From: Evan Riehl

Subject: "Do standardized college admission exams reflect or create inequality? Evidence from Brazil" (extended abstract)

Date: December 15, 2019

Motivation and summary

There is a growing backlash against the use of standardized exams in college admissions. Critics argue that wealthy students have greater access to test preparation services, and thus gain a leg up when standardized exams are used in admissions. Such criticisms have led several high-profile U.S. colleges—including the University of Chicago—to adopt test-optional admissions in recent years. Defenders of admission exams, however, contend that test score gaps mainly reflect inequities at lower schooling levels, and that other admission systems would be even more unequal. It is difficult to settle this debate without observing the gaps in college preparation that would exist under alternate admission methods.

This paper exploits a unique natural experiment in Brazil to examine how the use of standardized exams in college admissions affects access to top universities. From 2010–2015, Brazil's system of highly-selective federal universities transitioned from institution-specific admission exams to a national standardized test called the ENEM. Since the ENEM exam was also used for high school accountability, many high schools required their students to take the exam regardless of its role in college admissions.

This setting allows us to focus on comparable groups of students and ask how test score gaps changed as the ENEM's role in college admission grew. We use individual-level administrative data and a differences-in-differences design that exploits geographic variation in the rollout of the ENEM. The main result is that both racial and socioeconomic test score gaps increased by about 10 percent when federal universities began to use the ENEM for college admissions. This shows that college admission exams can overstate the gaps in academic preparation between demographic groups that would exist in lower-stakes exams. Future work on this project will use data on individual exam questions to explore the mechanisms underlying these effects, and examine the predictive power of these test score effects for students' college outcomes (Rothstein, 2004; Bettinger et al., 2013).

Details

This paper analyzes the rollout of a national standardized college admission exam in Brazil called the ENEM (*Exame Nacional do Ensino Médio*). The ENEM began in 1998 as a 63-question, multi-disciplinary test designed to evaluate high school quality. In 2009, the federal government expanded the exam so that it could also serve as a tool for college admissions. The new ENEM featured 180 questions in four distinct subject groups (math, language arts, natural sciences, and social sciences), and required two separate days of testing to complete. In its present form, the ENEM exam is taken by over five million students each November.

The ENEM transition was part of a broader push to centralize university admissions. The most prestigious colleges in Brazil are the *federal universities*, which are public, tuition-free, and highly selective. There are 63 federal universities distributed across the 27 states of Brazil. Prior to 2009, admission to federal universities was typically offered to students who obtained the highest scores on institution-specific admission tests (*vestibular* exams). The expansion of the ENEM exam was motivated in part by a desire to standardize and centralize federal university admissions.

This paper exploits two features of the ENEM transition to isolate the effect of standardized exams on college access. First, federal universities varied in the timing at which they switched from institution-specific admission tests to the EXAM exam. Figure 1 shows that roughly one-third of federal universities began using the ENEM exam right away for admission to the 2010 cohort, while others adopted the exam over the next five years. Although the ENEM may have increased geographic mobility (Machado and Szerman, 2017), roughly 80 percent of federal university students still attend an institution in their home state. Thus the staggered adoption of the ENEM by federal universities created geographic variation in the exam's stakes.

Second, because the ENEM exam was originally designed for high school accountability, many high schools required their students to take the test regardless of its role in college admissions. We identify a set of over 2,000 high schools where nearly all graduating seniors took the ENEM exam in each year from 2009–2017. This *high school graduate sample* allows us to examine how test score gaps changed in comparable groups of students as the stakes of the ENEM increased.

Our analysis uses three individual-level administrative data sources:

- A census of all high school graduating seniors in Brazil from 2009–2017;
- Scores for all ENEM test takers from 2009–2017; and
- A census of all Brazilian college enrollees from 2009–2018.

Our main empirical specification is the simple differences-in-difference regression:

$$y_{ist} = \gamma_s + \gamma_t + \beta \text{ENEM}_{st} + \epsilon_{ist} \tag{1}$$

where y_{ist} is an outcome for individual i who graduated from high school in state s and year t. Equation (1) includes state fixed effects, γ_s, year fixed effects, γ_t, and the variable of interest, ENEM_{st}, which measures whether federal universities in state s and year t had adopted the ENEM exam in admissions. We use two different versions of ENEM_{st}: 1) a continuous variable defined as the fraction of federal universities enrollees who entered through the ENEM exam in each state-year; and 2) a binary variable that measures the year in which the fraction of ENEM admits changed most sharply in each state. The binary treatment variable works because most universities that switched to ENEM admissions did so in a single year. Figure 2 shows that, on average, the fraction of federal universities enrollees in a given state who were admitted through the ENEM increased sharply by about 50 percentage points in the first year that we define $\text{ENEM}_{st} = 1$.

