year students).

We conduct a few robustness tests. We "falsify" our estimates by using the same equations but with dependent variables from the previous periods. Robustness requires that predetermined characteristics exhibit no discontinuities at the thresholds that define the CPC levels (Lee and Lemieux (2010); Cattaneo et al. (2019)).

In addition, we examined modifications to $\mathbf{P}_{jt}$, varying the order of the polynomial; i.e., we tested quadratic and quartic polynomials (these results are presented in the appendix C). We also obtained cutoff-specific estimates from local polynomial estimation and robust bias-corrected inference procedures from Cattaneo et al., 2020 (results are presented in appendix D).

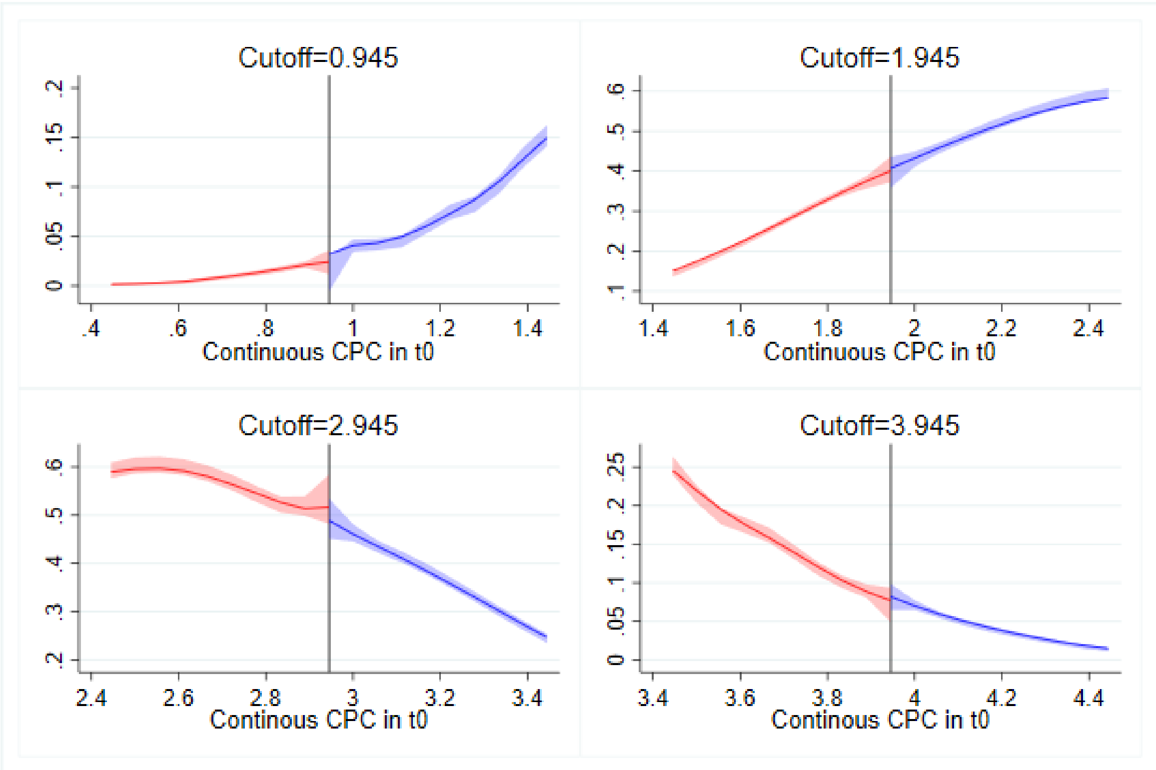
Although the accountability system for higher education in Brazil is supposed to be an exogenous system of evaluation, the manipulation of CPC levels around each CPC_{score} threshold could be a risk for our RDD specification if programs or institutions can perfectly determine the outcomes of their evaluation process. To be sure that programs do not perfectly determine their outcomes, we also test for threshold manipulation by plotting the histogram of the continuous CPC scores and testing the density around each cutoff based on a nonparametric density estimator, applying the method proposed by Cattaneo et al. (2020).

Figure 6 presents a histogram of the CPC_{score} with no signs of manipulation. Figure 7 implements manipulation tests for CPC_{score} at each CPC_{level} threshold. The figure does not suggest that there is manipulation around any of the thresholds. This is a necessary condition for conducting a credible sharp RDD analysis.



Source: SINAES Tables (INEP). Authors' elaboration.

Figure 6: Histogram of the CPC_{score}



Source: SINAES Tables (INEP). Authors' elaboration.

Figure 7: Density around thresholds of level assignments

Notes: The graphs plot the test of difference of densities around the thresholds according to the method of McCrary (2008).

5 Results

5.1 Program quality

First, we present a graphical analysis of our estimation strategy. We plot the linear results of a locally weighted Fan regression of the quality indexes against the CPC_{score} of each undergraduate program according to the method proposed by Fan et al. (1995). We estimate the local regressions separately for each group of programs that received the same CPC_{level}, including a quartic polynomial as a control. Jumps at the thresholds indicate that the outcomes are sensitive to CPC_{level} assignment.

Figure 8 presents graphs of the quality indicators and composite indexes measured 3 years after the evaluation, i.e., in the next evaluation cycle. Because these indicators feed into the final CPC, we expect a positive relationship between each indicator and the previous CPC_{score}. We identify a jump around threshold 1.945 (which separates CPC levels 2 and 3) for the ENADE and IDD scores, Infrastructure, Teaching and Learning, Opportunity, and CPC indexes. In these cases, the programs next to and below the threshold achieved higher outcomes in the next SINAES evaluation. We also identify a potential discontinuity around threshold 3.945 in the Opportunity index, suggesting that students identify more professional and academic opportunities in programs that received the highest score ($CPC_{level} = 5$). Around the other thresholds, jumps are less visible, indicating potentially lower impacts on the incentives for undergraduate programs related to the next evaluation among programs with a CPC level of 4 or 5.

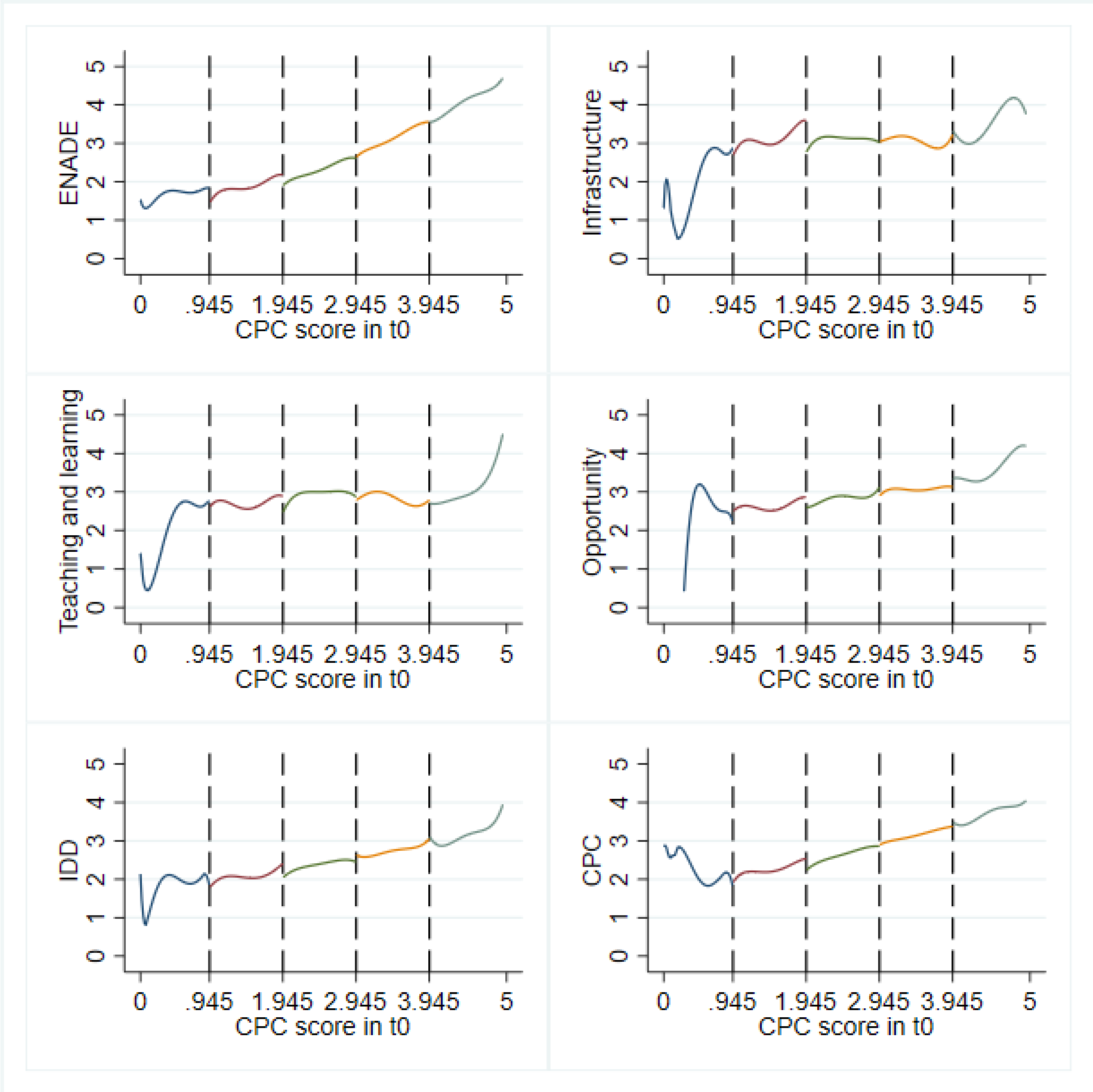


Figure 8: Program quality in t+3, by initial CPC_{score}

Notes: Quality indexes are continuous variables and range from 0 to 5. The graphs plot the linear results of a locally weighted Fan regression of the quality indexes against the CPC_{score} of each undergraduate program (Fan et al. (1995)). The local regressions are estimated separately for each group of programs that received the same CPC_{level}, and include a quartic polynomial as a control. Jumps at the thresholds indicate that the outcomes are sensitive to CPC_{level} assignment.

Table 4 shows the sharp RDD results with a quartic polynomial function in CPC_{score}. The results confirm the visible differences in figure 8. Being classified into $CPC_{level} = 2$ in year t increases the corresponding quality indexes in the next evaluation cycle (t+3). Thus, we obtain negative and statistically significant estimates at the CPC_{score} threshold of 1.945. The most likely explanation is that programs below the Ministry of Education recognition level overreact to the threat of punishment. Falling below level 3 triggers additional evaluations and supervisory processes that require improvements in several outcomes under penalty of suspension or closure should the program fail to fulfill its commitments. Level 2 programs have advantages around threshold 1.945 in terms of the indicators measured in t+3, such as ENADE (scores are higher by 0.195 points), Infrastructure (0.362 points higher), Teaching and learning (0.174 points higher), Opportunity (0.182 points higher), IDD (0.148 points higher) and the composite index CPC_{score} (0.158 points higher).

Alternatively, these results may reflect the program administrators' fears not only of the regulatory agency but also of bad "propaganda" that could reduce student demand for slots in the program.

At other thresholds, the evidence is inconsistent. In particular, the results for the Opportunity index at $CPC_{score} = 3.945$ are insignificant even after accounting for the potential discontinuity for this variable at this threshold visible in figure 8. Only at $CPC_{score} = 3.945$ do we see statistically significant differences for the ENADE and IDD, equal to 0.18 and 0.099, respectively, in favor of programs just below the threshold (for private institutions only). Nonetheless, these results are not robust to the robustness tests performed – see tables 5 and C.1 in the appendix.

We also run separate regressions by type of administration – public or private – to assess the potential heterogeneity in the impacts. We find stronger impacts for private HEIs, with differences still concentrated on $CPC_{score} = 1.945$ for both groups. Such differences in the responses to accountability between private and public institutions suggest that program administrators and faculty react differently to different incentive schemes, a result that is similar to the findings of Camargo et al. (2018) for secondary education in Brazil. Because teachers and managers enjoy job security in Brazilian public institutions and do not receive bonuses or salary increases for good performance, they do not face the same market incentives as their peers in private colleges. Furthermore, along with the consolidation of SINAES, the government has been expanding public institutions despite their performance results, which also reduces the incentives for public HEI managers to improve quality. For example, while the private programs classified into $CPC_{level} = 1$ or $CPC_{level} = 2$ reduced by 5% their slots in three years, the public programs classified into the same levels increased their slots by 3%, according to Higher Education Census (INEP).

Table 4: The impact of accountability on program quality

	(1)	(2)	(3)	(4)	(5)	(6)
	ENADE	Infrastructure	Teaching and learning	Opportunity	IDD	CPC
All sample						
CPC_{level}=2	-0.173	0.036	0.110	-0.043	-0.122	0.027
	(0.164)	(0.218)	(0.201)	(0.274)	(0.261)	(0.101)
CPC_{level}=3	-0.195***	-0.362***	-0.174***	-0.182***	-0.148***	-0.158***
	(0.064)	(0.084)	(0.045)	(0.054)	(0.038)	(0.018)
CPC_{level}=4	-0.017	0.042	0.006	0.082	0.032	0.011
	(0.026)	(0.061)	(0.074)	(0.055)	(0.022)	(0.015)
CPC_{level}=5	-0.180*	0.059	0.086	0.099	-0.099**	-0.039
	(0.094)	(0.110)	(0.089)	(0.091)	(0.043)	(0.050)
n	34,405	35,052	35,052	29,640	33,637	33,437
Programs in private institutions						
CPC_{level}=2	-0.285**	0.068	0.187	0.197	-0.181	-0.016
	(0.124)	(0.225)	(0.215)	(0.289)	(0.235)	(0.143)
CPC_{level}=3	-0.189***	-0.354***	-0.167***	-0.195***	-0.161***	-0.166***
	(0.067)	(0.076)	(0.049)	(0.065)	(0.045)	(0.023)
CPC_{level}=4	-0.006	0.030	0.010	0.138	0.051**	0.025
	(0.032)	(0.085)	(0.106)	(0.097)	(0.023)	(0.019)
CPC_{level}=5	-0.162**	0.133	0.128	0.215*	-0.254***	-0.077
	(0.065)	(0.125)	(0.122)	(0.124)	(0.071)	(0.058)
n	24,231	24,780	24,780	21,373	23,586	23,603
Programs in public institutions						
CPC_{level}=2	-0.058	-0.195	-0.202	-0.636**	0.032	0.032
	(0.378)	(0.447)	(0.396)	(0.240)	(0.403)	(0.219)
CPC_{level}=3	-0.140	-0.245**	-0.247***	-0.100*	-0.103*	-0.085*
	(0.118)	(0.101)	(0.055)	(0.055)	(0.056)	(0.042)
CPC_{level}=4	-0.010	0.087*	0.001	-0.014	-0.021	-0.004
	(0.034)	(0.043)	(0.021)	(0.046)	(0.036)	(0.014)
CPC_{level}=5	-0.229**	-0.025	0.077	0.016	-0.050	-0.062
	(0.093)	(0.114)	(0.085)	(0.119)	(0.085)	(0.048)
n	10,174	10,272	10,272	8,267	10,051	9,834

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Quality indexes are continuous measures and range from 0 to 5. The regressions are weighted by the number of senior students taking the ENADE exam. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Proper analysis of causal effects requires that the previous results be subjected to robustness tests. One type of robustness test is a falsification test. We conduct falsification tests by regressing the quality outcomes from the pre-treatment period on the same covariates as in our models in table 4. To claim causality, the estimates must not be in the same direction as the main estimates or have similar magnitudes. Table 5 shows that there are no pre-treatment jumps around threshold $CPC_{score} = 1.945$ except for the ENADE indicator, and even in this case, the direction is opposite that of the estimates in table 4. Thus, our results suggest that the estimates in table 4 are plausibly causal. Table C.1 in the appendix also reports estimates with different polynomial functions (cubic and quadratic) as controls. Only the differences at $CPC_{score} = 3.945$ ($CPC_{level} = 5$) are no longer statistically significant. Other estimates with different polynomials remain similar around threshold $CPC_{score} = 1.945$. We perform further robustness tests in section 5.4, wherein we estimate local regressions within specific bandwidths to confirm our results.

Finally, to exclude the possibility that the exams were manipulated, i.e., that the accountability system was gamed, we conduct an additional test in which the ratio of the number of students taking the ENADE to the total program enrollment is used as the dependent variable. Despite the fact that institutions might have incentives to manipulate the number of students participating in the ENADE, we do not find evidence of manipulation, as shown in table B.1 in the appendix.

Table 5: The impact of accountability scores on pre-treatment program quality

	(1)	(2)	(3)	(4)	(5)
	ENADE	Infrastructure	Teaching and learning	IDD	CPC
CPC_{level}=2	0.267	-0.370	-0.117	0.238	0.154
	(0.170)	(0.405)	(0.181)	(0.286)	(0.104)
CPC_{level}=3	0.093**	-0.071	-0.036	0.017	-0.002
	(0.042)	(0.043)	(0.038)	(0.083)	(0.057)
CPC_{level}=4	0.053	-0.008	-0.041	0.100	0.017
	(0.059)	(0.026)	(0.038)	(0.076)	(0.045)
CPC_{level}=5	-0.162	0.046	0.041	-0.253*	-0.147
	(0.100)	(0.096)	(0.110)	(0.147)	(0.088)
n	22,852	25,063	25,063	21,619	22,010

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Quality indexes are continuous measures ranging from 0 to 5 and refer to the pre-treatment measurement (i.e in t-3). The regressions are weighted by the number of senior students taking the ENADE exam. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

5.2 Program status and flow indicators

Figure 9 displays graphs of slots, applications and new students (measured as the sum over three years, i.e., one evaluation cycle). The figure also shows total enrollments, the dropout rate and activity status (i.e., whether the programs are still open), which refer to the last year of the following evaluation cycle. Slots, applications, new students and enrollments are in logarithms, while the dropout rate and activity status are measured in percentages. In general, the supply side indicators increase until $CPC_{score} = 2.945$. For slots, new students, enrollment and the dropout rate, the indicators increase up to $CPC_{score} = 2.945$, and for applications, up to $CPC_{score} = 3.945$). There are small but visible jumps around thresholds $CPC_{score} = 0.945$ and $CPC_{score} = 1.945$ for slots, applications, new students and enrollment. In these cases, higher CPC_{level} assignments are associated with a greater number of slots, applications, and new students and higher enrollment. At $CPC_{score} = 3.945$ and $CPC_{score} = 4.945$, higher CPC_{level} assignments lead to a decrease in program slots, applications, new students, enrollment and the dropout rate. Nevertheless, regressions around each of the thresholds show no statistical significance, which render these results less reliable.