We first show that when federal universities switched to ENEM admissions, the *total* number of exam takers in that state increased sharply, but there were no changes in the size or composition of our *high school graduate sample*. Table 1 shows that a 50 percentage point increase in the number of federal university ENEM slots led to a 10 percent increase in the number of ENEM takers in that state. This is consistent with more students opting to take the ENEM exam when it had higher stakes. Within our high school graduate sample, however, the ENEM adoption by federal universities is unrelated to the number of ENEM exam takers in that state. This is consistent with these high schools requiring their graduating seniors to take the exam regardless of its college admissions role.

Lastly, we use the high school graduate sample to examine how an increase in the stakes of the ENEM exam affected test score gaps between demographic groups. The main result is that

ENEM adoption by nearby federal universities caused racial and socioeconomic test score gaps to expand. For example, Figure 3 shows that the math test score gap between white and black students increased by about 0.05 standard deviations in the years following ENEM adoption by federal universities.[1] This increase is roughly 10 percent of the pre-ENEM racial test score gap of 0.4 standard deviations. ENEM adoption also caused the math test score gap between private and public high schools students to increase by 0.12 standard deviations off a base of 1.0 standard deviations. We observe smaller and moderately significantly increases in racial and socioeconomic test score gaps on the language arts, natural sciences, and social sciences components, and we find limited effects on gender test score gaps (see Table 2).

Future work

A note to the organizers: This is an early-stage project, but we are confident in the first stage and main results described above. As a next step, we plan to extend the analysis to individual questions on the exam. This will allow us to show, for example, whether changes in test scores gaps arose from easy/hard questions, or from questions that are more/less amenable to "test prep." We plan to have a draft of the paper with all of these analyses by April.

In addition, we hope to conclude the paper by analyzing college enrollment and persistence outcomes for students who were affected by the ENEM transition. This will show whether or not the test score changes induced by ENEM adoption were predictive of students' college success. This last analysis requires some data work done on-site in Brazil, and will not likely be carried out until the summer.

References

Bettinger, E. P., B. J. Evans, and D. G. Pope (2013). Improving college performance and retention the easy way: Unpacking the ACT exam. *American Economic Journal: Economic Policy 5*(2), 26–52.

Machado, C. and C. Szerman (2017). Centralized admissions and the student-college match. SSRN working paper.

Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics 121*(1), 297–317.

[1] In Brazilian datasets, racial groups are typically defined as *branca* (white), *parda* (brown), *preta* (black), *amarela* (yellow), and *indígena* (indigenous). In our analysis we use the term "white" to refer to students who identify as *branca* and *yellow*, and we use "black" to refer to students who identify as *parda*, *preta*, and *indígena*.

Figures and tables

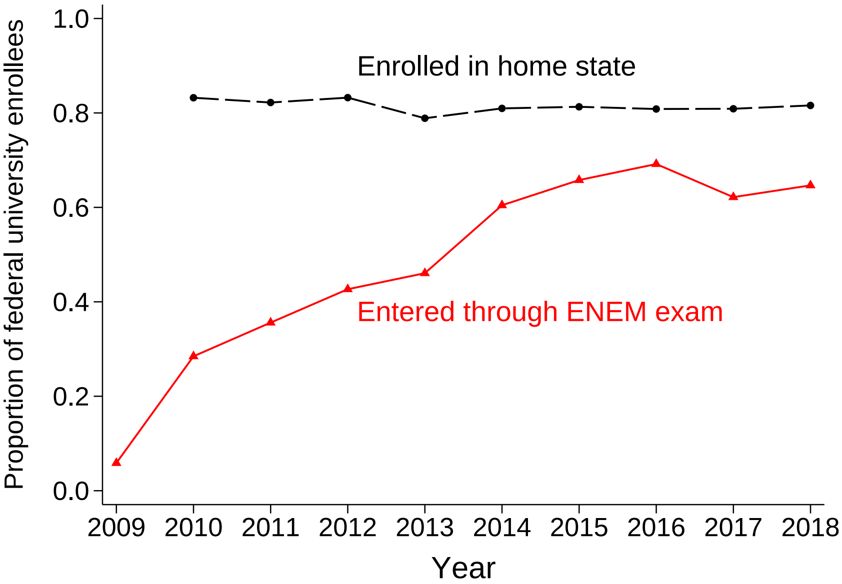


Figure 1: Characteristics of federal university enrollees by year

Notes: The sample includes all first-time enrollees in federal universities. The x-axis is the year of individuals' college enrollment. For on-time students, this is one year after they took the ENEM exam.