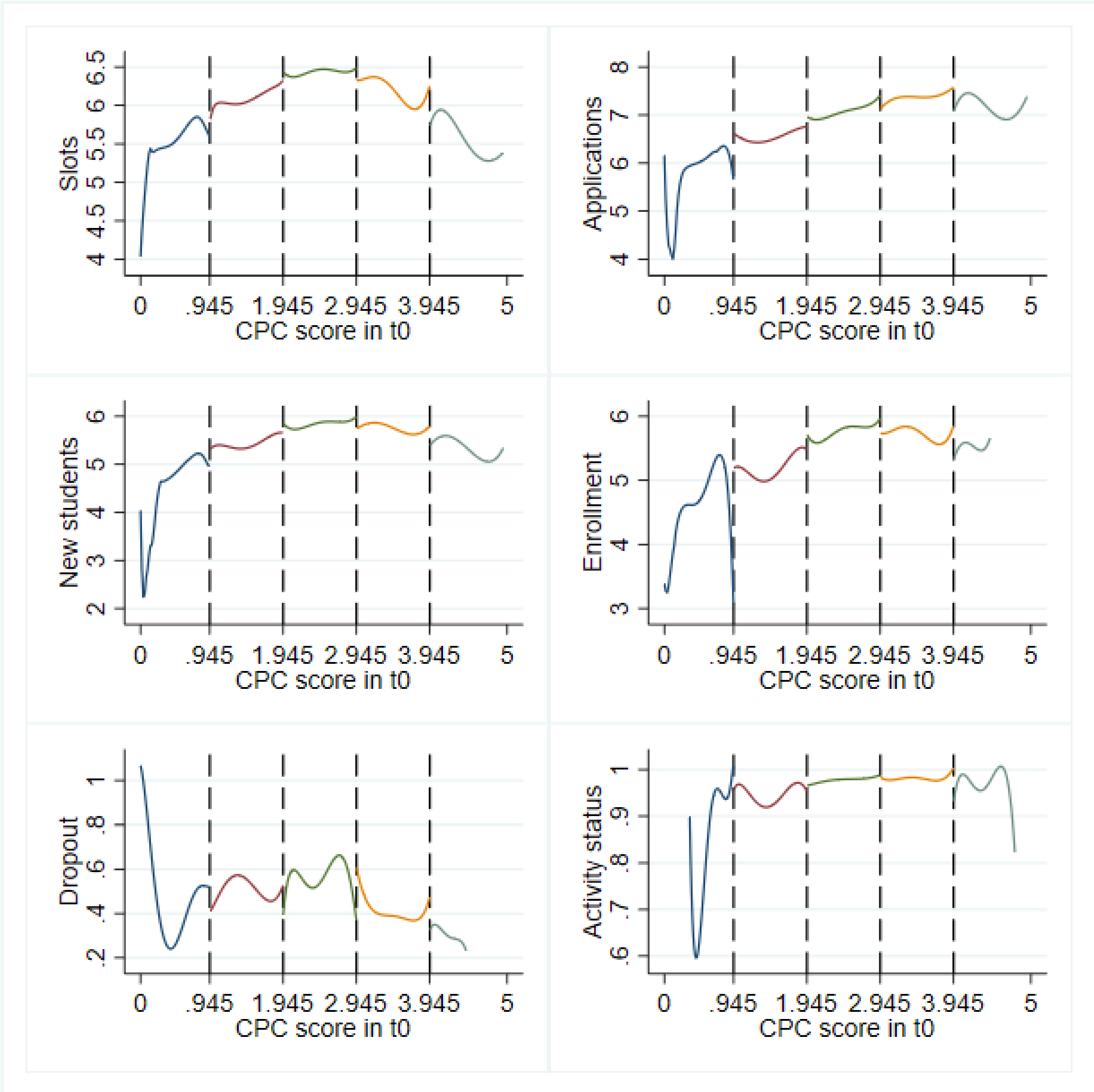


Figure 9: Program status and flow indicators in t+3, by initial CPC_{score}

Notes: Slots, applications and new students are the sum of the variables for the period from t+1 to t+3. Enrollments, dropout and activity status are measured in t+3. Outcome variables such as slots, applications, new students and enrollments are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. The regressions are weighted by the total number of students enrolled in the year of evaluation. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether courses are still open in t+3. The graphs plot the linear results of a locally weighted Fan regression of the quality indexes against the CPC_{score} of each undergraduate program (Fan et al. (1995)). The local regressions are estimated separately for each group of programs that received the same CPC_{level}, and include a quartic polynomial as a control. Jumps at the thresholds indicate that the outcomes are sensitive to CPC_{level} assignment.

Table 6 partially confirms the results presented in figure 9. At threshold

$CPC_{score} = 1.945$, the number of slots increases by 13.4% and the number of applications and new students increase by 12.5% and 10.3%, respectively, at recognized institutions. As expected, these results are driven by private institutions, while we do not find evidence that public institutions increase the number of program openings, new students or applications because of the inflexibility of state-led institutions.

The falsification test presented in table 7 – in which the pre-treatment outcomes are regressed on the same covariates as in table 6 – shows that the statistically significant results around threshold $CPC_{score} = 1.945$ in table 6 are not statistically different from zero when using pre-treatment outcomes. This result reinforces the plausibility that the estimates in table 6 are causal.

Finally, the results in table 6 suggest that legal recognition by the federal regulator increases the number of program slots, applications and new students. Demand-related explanations are the most likely, as applications increase with recognition. This recognition effect may also reflect positive reinforcement, as recognition increases student access to scholarships and loans.

Table 6: The impact of accountability on program status and flow indicators

	(1)	(2)	(3)	(4)	(5)	(6)
	Slots	Applications	New students	Total enrollment	Dropout	Activity
All sample						
CPC_{level}	0.106	0.124	0.173	0.122	-0.060	0.010
	(0.133)	(0.240)	(0.174)	(0.149)	(0.131)	(0.035)
CPC_{level}=3	0.134***	0.125***	0.103**	0.049	0.021	0.004
	(0.041)	(0.041)	(0.039)	(0.043)	(0.072)	(0.005)
CPC_{level}=4	-0.008	-0.024	-0.025	-0.044*	0.027	0.006*
	(0.032)	(0.039)	(0.028)	(0.023)	(0.086)	(0.003)
CPC_{level}=5	0.014	-0.053	-0.016	-0.013	-0.012	0.018
	(0.043)	(0.086)	(0.053)	(0.058)	(0.083)	(0.013)
n	37,861	37,484	37,277	38,819	37,115	32,978
Programs in private institutions						
CPC_{level}=2	0.078	0.133	0.325*	0.230	-0.085	0.005
	(0.129)	(0.267)	(0.160)	(0.145)	(0.211)	(0.043)
CPC_{level}=3	0.182***	0.167***	0.139***	0.078	0.030	0.004
	(0.050)	(0.052)	(0.046)	(0.050)	(0.083)	(0.006)
CPC_{level}=4	-0.005	-0.050	-0.017	-0.045	0.055	0.003
	(0.044)	(0.040)	(0.043)	(0.035)	(0.120)	(0.004)
CPC_{level}=5	-0.084	-0.159*	-0.083	-0.052	0.105	0.019
	(0.086)	(0.080)	(0.092)	(0.087)	(0.065)	(0.018)
n	27,216	26,892	26,675	27,727	26,352	24,091
Programs in public institutions						
CPC_{level}=2	0.229	0.082	-0.003	0.019	-0.195*	0.000
	(0.188)	(0.273)	(0.240)	(0.162)	(0.111)	(0.055)
CPC_{level}=3	-0.066	0.027	-0.056	-0.069	-0.029	0.003
	(0.072)	(0.050)	(0.077)	(0.052)	(0.052)	(0.010)
CPC_{level}=4	-0.042	0.004	-0.033	-0.015	-0.022	0.009
	(0.040)	(0.079)	(0.038)	(0.027)	(0.029)	(0.006)
CPC_{level}=5	0.102*	0.067	0.055	-0.008	-0.114	0.013
	(0.055)	(0.082)	(0.077)	(0.075)	(0.121)	(0.014)
n	10,645	10,592	10,602	11,092	10,763	8,887

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Columns (1)-(3) refer to the sum of new slots, applications and new students over the period t+1 to t+3. Columns (4)-(6) refer to enrollment, the dropout rate and the activity status in t+3. Outcome variables such as new slots, applications, new students and enrollment are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. The regressions are weighted by the total number of students enrolled in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Table 7: The impact of accountability scores on pre-treatment program status and flow indicators

	(7)	(8)	(9)	(10)	(11)	(12)
	Slots	Applications	New students	Total enrollment	Dropout	Activity
CPC_{level}=2	0.056	0.075	0.144	0.094	-0.063**	0.006
	(0.134)	(0.237)	(0.170)	(0.158)	(0.030)	(0.016)
CPC_{level}=3	0.039	0.014	0.047	0.042	0.014	0.001
	(0.041)	(0.041)	(0.044)	(0.038)	(0.009)	(0.002)
CPC_{level}=4	0.013	-0.033	-0.001	-0.040	0.002	-0.000
	(0.025)	(0.031)	(0.026)	(0.025)	(0.008)	(0.001)
CPC_{level}=5	-0.006	-0.118	0.017	0.045	-0.008	0.002
	(0.050)	(0.070)	(0.074)	(0.066)	(0.025)	(0.002)
n	37,845	37,517	37,616	39,425	39,425	17,798

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Offer variables refer to measures in t0 (pre-treatment measurement). Outcome variables such as new slots, applications, new students and enrollment are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. The regressions are weighted by the total number of students enrolled in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

5.3 Program faculty

Faculty profiles can also change in response to evaluation scores. Figure 10 presents graphs for faculty profiles 3 years after the evaluation. For the number of students per faculty member, we identify an almost flat relation with CPC, with no clear jumps around thresholds. Except for faculty with an MA, which is negatively related to CPC_{score}, the other faculty indicators are positively related to CPC. In general, we do not find straightforward jumps around the thresholds. There are small visible jumps around $CPC_{score} = 3.945$, with more faculty with a PhD and fewer with an MA in programs evaluated at CPC level 5.

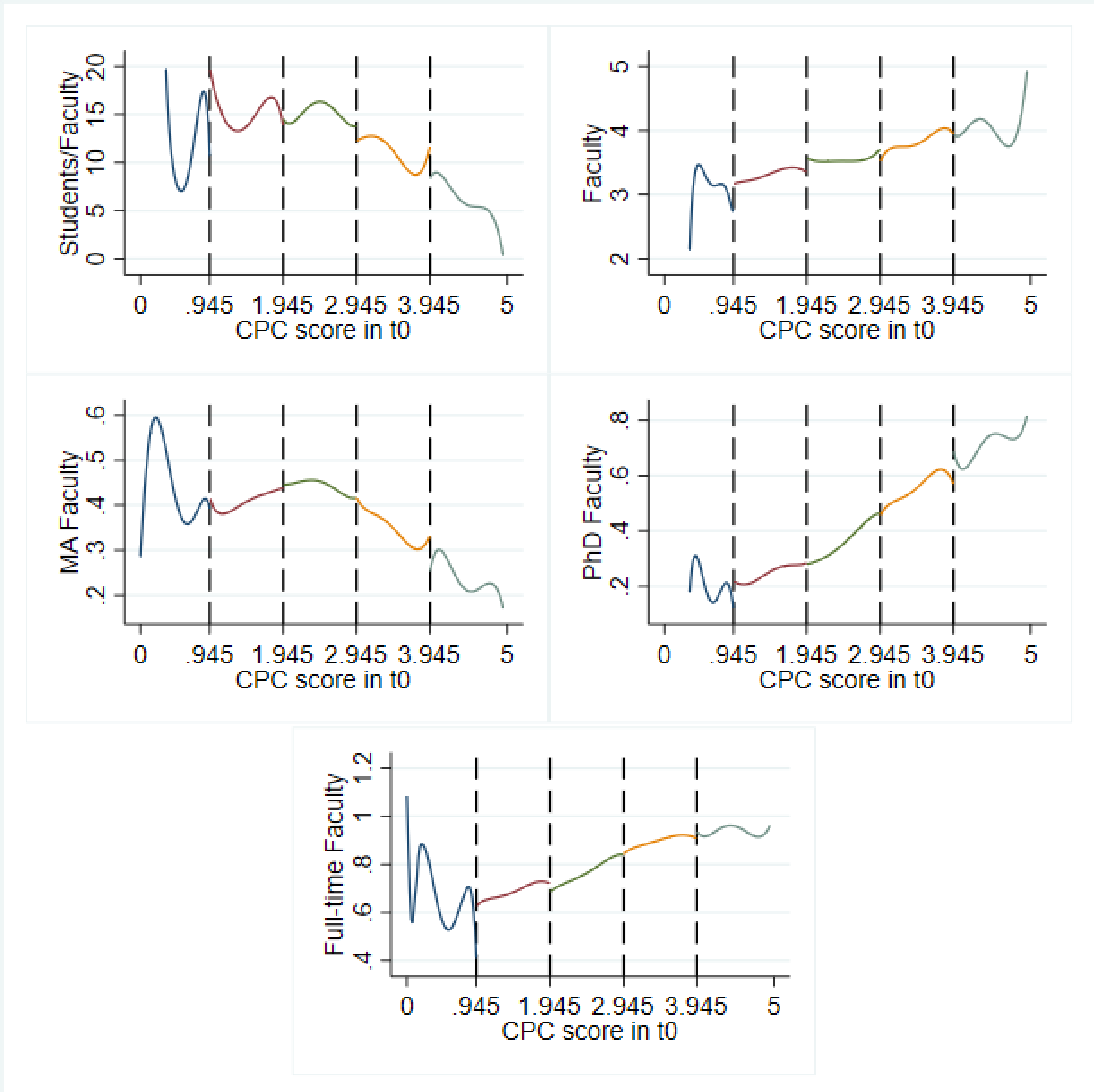


Figure 10: Program faculty profile in t+3, by initial CPC_{score}

Notes:Faculty variables are measured in t+3. Students/faculty is the ratio between the number of students and the number of professors associated with the program. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. The graphs plot the linear results of a locally weighted Fan regression of the quality indexes against the CPC_{score} of each undergraduate program (Fan et al. (1995)). The local regressions are estimated separately for each group of programs that received the same CPC_{level}, and include a quartic polynomial as a control. Jumps at the thresholds indicate that the outcomes are sensitive to CPC_{level} assignment.

The results in table 8 indicate significant impacts around threshold $CPC_{score} = 1.945$ for PhD and full-time faculty, as shown in columns (4) and (5). In contrast, we do not find any significant impact on the number of students per faculty member, faculty size or the percentage of faculty with an MA degree.

Our results suggest that programs that fall below the recognition threshold, i.e., $CPC_{score} = 1.945$, overreact to their evaluation by hiring more PhD and full-time faculty. Programs under the supervision of the regulatory authority, i.e., with a CPC_{score} just below 1.945, increase the percentage of professors with a PhD by 2.7 percentage points and the percentage with full-time contracts by 3.5 percentage points by the next evaluation.

In addition, table 8 suggests that only private institutions react to low scores. Around CPC_{score}=1.945, we also find that private institutions below the threshold increase the number of faculty members by 9.6%. In general, public institutions do not react to evaluations by changing faculty inputs, as the hiring process depends on public funding and such positions include job security, which prevents administrators from adjusting these inputs.

To evaluate the robustness of these findings, we also estimate the same regression over variables measured in the last three years in columns 1 through 6 of table 9 . Around $CPC_{score} = 1.945$, we do not find any significant estimates, which leads us to conclude that there are no previous discontinuities around that threshold.

These results confirm those we find for other outcomes. Undergraduate programs overreact to bad evaluations. Perhaps because they have imperfect control over their outcomes, program administrators adopt several measures to improve their indicators and be re-classified in the next evaluation cycle. We do not find evidence of a "score effect", wherein better-evaluated programs invest in the improvement of their indicators to maintain and, whenever possible, increase their scores.

Table 8: The impact of accountability on program faculty profile

	Students/ faculty	Faculty	MA	PhD	Full-time
All sample					
CPC_{level}=2	6.042	-0.192	-0.033	0.000	-0.010
	(6.305)	(0.138)	(0.034)	(0.047)	(0.050)
CPC_{level}=3	0.478	0.053	0.004	-0.027***	-0.035***
	(0.477)	(0.036)	(0.008)	(0.006)	(0.005)
CPC_{level}=4	-0.397	0.012	0.005	-0.014*	0.003
	(0.506)	(0.034)	(0.006)	(0.008)	(0.008)
CPC_{level}=5	0.508	0.030	-0.011	0.019	0.012
	(0.532)	(0.064)	(0.018)	(0.017)	(0.009)
n	25,043	25,117	34,804	34,804	34,804
Programs in private institutions					
CPC_{level}=2	-0.419	-0.015	-0.067*	-0.027	-0.011
	(3.081)	(0.125)	(0.033)	(0.057)	(0.060)
CPC_{level}=3	0.426	0.096***	0.005	-0.038***	-0.042***
	(0.630)	(0.034)	(0.007)	(0.008)	(0.007)
CPC_{level}=4	-0.128	0.021	0.001	-0.011	0.006
	(0.653)	(0.036)	(0.006)	(0.007)	(0.009)
CPC_{level}=5	0.406	0.054	-0.025*	0.027*	0.024
	(1.092)	(0.054)	(0.014)	(0.014)	(0.017)
n	16,710	16,749	24,638	24,638	24,638
Programs in public institutions					
CPC_{level}=2	23.552	-0.543**	0.035	0.027	-0.061
	(18.551)	(0.220)	(0.043)	(0.029)	(0.060)
CPC_{level}=3	-0.701	-0.109	0.011	0.015	-0.006
	(1.136)	(0.073)	(0.020)	(0.009)	(0.008)
CPC_{level}=4	-1.165	0.012	0.002	-0.006	0.003
	(0.924)	(0.055)	(0.006)	(0.008)	(0.004)
CPC_{level}=5	0.524	0.009	0.013	-0.002	-0.002
	(0.445)	(0.057)	(0.012)	(0.014)	(0.004)
n	8,333	8,368	10,166	10,166	10,166

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Faculty variables are measured in t+3. Students/faculty is the ratio between the number of students and the number of professors associated with the program. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. Regressions are weighted by the number of enrollments in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Table 9: The impact of accountability scores on pre-treatment program faculty profile

	(1)	(2)	(3)	(4)	(5)
	Students/ faculty	Faculty	MA	PhD	Full-time
CPC_{level}=2	3.097	0.153	-0.018	-0.004	0.039
	(2.247)	(0.228)	(0.067)	(0.024)	(0.074)
CPC_{level}=3	0.877	0.043	-0.011	0.012	0.000
	(0.589)	(0.041)	(0.010)	(0.008)	(0.013)
CPC_{level}=4	-0.256	-0.002	-0.006	-0.015***	-0.009
	(0.475)	(0.030)	(0.004)	(0.004)	(0.009)
CPC_{level}=5	4.012	-0.090	-0.047***	-0.024**	-0.023*
	(2.858)	(0.063)	(0.015)	(0.012)	(0.013)
n	19,740	19,858	23,229	25,373	25,373