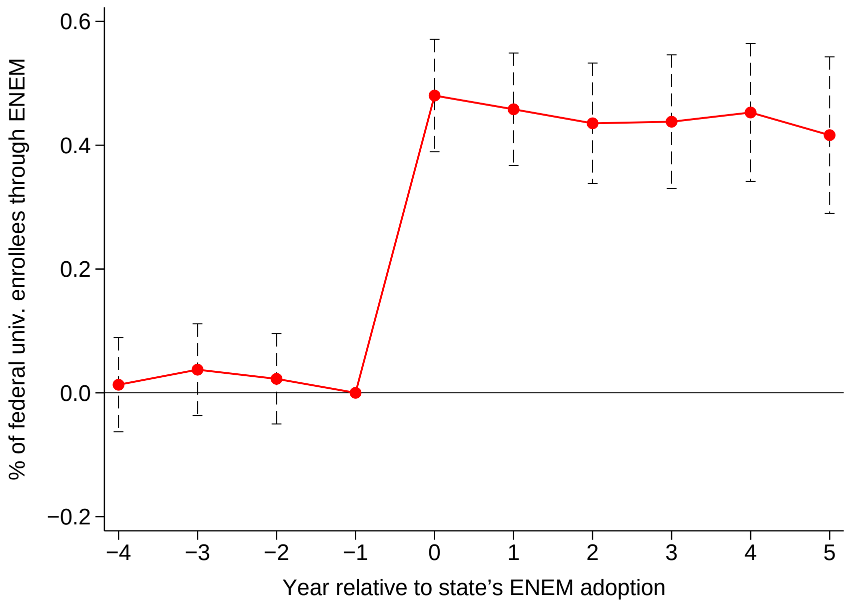
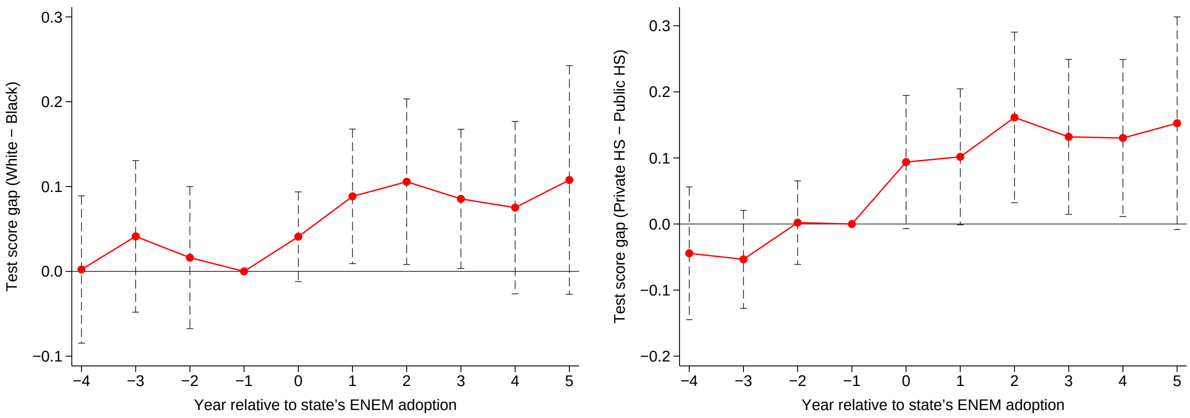


Figure 2: Event study: Proportion of federal university enrollees admitted with ENEM scores

Notes: The sample includes all first-time enrollees in federal universities. The x-axis is normalized to equal zero in the first year that $ENEM_{st} = 1$ for each state s.



Panel A. White − Black

Panel B. Private HS − Public HS

Figure 3: Event study: ENEM math test score gaps

Notes: The sample includes students in our *high school graduate* sample. The x-axis is normalized to equal zero in the first year that $ENEM_{st} = 1$ for each state s. Test score gap effects on the y-axis are in standard deviation units.

Table 1: Effects of federal university ENEM adoption on the number of exam takers

	(A)	(B)	(C)	(D)
	All exam takers		High school graduate sample	
	# exam takers (1000s)	Log # exam takers	# exam takers (1000s)	Log # exam takers
ENEM adoption	18.765*	0.103**	−0.047	0.030
	(10.072)	(0.041)	(0.089)	(0.022)
N	405	405	243	243

Notes: This table displays β coefficients from equation (1) using the binary treatment variable ENEM_{st}. Regressions are at the state-year level with robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Effects of federal university ENEM adoption on test score gaps

	(A)	(B)	(C)	(D)
	Exam subject			
	Math	Language arts	Natural science	Social science
Panel A. Test score gap in pre-ENEM cohorts (SD units)				
White − Black	0.406	0.213	0.290	0.256
Private HS − Public HS	1.037	0.561	0.814	0.726
Male − Female	0.510	−0.057	0.268	0.159
N	176,721	176,726	178,988	178,993
Panel B. Effect of ENEM adoption on test score gap (SD units)				
White − Black	0.044***	0.010	0.019**	0.020**
	(0.013)	(0.008)	(0.008)	(0.008)
Private HS − Public HS	0.117**	0.041	0.032	0.028
	(0.046)	(0.026)	(0.029)	(0.025)
Male − Female	0.021	0.012	0.004	0.004
	(0.016)	(0.008)	(0.008)	(0.010)
N	1,380,471	1,382,907	1,389,641	1,392,077

Notes: Panel A displays test score gaps by demographic group and exam subject in pre-ENEM cohorts (state-years where the binary $\text{ENEM}_{st} = 0$) using the high school graduate sample. Panel B shows the effects of ENEM adoption on these test score gaps as estimated by equation (1) with the the binary treatment variable ENEM_{st}. Parentheses contain standard errors clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$