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Faculty variables refer to the pre-treatement measurement (i.e in t-3). Students/faculty is the ratio between the number of students and the number of professors associated with the program. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. Regressions are weighted by the number of enrollments in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

5.4 Local regressions

This section presents robustness tests of the previous estimates around the threshold $CPC_{score} = 1.945$. We estimate local linear regressions of equation 2 with bandwidths $h \leq 1$, $h \leq 0.5$ and $h \leq 0.25$ over the same set of outcomes analyzed in the previous sections. Local regressions within small enough bandwidths reduce bias from selection on unobservables. Table 10 shows estimates for the quality outputs, which include the following indicators and indexes: 1) ENADE, 2) Infrastructure, 3) Teaching and learning, 4) Opportunity, 5) IDD and 6) CPC_{score}. As we reduce the bandwidth, we notice that the magnitudes, signs, and statistical significance of the results remain similar. In fact, compared to the previous results, the estimates from the regression with the smallest bandwidth, $h \leq 0.25$, seem to be greater in magnitude. Thus, our results corroborate and reinforce our previous conclusions. Undergraduate programs evaluated below the recognition level overreact in order to improve their performance on the ENADE and IDD indexes, as well as to improve their program infrastructure, teaching and learning, and opportunity indicators and attain recognition during the next evaluation cycle. We do not report the results for the other thresholds, as none of them are statistically significant here or in the previous sections.

Table 10: Local regressions of the impact of accountability on program quality

	(1)	(2)	(3)	(4)	(5)	(6)
	ENADE	Infrastructure	Teaching and learning	Opportunity	IDD	CPC
Unlimited distance to cutoff						
CPC_{level}=3	-0.189***	-0.363***	-0.177***	-0.143***	-0.180***	-0.157***
	(0.064)	(0.085)	(0.045)	(0.037)	(0.059)	(0.019)
n	34,405	35,052	35,052	33,637	29,640	33,437
Distance≤1						
CPC_{level}=3	-0.181**	-0.439***	-0.267***	-0.199***	-0.268***	-0.181***
	(0.087)	(0.096)	(0.042)	(0.071)	(0.065)	(0.047)
n	23,524	24,015	24,015	22,894	19,911	22,731
Distance≤0.5						
CPC_{level}=3	-0.190*	-0.427***	-0.357***	-0.265*	-0.207**	-0.254***
	(0.099)	(0.089)	(0.083)	(0.138)	(0.083)	(0.089)
n	12,153	12,416	12,416	11,742	9,849	11,649
Distance≤0.25						
CPC_{level}=3	-0.154**	-0.472***	-0.353***	-0.173**	-0.192**	-0.224***
	(0.065)	(0.064)	(0.084)	(0.083)	(0.085)	(0.054)
n	6,063	6,215	6,215	5,834	4,821	5,786

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Quality indicators are continuous and range from 0 to 5. The regressions are weighted by the number of senior students taking the ENADE exam. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Table 11 presents estimates of the level changes that occur at the threshold $CPC_{score} = 1.945$ in the number of slots, applications, and new students and in enrollment and the dropout rate. Within the smallest bandwidth of $h \leq 0.25$, estimates are very similar to the parametric estimates, except that the standard errors are larger and only the number of new students remains statistically significant, though only at the 10% level. Nevertheless, altogether, the results confirm the parametric estimates and suggest that the obtaining recognition, i.e., $CPC_{level} = 3$ or higher, results in an increase in the number of slots, applications and new students.

Table 11: Local regressions of the impact of accountability on program status and flow indicators

	(1)	(2)	(3)	(4)	(5)	(6)
	Slots	Applications	New students	Total enrollment	Dropout	Closure situation
Unlimited distance to cutoff						
CPC_{level}=3	0.132***	0.124***	0.101**	0.047	0.022	0.003
	(0.041)	(0.042)	(0.040)	(0.043)	(0.070)	(0.005)
n	37,861	37,484	37,277	38,819	37,115	32,978
Distance≤1						
CPC_{level}=3	0.104**	0.126**	0.086*	0.021	0.029	0.015*
	(0.048)	(0.046)	(0.042)	(0.046)	(0.092)	(0.008)
n	26,266	25,951	25,765	26,917	25,613	22,455
Distance≤0.5						
CPC_{level}=3	0.096	0.096*	0.102	0.084	-0.027	-0.001
	(0.066)	(0.054)	(0.076)	(0.070)	(0.065)	(0.008)
n	13,825	13,628	13,518	14,194	13,396	11,357
Distance≤0.25						
CPC_{level}=3	0.106	0.126*	0.106	0.077	-0.023	-0.004
	(0.083)	(0.070)	(0.086)	(0.083)	(0.061)	(0.012)
n	6,975	6,882	6,800	7,159	6,733	5,617

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Columns (1)-(3) refer to the sum of new slots, applications and new students over the period t+1 to t+3. Columns (4)-(6) refer to enrollment, the dropout rate and the activity status in t+3. Outcome variables such as new slots, applications, new students and enrollment are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. The regressions are weighted by the total number of students enrolled in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Finally, table 12 presents estimates of the level changes in the number of students per faculty member, the number of faculty members, the number of faculty members with an MA degree, the number with a PhD, and the number that are full-time at threshold $CPC_{score} = 1.945$. Within the smallest bandwidth, $h \leq 0.25$, the estimates are very similar to those from the parametric regressions, but the standard errors are larger and none of the estimates are statistically significant.

Table 12: Local regressions of the impact of accountability on program faculty profile

	(1)	(2)	(3)	(4)	(5)
	Students/ faculty	Faculty	MA	PhD	Full-time
Unlimited distance to cutoff					
CPC=3	0.390	0.055	0.005	-0.027***	-0.035***
	(0.481)	(0.036)	(0.008)	(0.005)	(0.005)
n	25,043	25,117	34,804	34,804	34,804
Distance≤1					
CPC_{level}=3	-0.827	0.096	0.003	-0.018	-0.027***
	(0.957)	(0.068)	(0.009)	(0.013)	(0.008)
n	16,922	16,972	23,833	23,833	23,833
Distance≤0.5					
CPC_{level}=3	-0.013	0.128	0.006	-0.014	-0.019
	(0.918)	(0.098)	(0.011)	(0.018)	(0.012)
n	8,711	8,741	12,311	12,311	12,311
Distance≤0.25					
CPC_{level}=3	0.146	0.115	0.015	-0.020	-0.017
	(1.429)	(0.120)	(0.016)	(0.023)	(0.015)
n	4,345	4,361	6,168	6,168	6,168

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Faculty variables are measured in t+3. Students/faculty is the ratio between the number of students and the number of professors associated with the program. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. Regressions are weighted by the number of enrollments in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Local estimates corroborate the main parametric estimates obtained in the previous sections. Results for the quality outcomes, as measured by the ENADE, IDD, CPC, Infrastructure, Teaching and learning, and Opportunity indexes, seem to be robust to falsification, specification and bandwidth tests. Outcomes related to offers and faculty seem to be robust to falsification and specification tests, but the standard errors are quite large within small bandwidths and we cannot reject the null hypothesis of no effect, although the magnitudes and sign remain stable.

5.5 Heterogeneous effects by field of study

In the previous estimations, we examined the heterogeneity in the effects by the type of administration. In addition, we also expect to find heterogeneous effects by field of study. To this extent, we estimate equation 2 for each subgroup of programs within each field of study. In this section, we mainly explore the results for the quality (measured by the CPC score in t+3) and offer (measured by the number of new slots in the three years following each evaluation)[36] outcomes.

The table 13 reports the impact of accountability on CPC scores in t+3 by area. The results confirm the pattern found in the previous estimations: the impact mainly occurs around the cutoff for $CPC_{level} = 3$ and for programs in private institutions. Additionally, at the threshold $CPC_{score} = 1.945$, the CPC index increases for programs in Social sciences, business and law, Health, and Education that are just under the cutoff in t+3 relative to the CPC index for programs in the same areas that are just above the same cutoff. This result is found only in regressions in which the sample is restricted to programs in private institutions.

Table 14 displays the impact of accountability on the number of new slots in the following three years after each evaluation by area. In the estimations over the full sample, the impact around threshold $CPC_{score} = 1.945$ is positive and programs that received a $CPC_{level} = 3$ (i.e., that are above this cutoff) increase the number of slots in Engineering and related fields, the Social sciences, business and law, and Health by 13.2% (p-value<0.1), 16.0% (p-value<0.05) and 20.1% (p-value<0.05), respectively. Again, similar results are observed for the estimations by area over the sample of programs in private institutions.

In both tables, the results for programs within the Sciences, math and computation and Other areas do not demonstrate a consistent pattern around the thresholds. Specifically, there is no significant impact of accountability on the number of new slots for theses programs in the estimations that use the full sample or those restricted to private institutions (see table 14).

[36]For simplicity, we do not report regression results for all variables by area, but these results can be obtained upon request to the authors.

Table 13: The impact of accountability on CPC score in t+3, by field of study

	(1)	(2)	(3)	(4)	(5)	(6)
	Engineering and related	Social sciences, business and law	Health	Education	Sciences, math and computation	Others
All sample						
CPC=2	0.177	-0.038	0.881**	0.043	-0.411*	-0.540
	(0.392)	(0.263)	(0.392)	(0.265)	(0.210)	(0.402)
CPC=3	-0.055	-0.199***	-0.156***	-0.165***	-0.012	0.000
	(0.034)	(0.032)	(0.051)	(0.039)	(0.038)	(0.076)
CPC=4	-0.003	0.017	0.025	0.011	0.036	-0.083
	(0.037)	(0.021)	(0.042)	(0.022)	(0.036)	(0.051)
CPC=5	0.045	-0.067	0.084	-0.117**	-0.186**	0.155
	(0.110)	(0.072)	(0.084)	(0.049)	(0.084)	(0.174)
n	3,953	11,766	3,833	7,878	3,946	2,052
Programs in private institutions						
CPC=2	0.368	0.003	0.374	-0.196	-0.777***	-0.928**
	(0.361)	(0.282)	(0.426)	(0.576)	(0.262)	(0.392)
CPC=3	-0.035	-0.192***	-0.125**	-0.215**	0.010	-0.104
	(0.041)	(0.041)	(0.059)	(0.101)	(0.053)	(0.119)
CPC=4	0.014	0.025	0.055	-0.011	-0.024	0.078
	(0.038)	(0.021)	(0.048)	(0.050)	(0.043)	(0.074)
CPC=5	-0.106	-0.074	0.254**	-0.351**	-0.118	0.237
	(0.243)	(0.078)	(0.096)	(0.146)	(0.140)	(0.172)
n	2,552	9,966	3,049	4,203	2,651	1,179
Programs in public institutions						
CPC=2	-0.378	-0.381	1.095***	0.005	-0.767**	-0.733*
	(0.538)	(0.503)	(0.378)	(0.321)	(0.332)	(0.394)
CPC=3	-0.217**	-0.145	-0.048	-0.038	-0.009	0.082
	(0.088)	(0.102)	(0.081)	(0.061)	(0.108)	(0.138)
CPC=4	-0.021	0.011	-0.158**	0.048	0.122**	-0.185**
	(0.055)	(0.036)	(0.065)	(0.033)	(0.046)	(0.072)
CPC=5	0.068	-0.202	0.045	-0.069	-0.136	0.107
	(0.098)	(0.119)	(0.113)	(0.101)	(0.097)	(0.164)
n	1,401	1,800	784	3,675	1,295	873

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. The CPC in t+3 is continuous and range from 0 to 5. The regressions are weighted by the number of senior students taking the ENADE exam. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Table 14: The impact of accountability on the total number of slots, by field of study

	(1)	(2)	(3)	(4)	(5)	(6)
	Engineering and related	Social sciences, business and law	Health	Education	Sciences, math and computation	Others
All sample						
CPC=2	0.239	-0.320	-0.389	0.377**	0.506	0.058
	(0.185)	(0.236)	(0.412)	(0.179)	(0.302)	(0.499)
CPC=3	0.132*	0.160**	0.201**	-0.012	0.022	0.099
	(0.074)	(0.072)	(0.084)	(0.103)	(0.074)	(0.101)
CPC=4	-0.012	-0.077	0.132**	0.074	-0.029	-0.000
	(0.066)	(0.062)	(0.057)	(0.066)	(0.069)	(0.070)
CPC=5	0.091	-0.068	-0.014	0.180	-0.005	0.131
	(0.096)	(0.075)	(0.069)	(0.141)	(0.114)	(0.118)
n	4,374	13,048	4,228	9,197	4,510	2,504
Only private institutions						
CPC=2	0.442**	-0.440*	-0.549	0.503	0.186	0.239
	(0.166)	(0.255)	(0.448)	(0.533)	(0.333)	(0.729)
CPC=3	0.166**	0.201**	0.215*	0.007	0.023	0.064
	(0.072)	(0.081)	(0.105)	(0.087)	(0.068)	(0.084)
CPC=4	-0.021	-0.077	0.189**	0.226***	-0.070	0.053
	(0.077)	(0.071)	(0.073)	(0.042)	(0.082)	(0.082)
CPC=5	0.233	-0.087	0.095	0.279***	-0.098	-0.524
	(0.199)	(0.128)	(0.155)	(0.045)	(0.222)	(0.346)
n	2,821	11,144	3,391	5,172	3,159	1,529
Only public insitutions						
CPC=2	-2.298***	-0.230	-0.294	0.304	1.005**	0.700
	(0.554)	(0.591)	(0.381)	(0.207)	(0.415)	(0.609)
CPC=3	-0.204	-0.077	0.069	-0.060	-0.008	-0.069
	(0.167)	(0.076)	(0.189)	(0.142)	(0.131)	(0.147)
CPC=4	-0.103	-0.020	0.050	-0.071	-0.084	0.016
	(0.082)	(0.070)	(0.077)	(0.088)	(0.081)	(0.081)
CPC=5	0.118	-0.028	-0.025	0.186	0.180*	0.234*
	(0.099)	(0.117)	(0.131)	(0.177)	(0.094)	(0.114)
n	1,553	1,904	837	4,025	1,351	975

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. The dependent variable is the sum of new slots over the period t+1 to t+3 and is measured in logarithm form. The regressions are weighted by the total number of students enrolled in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

6 Conclusions

Higher education accountability is meant to provide information about the quality of undergraduate programs and to support public regulation of HEIs. Our results suggest that SINAES impacts Brazilian HEIs in the years following the publication of results, mainly for HEIs around the cutoff that determines the minimum level required for federal approval of the program. The programs that receive a low score ($CPC_{score} < 1.945$) in a certain period achieve higher quality indexes in terms of student performance, infrastructure, faculty and quality overall in the next evaluation cycle. As a result, those programs also obtain a higher CPC_{score} 3 years after the evaluation than those programs that are just above the threshold ($CPC_{score} \geq 1.945$). On the other hand, programs just above this threshold increase the number of slots available, receive more applications and admit more new students than programs just below the same threshold. These results suggest that program administrators respond to the threat of punishment related to this threshold.

Even though we expected administrators to use their results as an advertisement when programs achieved higher scores (i.e., CPC_{level}=4 or 5), we do not find consistent impacts from reaching this level on either program effort or candidate perceptions of future returns.

In addition to identifying impacts mainly around $CPC_{score} = 1.945$, our main results are stronger for private HEIs, which we argue are related to the competitive pressure and positive incentives (such as access to public programs that offer scholarships and student loans) that private institutions face.

Although we discuss the potential mechanisms that explain administrators' behavioral changes, questions related to how society at large reacts to evaluation results remain unanswered. For example, how do candidates for higher education use accountability scores to decide which program to attend? Why do students still decide to attend programs with low scores? On the other hand, do employers take into account the quality of undergraduate programs when selecting prospective employees? More research is needed to answer these and other questions regarding how different agents respond to higher education accountability and to contribute to our understanding of the effects of this evaluation policy.

References

Bowman, N. A. and M. N. Bastedo (2009). Getting on the front page: organizational reputation, status signals, and the impact of US News and World

Report on student decisions. *Research in Higher Education 50*(5), 415–436.

Bowman, N. A. and M. N. Bastedo (2013). Anchoring effects in world university rankings: exploring biases in reputation scores. In *Higher Education in the Global Age*, pp. 271–287. Routledge.

Camargo, B., R. Camelo, S. Firpo, and V. Ponczek (2018). Information, market incentives, and student performance evidence from a regression discontinuity design in Brazil. *Journal of Human Resources 53*(2), 414–444.

Canaan, S. and P. Mouganie (2018). Returns to education quality for low-skilled students: evidence from a discontinuity. *Journal of Labor Economics 36*(2), 395–436.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2019). *A practical introduction to regression discontinuity designs: foundations*. Cambridge University Press.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association 115*(531), 1449–1455.

Cattaneo, M. D., R. Titiunik, and G. Vazquez-Bare (2020). Analysis of regression-discontinuity designs with multiple cutoffs or multiple scores. *The Stata Journal 20*(4), 866–891.

Chakrabarti, R. (2014). Incentives and responses under No Child Left Behind: credible threats and the role of competition. *Journal of Public Economics 110*, 124–146.

Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics 93*(9-10), 1045–1057.

Corbucci, P. R., L. C. Kubota, and A. P. B. Meira (2016). Evolução da educação superior privadano Brasil: da reforma universitário de 1968 à década de 2010. *Radar: tecnologia, produção e comércio exterior* (46), 7–12.

Craig, S. G., S. A. Imberman, and A. Perdue (2013). Does it pay to get an A? School resource allocations in response to accountability ratings. *Journal of urban Economics 73*(1), 30–42.

Craig, S. G., S. A. Imberman, and A. Perdue (2015). Do administrators respond to their accountability ratings? The response of school budgets to accountability grades. *Economics of Education Review 49*, 55–68.

De Hoyos, R., V. A. Garcia-Moreno, and H. A. Patrinos (2017). The impact of an accountability intervention with diagnostic feedback: evidence from Mexico. *Economics of Education Review 58*, 123–140.

Deming, D. J., S. Cohodes, J. Jennings, and C. Jencks (2016). School accountability, postsecondary attainment, and earnings. *Review of Economics and Statistics 98*(5), 848–862.

Deming, D. J. and D. Figlio (2016). Accountability in us education: applying lessons from K-12 experience to higher education. *Journal of Economic Perspectives 30*(3), 33–56.

Fan, J., N. E. Heckman, and M. P. Wand (1995). Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. *Journal of the American Statistical Association 90*(429), 141–150.

Feng, L., D. Figlio, and T. Sass (2018). School accountability and teacher mobility. *Journal of Urban Economics 103*, 1–17.

Figlio, D. N. and C. E. Rouse (2006). Do accountability and voucher threats improve low-performing schools? *Journal of Public Economics 90*(1-2), 239–255.

Figlio, D. N. and J. Winicki (2005). Food for thought: the effects of school accountability plans on school nutrition. *Journal of public Economics 89*(2-3), 381–394.

Hastings, J. S. and J. M. Weinstein (2008). Information, school choice, and academic achievement: evidence from two experiments. *The Quarterly journal of economics 123*(4), 1373–1414.

Holbein, J. B. and H. F. Ladd (2017). Accountability pressure: regression discontinuity estimates of how No Child Left Behind influenced student behavior. *Economics of Education Review 58*, 55–67.

Inep (2009). *Sistema Nacional de Avaliação da Educação Superior: da concepção à regulamentação* (5 ed.). Brasilia: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of economic literature 48*(2), 281–355.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: a density test. *Journal of econometrics 142*(2), 698–714.

Mizala, A. and M. Urquiola (2013). School markets: the impact of information approximating schools' effectiveness. *Journal of Development Economics 103*, 313–335.

Neal, D. and D. W. Schanzenbach (2010). Left behind by design:pproficiency counts and test-based accountability. *The Review of Economics and Statistics 92*(2), 263–283.

Nunes, L. C., A. B. Reis, and C. Seabra (2015). The publication of school rankings: a step toward increased accountability? *Economics of Education Review 49*, 15–23.

OECD (2018). *Rethinking Quality Assurance for Higher Education in Brazil, Reviews of National Policies for Education.* Paris: OECD Publishing.

Rezende, M. (2010). The effects of accountability on higher education. *Economics of Education Review 29*(5), 842–856.

Rockoff, J. and L. J. Turner (2010). Short-run impacts of accountability on school quality. *American Economic Journal: Economic Policy 2*(4), 119–47.

Rouse, C. E., J. Hannaway, D. Goldhaber, and D. Figlio (2013). Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure. *American Economic Journal: Economic Policy 5*(2), 251–81.

Woo, S., S. Lee, and K. Kim (2015). Carrot and stick? Impact of a low-stakes school accountability program on student achievement. *Economics Letters 137*, 195–199.

Appendices

A The use of the higher education accountability results as advertisement

Figure 11: Example of the CPC level used as positive advertising

Note: The name on the building's facade should be IESB, the name of the institution, but managers replace the letter "S" with the number 5 to advertise their performance in the evaluation.

B Robustness check on the number of students taking the ENADE exam

Table B.1: Accountability and the number of students taking the ENADE

	(1)	(2)
	t=0	t=3
CPC_{level}=2	0.015	-0.028
	(0.012)	(0.028)
CPC_{level}=3	-0.004	-0.003
	(0.003)	(0.003)
CPC_{level}=4	-0.001	0.001
	(0.002)	(0.002)
CPC_{level}=5	-0.002	-0.010
	(0.004)	(0.008)
n	29,087	29,087

*** p<0.01, ** p<0.05, * p<0.1
Notes: The dependent variable is the the ratio of the number of students taking the ENADE to the total program enrollment. Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Regressions are weighted by total enrollments. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

C Robustness check varying the polynomial order

Table C.1: The impact of accountability on program quality varying the order of the polynomial

	(1)	(2)	(3)	(4)	(5)	(6)
	ENADE	Infrastructure	Teaching and learning	Opportunity	IDD	CPC
Cubic polynomial						
CPC_{level}=2	-0.263	-0.258	-0.201	-0.288	-0.216	-0.163
	(0.175)	(0.203)	(0.214)	(0.287)	(0.212)	(0.110)
CPC_{level}=3	-0.188***	-0.324***	-0.134**	-0.154***	-0.139***	-0.143***
	(0.063)	(0.079)	(0.052)	(0.055)	(0.040)	(0.017)
CPC_{level}=4	-0.036	0.014	-0.019	0.060	0.019	-0.013
	(0.029)	(0.057)	(0.065)	(0.054)	(0.027)	(0.027)
CPC_{level}=5	-0.110	0.167	0.183*	0.196**	-0.048	0.054
	(0.098)	(0.129)	(0.106)	(0.073)	(0.078)	(0.058)
n	34,405	35,052	35,052	29,640	33,637	33,437
Quadratic polynomial						
CPC_{level}=2	-0.125	-0.199	-0.270*	-0.348**	-0.060	-0.117
	(0.134)	(0.177)	(0.152)	(0.157)	(0.191)	(0.080)
CPC_{level}=3	-0.181**	-0.326***	-0.138**	-0.169**	-0.128***	-0.144***
	(0.067)	(0.086)	(0.057)	(0.066)	(0.034)	(0.019)
CPC_{level}=4	-0.052	0.012	-0.016	0.073	0.001	-0.017
	(0.032)	(0.046)	(0.051)	(0.045)	(0.024)	(0.031)
CPC_{level}=5	-0.021	0.185	0.151	0.135*	0.056	0.080
	(0.071)	(0.115)	(0.119)	(0.078)	(0.100)	(0.051)
n	34,405	35,052	35,052	29,640	33,637	33,437

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Quality indicators are continuous and range from 0 to 5. The regressions are weighted by the number of senior students taking the ENADE exam. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Table C.2: The impact of accountability on program status and flow indicators varying the order of the polynomial

	(1)	(2)	(3)	(4)	(5)	(6)
	Slots	Applications	New students	Enrollment	Dropout	Acitivity status
Cubic polynomial						
CPC=2	0.102	0.056	0.057	0.013	-0.043	0.025
	(0.116)	(0.195)	(0.144)	(0.124)	(0.162)	(0.032)
CPC=3	0.129***	0.132***	0.113***	0.057	0.020	0.002
	(0.036)	(0.039)	(0.034)	(0.037)	(0.073)	(0.004)
CPC=4	-0.007	-0.034	-0.038	-0.058**	0.030	0.007*
	(0.031)	(0.041)	(0.029)	(0.026)	(0.080)	(0.004)
CPC=5	0.025	-0.003	0.047	0.056	-0.026	0.011
	(0.038)	(0.085)	(0.053)	(0.066)	(0.087)	(0.010)
n	37,861	37,484	37,277	38,819	37,115	32,978
Quadratic polynomial						
CPC=2	-0.056	-0.084	-0.113	-0.071	0.002	0.019
	(0.098)	(0.161)	(0.181)	(0.154)	(0.162)	(0.032)
CPC=3	0.132***	0.131***	0.114***	0.058	0.021	0.002
	(0.039)	(0.037)	(0.037)	(0.039)	(0.075)	(0.004)
CPC=4	-0.006	-0.027	-0.034	-0.056**	0.026	0.007**
	(0.029)	(0.035)	(0.024)	(0.021)	(0.075)	(0.003)
CPC=5	-0.001	-0.044	0.008	0.038	0.011	0.008
	(0.058)	(0.070)	(0.041)	(0.052)	(0.062)	(0.008)
n	37,861	37,484	37,277	38,819	37,115	32,978

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Columns (1)-(3) refer to the sum of new slots, applications and new students over the period t+1 to t+3. Columns (4)-(6) refer to enrollment, the dropout rate and the activity status in t+3. Outcome variables such as new slots, applications, new students and enrollment are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. The regressions are weighted by the total number of students enrolled in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

Table C.3: The impact of accountability on program faculty profile varying the order of the polynomial

	(1)	(2)	(3)	(4)	(5)
	Students by Faculty	Faculty	MA	PhD	Full-time
Cubic polynomial					
CPC=2	3.306	-0.076	-0.015	-0.029	-0.027
	(5.137)	(0.128)	(0.025)	(0.046)	(0.056)
CPC=3	0.811	0.037	0.003	-0.024***	-0.033***
	(0.490)	(0.033)	(0.008)	(0.006)	(0.006)
CPC=4	-0.686	0.020	0.006	-0.017**	0.001
	(0.526)	(0.035)	(0.006)	(0.008)	(0.008)
CPC=5	1.364*	0.017	-0.018	0.032	0.018*
	(0.664)	(0.068)	(0.021)	(0.019)	(0.010)
n	25,043	25,117	34,804	34,804	34,804
Quadratic polynomial					
CPC=2	1.223	-0.001	-0.039*	-0.005	-0.053
	(3.477)	(0.121)	(0.019)	(0.037)	(0.067)
CPC=3	0.709	0.041	0.005	-0.026***	-0.034***
	(0.478)	(0.034)	(0.007)	(0.005)	(0.005)
CPC=4	-0.430	0.004	0.004	-0.015*	0.003
	(0.458)	(0.031)	(0.007)	(0.008)	(0.008)
CPC=5	0.097	0.083*	-0.016	0.030**	0.007
	(1.105)	(0.046)	(0.012)	(0.012)	(0.010)
n	25,043	25,117	34,804	34,804	34,804

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Faculty variables are measured in t+3. Students/faculty is the ratio between the number of students and the number of professors associated with the program. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. Regressions are weighted by the number of enrollments in the year of evaluation. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). Robust standard errors –in parentheses – are clustered at the state level.

D Robustness check applying the method of Cattaneo et al. (2020)

Table D.1: The impact of accountability on program quality applying the method of Cattaneo et al. (2020)

	(1)	(2)	(3)	(4)	(5)	(6)
	ENADE	Infrastructure	Teaching and learning	Opportunity	IDD	CPC
CPC=2	0.994	0.675	0.327	-5.460***	0.185	-0.285
	(0.627)	(0.966)	(0.666)	(1.362)	(0.948)	(0.616)
CPC=3	-0.145**	-0.170**	-0.222**	-0.233**	-0.102	-0.168***
	(0.064)	(0.076)	(0.095)	(0.105)	(0.073)	(0.055)
CPC=4	-0.006	-0.037	0.004	-0.040	-0.062	-0.038
	(0.058)	(0.066)	(0.060)	(0.088)	(0.080)	(0.041)
CPC=5	-0.267	0.368**	0.241	0.180	-0.121	-0.080
	(0.173)	(0.153)	(0.215)	(0.249)	(0.291)	(0.122)

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Quality indicators are continuous and range from 0 to 5. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). The table shows cutoff-specific treatment effects based on local polynomial estimation and robust bias-corrected inference procedures, following Cattaneo et al. (2020). Robust standard errors –in parentheses – are clustered at the state level.

Table D.2: The impact of accountability on program status and flow indicators applying the method of Cattaneo et al. (2020)

	(1)	(2)	(3)	(4)	(5)	(6)
	Slots	Applications	New students	Enrollments	Dropout	Closure situation
CPC=2	0.419	0.782**	1.602***	1.745***	-0.371	0.629***
	(0.282)	(0.356)	(0.539)	(0.572)	(0.305)	(0.219)
CPC=3	0.112	0.119	-0.097	-0.143	0.081	0.025
	(0.085)	(0.146)	(0.144)	(0.121)	(0.195)	(0.024)
CPC=4	-0.100*	-0.142*	-0.110*	-0.089	-0.009	-0.002
	(0.060)	(0.082)	(0.058)	(0.079)	(0.184)	(0.015)
CPC=5	-0.209*	-0.429**	-0.184	-0.274**	0.232	-0.052
	(0.115)	(0.177)	(0.192)	(0.139)	(0.319)	(0.062)

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Columns (1)-(3) refer to the sum of new slots, applications and new students over the period t+1 to t+3. Columns (4)-(6) refer to enrollment, the dropout rate and the activity status in t+3. Outcome variables such as new slots, applications, new students and enrollment are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). The table shows cutoff-specific treatment effects based on local polynomial estimation and robust bias-corrected inference procedures, following Cattaneo et al. (2020). Robust standard errors –in parentheses – are clustered at the state level.

Table D.3: The impact of accountability on program faculty profile applying the method of Cattaneo et al. (2020)

	(1)	(2)	(3)	(4)	(5)
	Students by Faculty	Faculty	MA	PhD	Full-time
CPC=2	22.148***	0.293	0.578**	-0.654	0.587
	(4.703)	(0.267)	(0.287)	(0.564)	(0.376)
CPC=3	-9.067	0.008	-0.143	0.033	0.010
	(8.408)	(0.465)	(0.513)	(0.093)	(0.068)
CPC=4	1.080	-0.071	-0.083	-0.092	-0.046
	(0.858)	(0.063)	(0.064)	(0.072)	(0.061)
CPC=5	-1.794	0.030	-0.002	0.113	-0.000
	(1.507)	(0.158)	(0.160)	(0.194)	(0.158)

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the programs to the left of the cutoff and the programs to the right of the same cutoff. Faculty variables are measured in t+3. Students/faculty is the ratio between the number of students and the number of professors associated with the program. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. All specifications include a vector of program characteristics control variables (the number of programs within the same field of study in the same HEI and dummies for type of administration (public or private), the year of evaluation, state, field of study, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on continuous indexes used to define the CPC_{level} (i.e., CPC_{score}, ENADE, IDD, infrastructure, teaching and learning, and faculty characteristics). The table shows cutoff-specific treatment effects based on local polynomial estimation and robust bias-corrected inference procedures, following Cattaneo et al. (2020). Robust standard errors –in parentheses – are clustered at the state level.

E Institution quality

This section adapts equation 2 to estimate the impacts of the higher education institution evaluation on institutional outcomes. As mentioned before, the quality of the institution is summarized in the IGC score, which follows the same rule as CPC for classifying courses from levels 1 to 5. The IGC level is updated every year based on the results of the courses evaluated in the same year and the results from courses evaluated over the last 2 years.

Thus, we run the following reduced-form regression specification:

$$Y_{jt+3} = \alpha + \lambda_L \mathbf{IGC}^L_{jt} + \beta \boldsymbol{f}(\mathbf{Q}_{jt}) + \gamma \mathbf{C}_{jt} + \varepsilon_{jt} \tag{3}$$

where Y_{jt+3} is the dependent variable for higher education institution j in year t such as the following IGC_{score}, ENADE score, MA and PhD grades –as evaluated by Capes –; $\mathbf{IGC}^{L}_{jt}$ is a dummy indicating whether an institution is at IGC level L or below that based on IGC score achieved in $t = 0$. $\mathbf{Q}_{jt}$ is a quartic polynomial on the IGC score. $\mathbf{C}_{jt}$ is a vector of covariates for institutional characteristics, such as the number of courses, a dummy for type of administration (whether public or private), dummies for years of evaluation, and state dummies; and ε_{jt} is the idiosyncratic error term.

Tables E.1, E.3 and E.2 display the results for equation 3. It is evident that there are no clear impacts of IGC level on either outcome. This indicates that accountability has stronger impacts at the course level, which is expected since it is somewhat rare to find an entire Higher Education Institution that is below the minimum threshold of recognition.

Table E.1: Impacts of HEI score on instituion quality

	(1)	(2)	(3)	(4)
	IGC	ENADE	Master	Doctorate
IGC_{level}=2	0.008	0.033	0.096*	0.176
	(0.132)	(0.130)	(0.054)	(0.122)
IGC_{level}=3	-0.020	-0.029**	-0.025	0.018
	(0.012)	(0.011)	(0.046)	(0.043)
IGC_{level}=4	-0.001	0.002	0.037*	0.044
	(0.015)	(0.016)	(0.019)	(0.044)
IGC_{level}=5	0.006	0.038	-0.031	0.062
	(0.024)	(0.049)	(0.023)	(0.046)
n	15,526	15,548	15,627	15,627

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the HEIs to the left of the cutoff and the HEIs to the right of the same cutoff. Quality indexes are continuous measures and range from 0 to 5. The regressions are weighted by the number of senior students taking the ENADE exam. All specifications include a vector of control variables for HEI characteristics (dummies for type of administration (public or private), the year of evaluation, state, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on IGC_{score} with interactions with IGC_{level}. Robust standard errors –in parentheses – are clustered at the state level.

Table E.2: Impacts of HEI score on flow indicators

	(1)	(2)	(3)	(4)	(5)
	slots	Applications	New students	Total enrollment	Dropout
IGC_{level} =2	0.670*	0.628	0.463	0.825*	-0.308**
	(0.338)	(0.537)	(0.331)	(0.419)	(0.126)
IGC_{level}=3	0.021	0.315*	0.141	0.205	-0.074
	(0.084)	(0.166)	(0.150)	(0.130)	(0.081)
IGC_{level}=4	0.342*	0.209*	0.131	0.194	0.040
	(0.199)	(0.102)	(0.098)	(0.133)	(0.028)
IGC_{level}=5	0.420*	0.587	0.146	0.671*	0.054
	(0.241)	(0.360)	(0.287)	(0.351)	(0.066)
n	15,381	15,220	15,610	15,085	15,610

*** p<0.01, ** p<0.05, * p<0.1

Notes: Each coefficient is interpreted as the difference in performance between the HEIs to the left of the cutoff and the HEIs to the right of the same cutoff. Columns (1)-(3) refer to the sum of new slots, applications and new students over the period t+1 to t+3. Columns (4)-(6) refer to enrollment, the dropout rate and the activity status in t+3. Outcome variables such as new slots, applications, new students and enrollment are in logarithm form. Dropout is the ratio of students who leave the program to the total enrollment. Activity status indicates whether programs are still working in the following years. The regressions are weighted by the total number of students enrolled in the year of evaluation. All specifications include a vector of control variables for HEI characteristics (dummies for type of administration (public or private), the year of evaluation, state, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on IGC_{score} with interactions with IGC_{level}. Robust standard errors –in parentheses – are clustered at the state level.

Table E.3: Impacts of HEI score on instituion faculty profile

	(1)	(2)	(3)	(4)	(5)
	Students/ faculty	Faculty	MA	PhD	Full- time
IGC_{level}=2	15.702	0.236	-0.051	-0.088*	-0.276**
	(13.756)	(0.242)	(0.055)	(0.050)	(0.101)
IGC_{level}=3	-0.316	0.039	-0.022**	0.009	0.033
	(2.142)	(0.056)	(0.010)	(0.013)	(0.021)
IGC_{level}=4	3.667**	0.071	0.001	-0.020	0.008
	(1.545)	(0.052)	(0.015)	(0.022)	(0.040)
IGC_{level}=5	-1.035	-0.027	0.013	-0.010	0.044
	(2.275)	(0.109)	(0.015)	(0.020)	(0.047)
n	15,615	15,615	15,615	15,615	15,615

*** p<0.01, ** p<0.05, * p<0.1

Notes:Each coefficient is interpreted as the difference in performance between the HEIs to the left of the cutoff and the HEIs to the right of the same cutoff. Faculty variables are measured in t+3. Students/faculty is the ratio between the number of students and the number of professors associated with the HEI. Faculty is the number of faculty members and is in logarithm form. MA, PhD and Full-time refer respectively to the percentage of faculty members with a master degree, percentage of faculty members with a doctoral degree, and percentage of faculty members in a full-time working contract. Regressions are weighted by the number of enrollments in the year of evaluation. All specifications include a vector of control variables for HEI characteristics (dummies for type of administration (public or private), the year of evaluation, state, type of academic organization, and the IGC level of the institution). Regressions also include a quartic polynomial on IGC_{score} with interactions with IGC_{level}. Robust standard errors –in parentheses – are clustered at the state level.